Manual for ESHRE guideline development

Authors: N. Le Clef, S. Mcheik, Y. Lauwers, N. Vermeulen, A. Alteri, R. Anderson, A. D'Angelo, A. Feki, S. Kolibianakis, S. Lensen, Z. Veleva and B. Woodward

eshre
SCIENCE MOVING PEOPLE MOVING SCIENCE

Version 5.0 – 2026

Approved by the ESHRE ExCO April 2026

Table of Contents

Background information

Evidence-based guidelines

Evidence-based guidelines are systematically developed statements, based on the best available scientific research and developed with transparent methodology, to assist care providers and patients' decisions about appropriate care for specific circumstances (Institute of Medicine Committee on Clinical Practice Guidelines, 1992). Evidence-based guidelines are intended as neither cookbook nor textbook, but should be helpful in everyday decision-making about appropriate and effective care. Therefore, evidence-based guidelines should be presented in formats that are easy to interpret.

Aim and structure of this manual

The principal aim of this manual is to provide stepwise advice to individual members of ESHRE guideline development groups (GDG).

This manual describes the process of development and updating of the ESHRE evidence-based guidelines. This manual is based on the ESHRE manual for guideline development 2019 and draws on the most up-to-date evidence on international guideline development methodology and resources available, including Grading of Recommendations Assessment, Development, and Evaluation (GRADE) (https://www.gradeworkinggroup.org/).

GRADE: Key Points

→ GRADE offers a transparent and structured process for developing and presenting summaries of evidence, including its quality, for systematic reviews and recommendations in healthcare.

→ GRADE provides guideline developers with a comprehensive and transparent framework for carrying out the steps involved in developing recommendations.

→ GRADE's use is appropriate and helpful irrespective of the quality of the evidence: whether high or very low.

→ Although the GRADE system makes judgments about the quality of evidence and the strength of recommendations in a systematic and transparent manner, it does not eliminate the inevitable need for judgments.

In addition, the manual is based on internationally acceptable criteria of methodological quality, as articulated by the Appraisal of Guidelines for Research and Evaluation in Europe (AGREE 2) instrument (AGREE Reporting Checklist, 2016, Brouwers et al., 2016). All 23 items of the AGREE 2 Reporting checklist were incorporated in the manual and listed as tips at the end of each chapter.

The structure of this manual follows guideline development from its proposal through to publication and beyond.

This ESHRE manual is intended to be a "living" publication and will be updated regularly based on new developments in guideline development and experiences in the GDG. Comments on either content or presentation are welcome and should be sent to guidelines@eshre.eu. At the time of change, GDG members will be notified.

Previous versions

→ Manual for ESHRE guideline development v1. 2007

→ Manual for ESHRE guideline development v2. 2014

→ Manual for ESHRE guideline development v3. 2017

→ Manual for ESHRE guideline development v4. 2019

Details on the update 2026

In addition to some minor adaptations and corrections, 4 major adaptations were made in the current manual:

→ Adaptation of the methodology for the application procedure for new guidelines and the update of existing guidelines, which is now reviewed by the Guidelines Committee before approval by the Executive Committee (ExCO).

→ Adaptation of the methodology for forming a GDG, with all members and chairs selected via an application procedure and with more stringent rules on conflict of interest (COI). The composition of the GDG with review of the applications and COI is now handled by the Guidelines Committee before approval by the Executive Committee.

→ The scope and key questions will be made public for stakeholder comments for 2-4 weeks.

→ Stakeholders will need to pre-register to be able to submit comments during the stakeholder review period at the end of the guideline development process.

Abbreviations used in this document

COI	Conflict of interest
ExCO	Executive Committee
GC	Guidelines Committee
GDG	Guideline development group
Mo	Month
NGT	Nominal group technique
RS	Research Specialist
SIG	Special Interest Group
SIG SQART	SIG Safety and Quality in Assisted Reproductive Techniques
WG	Working Group

ESHRE guidelines

The main goal of ESHRE guideline development is the provision of recommendations to improve the quality of healthcare delivery within the European field of human reproduction and embryology (For more information on ESHRE visit www.eshre.eu).

ESHRE guideline development methodology is similar to the methodology of other societies (Hilbink et al., 2014, Scottish Intercollegiate Guidelines Network (SIGN), 2015 , The GRADE Working Group, 2013) and complies with the criteria used by the AGREE instrument for good quality guidelines (AGREE Reporting Checklist, 2016, Brouwers et al., 2016). Furthermore, all relevant items of the Guidelines International Network (GIN) Guideline Development Checklist were included (Schünemann et al., 2014) (https://macgrade.mcmaster.ca/resources/gin-mcmaster-guideline-development-checklist/gin-mcmaster-guideline-development-checklist/).

ESHRE evidence-based guidelines contain recommendations on several aspects of a particular clinical issue. These guidelines are based on the best available evidence (i.e., the most relevant and highest level of evidence)

and not on all evidence available. There is an explicit link between recommendations and their available evidence. Furthermore, scientific and clinical evidence takes precedence over expert judgement.

ESHRE guidelines will not include a formal analysis of cost effectiveness of recommended as compared to established practice, as this is not the main aim, and is sometimes impossible because of the obvious differences in current European economic and healthcare systems. The clinical and organisational impact of costs on recommendations will be considered in GDG meetings, and if relevant, described in the justification section. The economic feasibility of recommendations will not be covered.

ESHRE guidelines can be adapted and translated by National Societies ensuring more efficient use of resources and improvement of patient outcomes throughout Europe. ESHRE guidelines should therefore be flexible and adaptable such that individual circumstances can be taken into consideration. ESHRE has established a policy for translation of its guidelines to ensure quality and validity of translated documents.

Where terms such as “consensus”, “appropriate”, “if relevant”, “when necessary”, or “where feasible” are used, decisions are made by the Guidelines Committee or GDG based on predefined criteria, including relevance to the guideline topic, methodological considerations, and feasibility. Such decisions should be documented.

Medico-legal implications of ESHRE guidelines

Potential medico-legal implications of evidence-based guidelines have been of ongoing concern to medical practitioners (Moses and Feld, 2008). However, evidence-based guidelines are intended as an aid to judgement, not a replacement for it. The ultimate decision about a particular procedure or treatment will always depend on each individual patient’s condition, circumstances and wishes, and the judgement of the healthcare team as is represented within the disclaimer in the beginning of each guideline. Evidence-based guidelines are not intended to deprive clinicians of their medical freedom to treat, nor to relieve them of their responsibility to make appropriate decisions based on their own knowledge and experience.

To clarify the legal perspective all ESHRE guidelines carry the following statement in the disclaimer:

The European Society of Human Reproduction and Embryology (hereinafter referred to as ‘ESHRE’) developed the current clinical practice guideline, to provide clinical recommendations to improve the quality of healthcare delivery within the European field of human reproduction and embryology. This guideline represents the views of ESHRE, which were achieved after careful consideration of the scientific evidence available at the time of preparation. In the absence of scientific evidence on certain aspects, a consensus between the relevant ESHRE stakeholders has been obtained.

The aim of clinical practice guidelines is to aid healthcare professionals in everyday clinical decisions about appropriate and effective care of their patients.

However, adherence to these clinical practice guidelines does not guarantee a successful or specific outcome, nor does it establish a standard of care. Clinical practice guidelines do not replace the need for application of clinical judgment to each individual presentation, nor variations based on locality and facility type. Ultimately, healthcare professionals must make their own clinical decisions on a case-by-case basis, using their clinical judgment, knowledge, and expertise, and taking into account the condition, circumstances, and wishes of the individual patient, in consultation with that patient and/ or the guardian or carer.

ESHRE makes no warranty, express or implied, regarding the clinical practice guidelines and specifically excludes any warranties of merchantability and fitness for a particular use or purpose. ESHRE shall not be liable for direct, indirect, special, incidental, or consequential damages related to the use of the information contained herein. While ESHRE makes every effort to compile accurate information and to keep it up to date, it cannot, however, guarantee the correctness, completeness, and accuracy of the guideline in every respect. In any event, these clinical practice guidelines do not necessarily represent the views of all clinicians that are member of ESHRE.

The information provided in this document does not constitute business, medical or other professional advice, and is subject to change.

Guideline development in 12 steps

Guideline development, implementation, and evaluation are not a linear process, but a cycle of interdependent activities. Key steps within this process are: topic selection, synthesis of evidence, formulation of recommendations, consultation and review, dissemination and implementation, evaluation, and updating.

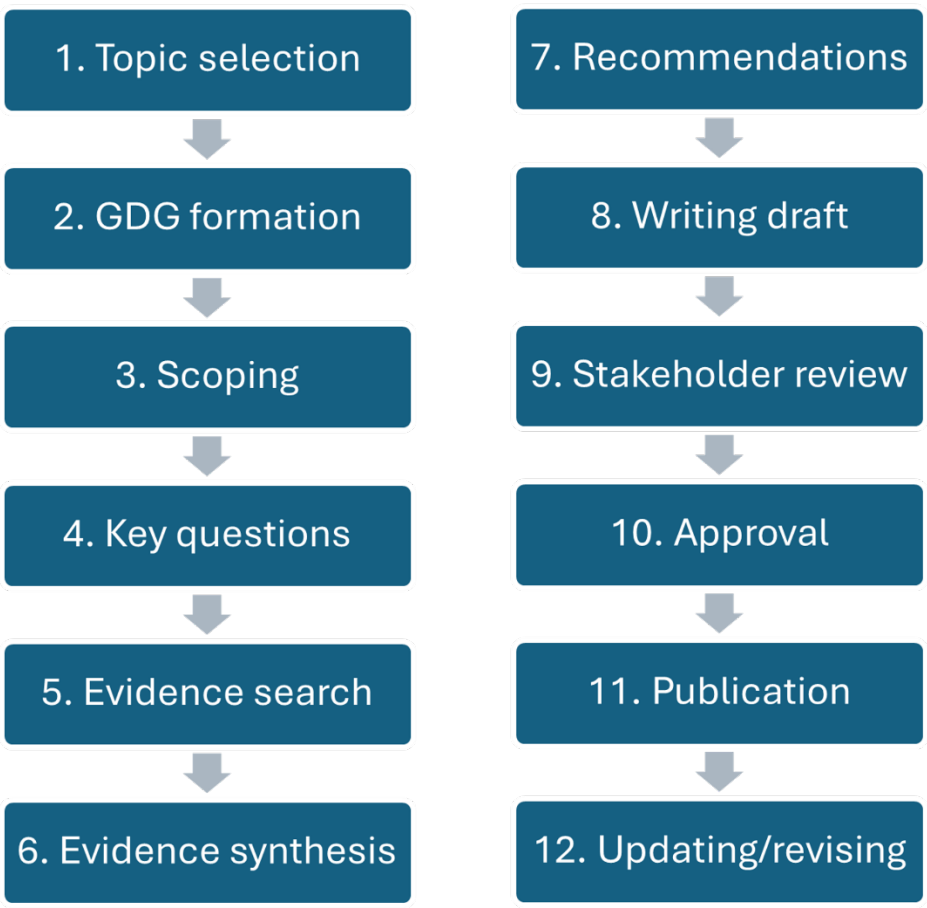


Figure 1: Overview of the 12-step guideline development process

Timelines

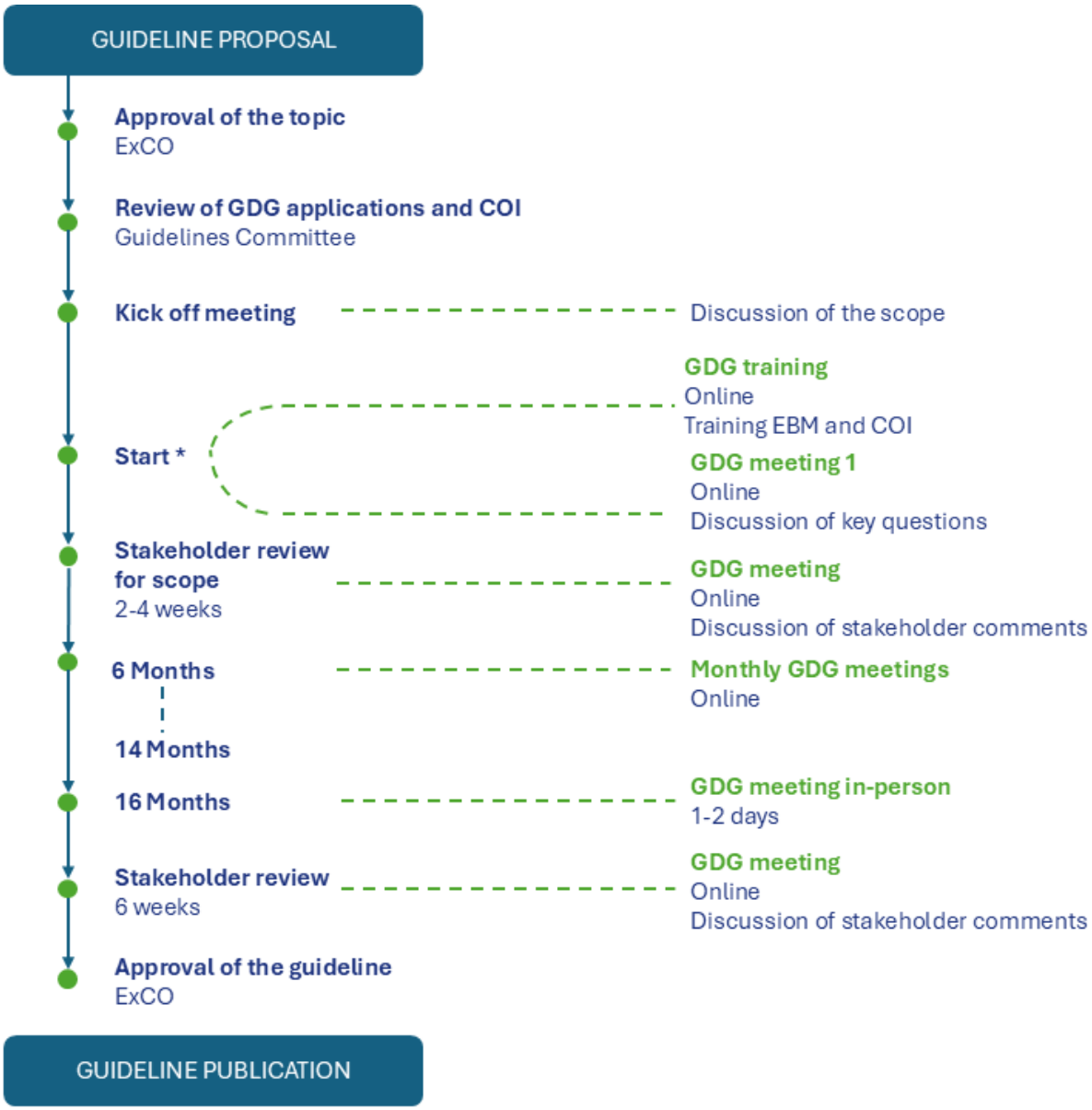
The time taken to develop an ESHRE guideline varies according to the scope of the topic, the volume of relevant literature, the amount of feedback received and the time needed to reach consensus about some topics. In general, it is recommended to keep the guideline to a reasonable size to ensure its development within an 18-24 month period from the first meeting to publication.

Budget

Most of the meetings of the GDG are virtual, with the exception of an in-person meeting during the ESHRE Annual meeting (if needed), and one in-person meeting at the end of the evidence synthesis process.

A fixed budget is set to cover the costs of necessary in-person meetings of a GDG. These expenses cover meeting costs, including travel (economy-class tickets), accommodation, food and meeting facilities. Costs are reimbursed upon request within four weeks, upon presentation of original receipts, invoices, bills, tickets etc., together with a provided ESHRE expense claim form.

Summary of meetings and timelines



** The guideline development should start as soon as a proposal is approved, but can be postponed due to other projects or workload.*

1. TOPIC SELECTION

Selection procedure

The coordinators of all Special Interest Groups (SIGs) are regularly invited to propose new guideline topics. These proposals are made on an application form (Appendix Ⓐ), concerning subjects within the field of reproductive medicine, including embryology, where a need for guidance has been identified, and aiming to assist physicians and laboratory staff in diagnosis and/or clinical management.

Individual ESHRE members wanting to present a guideline topic are encouraged to complete the application form and send it to the ESHRE research specialists (by emailing guidelines@eshre.eu). The research specialists will contact the relevant SIG coordinator for advice on the relevance of the topic for an ESHRE guideline.

ESHRE uses a 3-step selection procedure to decide which proposals for guideline development are accepted. In a first step, all proposals are checked by the ESHRE research specialists based on the following criteria:

- Is the proposal complete?
- Is the proposal clear and well organised?
- Are the details in the proposal correct?
- Is the topic appropriate for an evidence-based guideline?

The research specialist also adds information on existing guidelines and overlap with other ESHRE guidelines. To check the presence of existing guidelines the following websites are screened: Guidelines International Network (www.g-i-n.net), American Society for Reproductive Medicine (www.asrm.org), the National Institute for Clinical Excellence (www.nice.org.uk) and the Royal College of Obstetricians and Gynaecologists (www.rcog.org.uk). All guidelines found must be evaluated by using the AGREE II checklist. Adaptation of existing guidelines (rather than developing new guidelines) can be considered; methodologies for adaptation are available (Fervers et al., 2006, Darzi et al., 2017). The Cochrane Library should be consulted to estimate the available existing evidence.

If necessary, additional information is requested from the applicant to complete the proposal before submission to the ESHRE Guidelines Committee.

In a second step the ESHRE Guidelines Committee evaluates the application for guideline development and formulates an advice whether the proposal is acceptable for ESHRE guideline development. The ESHRE Guidelines Committee may suggest revisions to the application or may suggest postponing guideline development on the topic because of important ongoing research studies. The ESHRE Executive Committee revises the advice of the Guidelines Committee and decides if the proposal is acceptable for ESHRE guideline development. If not acceptable, the applicant will be informed of the reason. The decision of the Executive Committee and any comments will be communicated to the applicant by the research specialist.

Appropriate topic selection is important to ensure that an ESHRE guideline is relevant and addresses priority issues for the improvement of European reproductive medicine. Within the selection procedure priority is given on topics with:

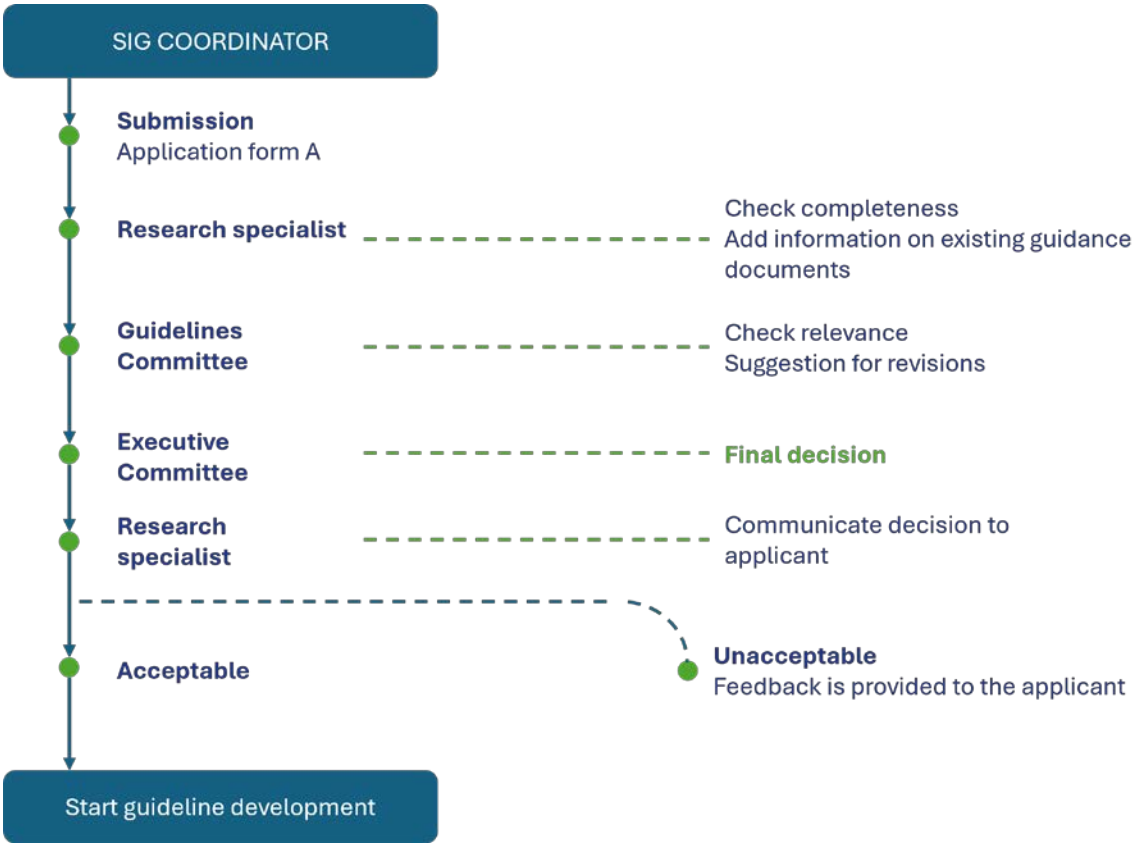
- high volume or rapidly changing evidence
- high costs
- major impact on patients (e.g. health burden or high risks)
- high practice variation (within Europe)
- high ethical/legal impact
- high improvement potential.

Application procedure

The guideline application form (Form Ⓐ) can be requested via email (guidelines@eshre.eu). Completed application forms should be sent to the ESHRE research specialists.

Proposals can be added at any time and will generally be evaluated at the next meeting of the Guidelines Committee and Executive committee.

Summary



Available forms/checklists:

Ⓐ Application form

2. FORMING GUIDELINE DEVELOPMENT GROUP

Convening an effective GDG is a crucial stage in producing a guideline; the GDG agrees on the key questions, considers the evidence and has considerable influence on the final guideline recommendations (Murphy et al., 1998). Therefore, it is strongly recommended that representatives of all key groups and disciplines affected by a guideline topic participate.

Composition of the guideline development group

Diversity is an essential feature of a GDG and its exact composition should be tailored to the guideline topic (and scope) and reflect the range of stakeholders involved. A GDG should comprise at least:

- content expert(s)
- non-expert clinician(s)
- a patient or their representatives
- allied healthcare provider(s) and
- an ESHRE research specialist.

Industry representatives are excluded from membership, ensuring no commercial influence.

A GDG should typically comprise 10 to 15 members, including the two co-chairs. Members are expected to participate in only one active[1] GDG at a time. In exceptional circumstances, participation in two active GDG may be considered, following justification and approval. Simultaneous membership of more than two active GDG is not permitted.

In composing a GDG, the following points should be considered:

- balance in geographical location; the group should reflect broad representation across Europe, including Northern, Eastern, Southern, Western, and Central regions.
- balance in gender; efforts should be made to ensure a reasonable and appropriate balance of gender within the group.
- balance in expertise; the group should include a mix of academic an non-academic members, as well as a range of seniority levels (seniors and juniors).

Depending on the guideline topic, a representative from a related society might be considered for membership of the GDG. In the case of a joint guideline development with partner organisations, the ESHRE Executive Committee must approve the collaboration (preferably at the same time as the application). A memorandum of understanding should be drafted and signed by both the organisations before the start of guideline development.

GDG selection procedure

When a topic is accepted for guideline development, an open call for GDG member applications should be organised. All candidates should submit a short CV, a motivation statement on why they should be included and a completed ESHRE COI disclosure form (form Ⓑ) and indicate whether they would like to be considered for a chair/co-chair position. The Guidelines Committee reviews the applications and selects the GDG members,

[1] *Active meaning from the first GDG meeting (scope) to the last GDG meeting (after stakeholder review)*

including the nomination of the chair/co-chairs, considering a balance in gender, geography, and expertise. The proposed composition is reviewed and approved by the ExCO before the GDG can be formalised.

At the start of the guideline development, all GDG members, except for patient representatives and invited experts, should be members of ESHRE.

Following ExCO ratification of the GDG composition, a formal invitation should be issued to each proposed GDG nominee. Once all invited nominees have agreed to participate, the GDG can become functional.

New members should usually not be added to the GDG once the development process has started. Additional needed expertise or the replacement of a GDG member should be discussed within the GDG and approved by the Guidelines Committee. The research specialist should ensure that new GDG members have all information on the previous steps in the guideline development and receive training similar to the rest of the GDG.

Leadership of the Guideline Development Group

The GDG is led by two co-chairs, at least one of whom must be free of any direct conflicts of interest and have no more than minor-level indirect interests, as defined in **table 2.1**. This requirement does not exclude clinicians who have a general interest in the topic through the provision of routine clinical care, nor individuals employed within publicly funded health or social care services. The co-chairs of the GDG are nominated by the Guidelines Committee after the application procedure described above.

The co-chairs should be GDG members with appropriate expertise, team-working skills, and should be respected content experts, preferably with experience in guideline development. They should also be experienced in group facilitation, maintaining constructive dynamics, identifying and resolving conflicts, remaining neutral and objective, and having methodological expertise.

The co-chairs, together with the research specialist, serve as the primary point of contact for the Guidelines Committee and are responsible for overseeing the development of the content and the timely production of the guideline.

Responsibilities of guideline development group members

To ensure that the GDG functions effectively and achieves its aims, all GDG members should engage to the following responsibilities:

- Attend all GDG meetings [2]
- Sign a statement of confidentiality at the start of the project
- Declare any COI at application (update in case of changes, and at least annually)
- Contribute to the formulation of key questions (PICO questions)
- Assess and summarise papers for a specific section of the guideline
- Write a summary of evidence and draft recommendations for a specific section of the guideline
- Participate in discussion and decision making, with acceptance and tolerance of varying viewpoints
- Approve of the final recommendations
- Support the dissemination of the guideline
- The GDG will be supported by an ESHRE research specialist who will be responsible for overall project management and organising the meetings in collaboration with the co-chairs of the guideline group. In addition, the research specialist will provide specific training on the different steps in guideline development during the first online meetings of the GDG. The aim of such training is to increase and

[2] If a GDG member cannot attend several meetings in a row, he/she may be asked to stand down by the co-chairs.

equalise the level of guideline development expertise within a GDG. Finally, the research specialist will perform the literature searches and collect all input in one guideline document.

Patient participation

Patient involvement in guideline development is important to ensure reflection of their needs, concerns and preferences, as they may have different perspectives on healthcare processes, priorities, and outcomes from those of healthcare professionals. Ideally, patients are involved starting from the scoping process (den Breejen et al., 2016). At least, one patient representative is included in the GDG and supported to actively contribute throughout the process and play a meaningful role in shaping the key questions and in formulating recommendations, ensuring that patients are placed at the center. Appropriate support- such as preparatory materials, clear explanation of methods and guidance during discussions- is provided for patient representatives to facilitate their participation. Their contributions are considered alongside other evidence and perspectives, and key patient-related considerations are documented in the guideline development process.

Patient needs and preferences should, for each guideline, at least be considered with respect to:

- information
- communication
- healthcare organisation
- financial constraints
- shared decision making, and
- self-management.

For the identification of patients' views other methodologies can also be applied, including literature search, patient (organisation) consultation e.g., by (focus group) interviews, and/or guideline review by patients or their representatives.

Handling Conflicts of Interest

Definition of a COI

A COI exists when a secondary interest – financial, professional, institutional, personal, or related to family or close associates – could reasonably be perceived to influence an individual's judgement or actions in the guideline development process.

Types of interests

ESHRE distinguishes two main types of interests: i) Direct interests which are interests personally held by the individual that are directly related to the topic of the guideline. These may be direct financial interests that can be measured by monetary units (e.g., salary, consultancy fees, honoraria, advisory roles, stock ownership, patents, royalties, personal research funding, or institutional grants on which the member is explicitly named as recipient or investigator, etc.), and direct non-financial interests that cannot be measured by monetary units (e.g., leadership roles, advocacy, authorship of topic related opinions, or strong intellectual positions, etc.). ii) Indirect interests which are not personally held by the individual and/or not directly related to the topic of the guideline but could be perceived as influencing judgement. These include indirect financial interests (e.g., institutional grants, funding received by the employer, or family- or close relatives- related financial interests, etc.) and indirect non-financial interests (e.g., institutional affiliations, professional networks, family relationships, etc.).

Disclosure

Because ESHRE aims to ensure objectivity, credibility and independence in its European guidelines and Recommendations for Good Practice, they are developed without external funding. All applicants must provide disclosure statements of all potential direct and indirect COI (form ®) on behalf of themselves and their first-

degree relatives. Disclosure would rely on self-reporting, accompanied by a CV and a formal "declaration of honour". No routine external verification is proposed, but spot checks are conducted to identify potential inconsistencies and, where necessary, seek clarification (e.g., recent publications, trial registries, other publicly available sources, etc.). The Guidelines Committee reviews the forms prior to the applicant being accepted as a nominated GDG member. The co-chairs and the majority of the GDG must be free from direct financial COI. In exceptional circumstances, participation of individuals with topic-relevant COI may be permitted following explicit assessment by the Guidelines Committee and approval by the ESHRE Executive Committee, with appropriate management measures in place.

The look-back period for reporting potential COI is 36 months for direct financial interests, 12 months for direct non-financial and indirect interests, prior to assuming GDG membership. Any known upcoming COI during the guideline development process must also be disclosed. An exception is made for publications related to the topic of the guideline, for which there is no time limit. Any interests arising during the guideline development period, or becoming relevant during this time, must be declared promptly and will be reassessed by the Guidelines Committee.

Monetary thresholds

Rather than focusing exclusively on the interests' monetary values, ESHRE follows an approach similar to that used by the American Thoracic Society (ATS) which prioritises the nature, scope and relevance of disclosed relationships to the topic of the guideline (e.g., topic of the lecture, consulting, research funding) (Schünemann et al., 2009). The Guidelines Committee assesses the weight of the COI by combining the type of the activity, its relevance to the guideline topic, the individual's role within the GDG (co-chair, member, advisor/reviewer). When appropriate, the magnitude of financial relationship is considered as a complementary factor. The assessment of the COI severity is based on a structured scale categorising interests as minor, moderate or severe. This assessment may be informed by indicative value categories (e.g., ≤€1000; €1001-€5000; €5001-€10000; etc.), together with relevance to the topic of the Recommendations for guideline and role-based considerations. In general, lower value ranges (≤€1000) are more likely to be associated with minor interests, mid-range values (€1001-€10000) with moderate interests, and higher value ranges (≥€10000) with more significant (potentially severe) interests; however, the final classification depends on the overall context, particularly the relevance of the relationship to the topic of the Recommendations for Good Practice document (e.g., a €500 payment could still be classified as more than minor if it is directly linked to the intervention under Recommendations for Good Practice document evaluation or if it reflects a pattern of repeated interactions with the same company). This approach supports a proportionate evaluation of direct financial, direct non-financial, and indirect COI recognising that these types of COI may differ in their potential to introduce bias.

Management plan

The Guidelines Committee reviews all COI declarations at the application stage to identify any potential COI relevant to the topic of the guideline. Disclosed interests are then categorised according to their likely impact on the development of the guideline typically as minor, moderate and major. All assessments and decisions are documented to ensure transparency and consistency.

Based on this assessment, the Guidelines Committee determines the appropriate management measures. These measures are defined before key methodological steps, such as nomination, finalising the scope, preparing the draft and formulating recommendations, and should be revisited throughout the development process as new interests arise or roles evolve. Management actions may include full participation, recusal from recommendations formulation, restricted participation in specific discussions, or, in case of major COI, exclusion from the nomination process.

Major COI generally leads to exclusion from the nomination process. Individuals with direct financial COI relevant to the topic of the guideline (e.g., marketing-related interests) should be excluded from authorship or leadership positions. They may instead serve in an advisory capacity in case their expertise would otherwise not be available

to the group. Where feasible, divestment or resignation from conflicted roles could be requested. Members with direct non-financial COI related to a substantial proportion (>50%) of the evidence base for a specific topic in a guideline can participate in the discussion of the evidence, however, they should be recused from the formulation of final recommendations and from voting on those recommendations. Indirect interests must also be declared for proper assessment. Where such interests are not posing a significant risk of bias, the GDG member concerned can engage in all aspects of the development process of the guideline.

Failure to disclose a relevant COI may result in exclusion from the GDG. Furthermore, industry interactions during the development process are strictly prohibited, and a confidentiality form (form ©) should be signed before appointment. Failure to sign this form could render an individual ineligible to participate.

Table 2.1. Structured COI severity assessment scale (ATS- aligned approach).

COI SEVERITY LEVEL	NATURE, SCOPE, ROLE	FINANCIAL MAGNITUDE	EXAMPLES	TYPICAL MANAGEMENT ACTIONS
Minor (Low risk of bias)	Indirect interests that are not personally held and/or not directly related to the topic of the guideline. Limited involvement (occasional, peripheral or unrelated activities) No leadership or decision-making role linked to the interest; no reasonable perception of influence on judgement.	Financial magnitude considered of limited relevance.	Occasional lectures or honoraria unrelated to the topic of the guideline; institutional grants not linked to the topic of the guideline and not involving the member personally; general academic publications without advocacy; institutional affiliations without direct stake in the recommendations; Indirect non-financial interests of a family member not relevant to the topic of the guideline.	Full participation in all GDG activities; disclosure recorded and monitored; no restrictions unless new interests arise.
Moderate (Potential risk of bias)	Direct non-financial interests personally held and directly related to the topic of the guideline and/or indirect institutional or family-related interest relevant to the topic of the guideline. Ongoing roles with partial relevance to the topic of the guideline, or involvement in substantial proportion of the evidence base. Professional or academic roles that could reasonably be perceived as influencing judgement	Financial magnitude considered alongside relevance, scope, and role rather than in isolation.	Leadership roles in professional societies linked to the topic; principal investigator of multiple included studies; strong publicly stated positions on the topic; institutional funding relevant to the topic of the guideline and known to the member (without personal financial benefit); research funding received by a close relative from entities operating in the same clinical area;	Participation permitted with restrictions; recusal from specific discussions where the interest is directly relevant; recusal from recommendation formulation and/or voting on affected topics; ongoing monitoring by the Guidelines Committee.

			employment of a spouse/first-degree relative in a company or organisation with interests related to the topic of the guideline.	
Major (High risk of bias)	Direct financial interests personally held and relevant to the topic of the guideline, strong direct non-financial interests and/or highly relevant indirect interests (institutional, employer, or family-related). Ongoing relationships with entities directly affected by the recommendations of the guideline. Leadership or highly influential role (e.g., co-chairs) with interests that could substantially influence judgement or recommendations.	Financial magnitude considered as an aggravating contextual factor when relevance is high.	Consultancy fees, speaker honoraria, or research funding from industry directly linked to the topic; patents, royalties, or stock ownership related to evaluated interventions/ tests; marketing-related or advisory roles; extensive authorship or intellectual leadership dominating a substantial proportion of the evidence base; spouse or first-degree relative employed by, holding shares in, or receiving consultancy fees from a company directly affected by the recommendations of the guideline; family ownership or financial stake in a commercial entity operating in the field of the guideline.	Exclusion from GDG membership or leadership positions; advisory role only (non-voting) if expertise is indispensable; mandatory recusal from recommendation formulation and voting; consideration of divestment or resignation from conflicted roles where feasible.

In line with the ATS-informed approach, the assessment of COI severity prioritises the nature, scope, and relevance of the interest to the topic of the guideline*, as well as the individual's role within the GDG. Financial magnitude and indicative value ranges may be considered as contextual factors to support consistency but are not used as the sole determinant of COI severity.*

Operationalisation

The research specialist, in consultation with the co-chairs, operationalises the post-appointment management actions decided by the Guidelines Committee and informs the individual concerned. In case of uncertainty or disagreement between the co-chairs and the research specialist regarding the appropriate management of a declared interest, the matter is referred to the Guidelines Committee for further consideration and decision.

The management actions should be formally presented to the group. The disclosure forms must be reviewed and validated prior to each in-person meeting and updated whenever any individual changes occur during the development process. In addition, the disclosure forms should be updated annually. It is the GDG members responsibility to declare any relevant interests at the earliest opportunity and ensure that their declaration remains accurate and up to date throughout their involvement. A final checklist (ICMJE form) should be completed at project completion before submission of the guideline summary paper.

Transparency

The disclosed COI, and the management decisions should be published alongside the guideline, with a summary presented in the methods section and a detailed version provided as supplementary material. Records should be retained long-term according to ESHRE policies.

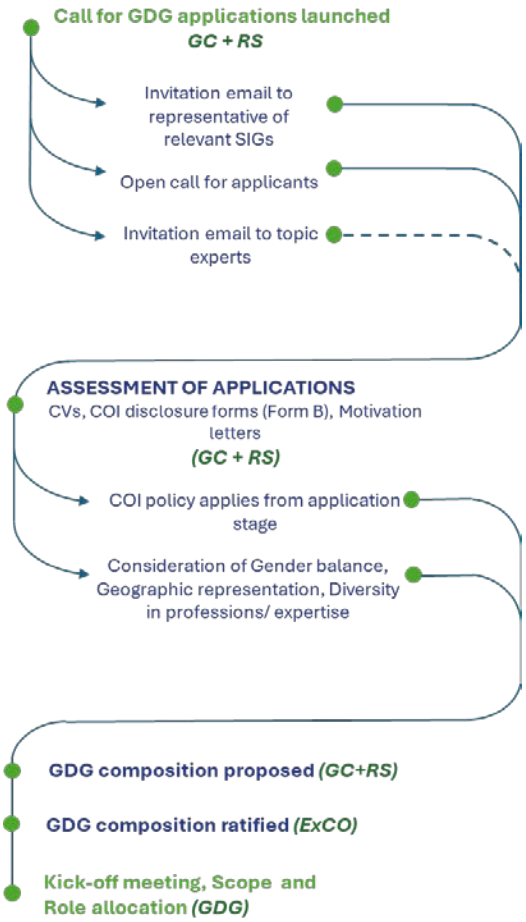
Training

The research specialist should provide a training on COI for the GDG members during the first online GDG meeting. The training should cover the types of COI that should be disclosed, how COIs are assessed and managed by the Guidelines Committee, why COI disclosure is essential for guideline activity and what a COI management plan is and how it will be executed.

Consensus

GDG members need to make collective decisions throughout the entire development of a guideline. Such consensus includes generating key questions, agreeing the best evidence to answer them, and formulating recommendations. There are many different approaches to making group decisions and reaching consensus, but there is no blueprint about which approach should be used. Resources for consensus development methods can be found in the systematic review by Murphy and others (Murphy et al., 1998). The most commonly used consensus development methods are the: nominal group technique, Delphi survey, and RAND/UCLA appropriateness method, which may incorporate structured discussion and when necessary, formal voting to reach agreement. ESHRE primarily relies on structured group discussions, with all viewpoints considered and documented in meeting minutes. Consensus is generally reached through iterative discussion and agreement among members. When consensus cannot be readily achieved, a formal voting may be considered to reach a qualified majority (≥75% agreement after structured discussion).

Summary



Tips

→ Document the guideline group member selection process and roles to ensure transparency.

→ Consider the optimum group size for the GDG, particularly the guideline panel (e.g. too small of a group may lack sufficient experience, content expertise and wide representation, too large of group may lack cohesiveness and effective group interaction).

→ Record within the guideline that its development was without external funding.

→ Set expectations and awareness of the group process through an introduction, training, and support for the GDG members (e.g. setting ideal conditions for group discussion and decision-making).

→ Set a quorum for meetings (e.g. 75% of group must be present to formulate guideline recommendations), but expect that all group members attend all meetings as far as possible.

→ Record the composition of the GDG (names, professions, represented organisations, geographical location) within the guideline.

→ Record competing interests of the GDG within the guideline, particularly where the conflicts bear on specific recommendations.

Available forms/checklists:

Ⓑ Disclosure form
Ⓒ Confidentiality form

3. SCOPING THE GUIDELINE

The aim of the scoping process is to define the overall objectives of the guideline (e.g. potential impact and benefits), the patients and target users to whom the guideline is meant to apply and its relation to other (ESHRE) documents. The scope also defines which populations, settings and activities/services/aspects of care will NOT be covered in the guideline.

Scoping procedure

In general, a scoping procedure starts with a kick-off meeting of the GDG. A preliminary literature search or a survey of target users and patients can be performed to provide input for the scope of the guideline.

The scoping checklist (form Ⓓ) is completed to document the consensus of the GDG on what is within and outside the scope of the guideline.

The scope should be accepted by the entire group before the GDG begins to formulate the key questions as the basis for literature searching.

The scope of the guideline, together with the key questions, will be published on the ESHRE website for stakeholder comments for 2-4 weeks in the format of a survey. Potential stakeholders should note that ESHRE can accept only one set of answers per reviewer. Comments submitted outside the official survey will not be considered as part of the stakeholder comments. ESHRE also reserves the right not to review comments that are hostile (e.g. personal attacks, offensive language), inappropriate (e.g. commercial promotion), or otherwise unsuitable for constructive review. All comments should be solely aimed at improving the quality and clarity of the guideline.

The GDG does not commit to altering the scope or key questions in all instances.

Summary

Tips

→ Guideline scoping provides the opportunity for patient consultation.

→ Establish a method and criteria to generate and prioritise a candidate list of topics to be addressed within the guideline (e.g. where evidence is most confusing or controversial, where there is currently uncertainty or inconsistency in practice, questions about screening, diagnosis, and treatment, etc.).

→ Consult appropriate stakeholders to ensure all relevant topics for the guideline have been identified and will meet the needs of the target audience(s).

→ Record the overall objectives of the guideline, and the specific health benefits.

→ Describe the population to whom the guideline is meant to apply.

→ Record the target users of the guideline.

Available forms/checklists:

Ⓓ Scoping checklist

4. FORMULATING KEY QUESTIONS

Effective and efficient guideline development involves asking and answering key questions. Key questions should be clear, focused and closely define the boundaries of the topic. They are important both as the starting point for the subsequent systematic literature review and as a guide for the development of recommendations. After the formulation of the key questions is finalised, the timelines for guideline development should be set.

Developing and selecting key questions

The key questions are developed from the guideline's scope. The scope is divided into different clinical stages (e.g. diagnosis, prognosis, treatment) and for each stage key questions are defined. It is generally not acceptable to define key questions on topics that have not been covered in the scope. Generally, a list of key questions is proposed by the co-chairs or after consulting the GDG members, and these are further defined and structured at the first GDG meeting.

Around 15-20 questions would be a reasonable number of key questions for guidelines taking 18-24 months to develop, but this depends largely on the complexity of the topic and the questions.

During the final selection of key questions within a guideline, the overall guideline outline should be kept in mind; each step of a clinical scenario needs to be addressed in a logical sequence. For example: diagnostics, treatment options, monitoring options, potential benefits/risks, outcome, prevention, and information provision. The GDG selects the definitive key questions by consensus.

A significant proportion of the key questions will focus on interventions: these questions should be formulated in a structured format, based on the defined PICO components (see below):

→ Should [intervention] vs. [comparison] be used for [health problem]/[population]?

For intervention questions, PICO components should be defined, a formal evidence synthesis should be carried out, and GRADE evidence profiles should be developed. A similar approach is suggested for questions on diagnosis, although it is not yet relevant to create GRADE evidence profiles for diagnostic questions. The proposed structured question is:

→ Should [intervention] vs. [comparison] be used to diagnose [target condition] in [health problem and/or population]

In addition to interventions and diagnosis, other types of questions may arise. Some of these questions will fit the PICO format (although some components may be non-applicable), and a systematic assessment of the available evidence can be relevant, and recommendations can be formulated. For other questions, a formal systematic assessment of evidence synthesis is often irrelevant. These questions are often answered in a narrative format and conclusions or good practice points, rather than recommendations, are formulated by the GDG. When defining these questions, the GDG should define whether a systematic or narrative review is relevant. Examples of questions, and how they can be handled are:

WHAT CAUSES THE PROBLEM?	AETIOLOGY, RISK FACTORS	PICO
WHAT IS THE FREQUENCY OF THE PROBLEM?	FREQUENCY, PREVALENCE	PO
WHO WILL GET THE PROBLEM?	PROGNOSIS, PREDICTION	PICO
WHAT IS THE DEFINITION?	DEFINITION	(narrative)
WHAT IS THE CLINICAL PRESENTATION?	DEFINITION	(narrative)

For these questions (not on diagnosis or interventions), it is not relevant to create GRADE evidence profiles.

Defining key questions as PICO questions

The PICO framework is a well-accepted methodology for framing clinical questions (Richardson et al., 1995). This framework divides each question into four components (see also template for PICO questions Ⓔ): Patients/population, Interventions, Comparisons and Outcomes.

Table 4.1. Definition of PICO components and factors to consider [3] ***(The Cochrane Handbook 2011, The GRADE Working Group, 2013)***

			EVIDENCE-BASED GUIDELINES
P	PATIENT POPULATION	The patients or population to whom the recommendations are meant to apply	➔ How is the disease/condition defined? ➔ What are the most important characteristics that describe the people? ➔ Are there any relevant demographic factors (e.g. age, sex, ethnicity)? ➔ What is the setting (e.g. hospital, community)? ➔ Who should make the diagnosis? ➔ Are there any other types of people who should be excluded from the review (because they are likely to react to the intervention in a different way)? ➔ How will studies involving only a subset of relevant participants be handled? ➔ *Consider the prevalence of multiple comorbidities in the population*
I	INTERVENTION	The therapeutic, diagnostic, or other intervention under investigation (e.g. the experimental intervention, or in observational studies the exposure factor)	➔ What are the experimental and control (comparator) interventions of interest? ➔ Does the intervention have variations (e.g. dosage/intensity, mode of de- livery, personnel who deliver it, frequency of delivery, duration of delivery, timing of delivery)? ➔ Are all variations to be included (for example is there a critical dose below which the intervention may not be clinically appropriate)? ➔ How will trials including only part of the intervention be handled? ➔ How will trials including the intervention of interest combined with another intervention (co intervention) be handled? ➔ *Identify whether multiple (treatment) comparisons should be included.*
C	COMPARISON	The alternative intervention; intervention in the control group	
O	OUTCOME	The outcome(s) of interest	➔ Main outcomes, for inclusion in the 'Summar y of findings' table, are those that are essential for decision making, and emphasis should be on patient important outcomes. ➔ Primary outcomes are the two or three outcomes among the main outcomes that the review would be likely to be able to address if sufficient studies are identified, to reach a conclusion about the effects (beneficial and adverse) of the intervention(s). ➔ Secondary outcomes include the remaining main outcomes (other than primary outcomes) plus additional outcomes useful for explaining effects. ➔ Ensure that outcomes cover potential as well as actual adverse effects. ➔ Consider outcomes relevant to all potential decision makers, including eco- nomic data. ➔ Consider the type and timing of outcome measurements. ➔ *Include both desirable (e.g. benefits, less burden, savings) and undesirable effects (e.g. harm, burden, costs, and decrease in patient autonomy).* ➔ *Do not ignore important outcomes for which evidence may be lacking.*

[3] Additions in Italic based on GIN-McMaster Guideline Development Checklist.

Defining the patient population and intervention for guideline question is challenging, and should consider the underlying biology. Defining a broad patient population or intervention may be okay if across the range of patients and interventions it is plausible that the magnitude of effect on the key outcomes is more or less the same. If that is not the case the review will generate misleading estimates for at least some subpopulations of patients and interventions, and the questions should be defined narrower or split up. Also, and different to a systematic review calculating relative risks, recommendations may differ across subgroups of patients at different baseline risk of an outcome, despite there being a single relative risk that applies to all of them. Thus, guideline panels must often define separate questions (and produce separate evidence summaries) for high- and low-risk patients, and patients in whom quality of evidence differs.

Another challenge is defining the comparators. Mostly, guideline groups will be strict in defining the intervention, but will define the comparator as "all other interventions". Clarity in choice of the comparator makes for interpretable guidelines, and lack of clarity can cause confusion.

To make sensible recommendations all relevant outcomes that are important or critical to patients for decision making must be considered and included. Recommendations cannot be made based on information about single outcomes and decision-making always involves a balance between health benefits and harms. GDGs must base the choice of outcomes on what is important, not on what outcomes are measured and for which evidence is available. If evidence is lacking for an important outcome, this should be acknowledged, rather than ignoring the outcome. Most systematic reviews do not summarise the evidence for all important outcomes, and evidence from other sources should be included.

In GRADE, outcomes should be classified on importance for decision-making in 3 categories; critical, important but not critical, and of limited importance. Ranking outcomes by their relative importance can help to focus attention on those outcomes that are considered most important, and help to resolve or clarify disagreements. Practically, to generate a list of relevant outcomes, one can use the following type of scales (The GRADE Working Group, 2013):

RATING SCALE								
1	2	3	4	5	6	7	8	9
Of **least** importance								Of **most** importance
OF LIMITED IMPORTANCE FOR MAKING A DECISION			**IMPORTANT, BUT NOT CRITICAL**			**CRITICAL FOR MAKING A DECISION**		
Not included in evidence profile			Included in evidence profile			Included in evidence profile		

It is important to realise that the importance of outcomes is likely to vary within and across cultures or when considered from the perspective of the target population (e.g. patients or the public), clinicians or policymakers. The perspective would generally be that of the patient, and a literature search can be conducted on patients' values and preferences about the intervention in question to inform the rating of the importance of outcomes. Reviewing the evidence may provide the panel with insight into the variability in patients' values, the patient experience of burden or side effects, and the weighing of desirable versus undesirable outcomes. However, often such evidence is not available, and panel members should use their prior experiences with the target population to assume the relevant values and preferences.

Lastly, basic selection criteria (e.g., duration of a follow-up period, the primary outcome measure, age limits) should be established by the GDG in advance of the evidence search.

The key questions of the guideline, together with the scope, will be published on the ESHRE website for stakeholder comments for two weeks in the format of a survey. Potential stakeholders should note that ESHRE can accept only one set of answers per reviewer. Comments submitted outside the official survey will not be

considered as part of the stakeholder comments. ESHRE also reserves the right not to review comments that are hostile, inappropriate, or otherwise unsuitable for constructive review. All comments should be solely aimed at improving the quality and clarity of the guideline.

The GDG does not commit to altering the scope or key questions in all instances.

Modifications to the key questions

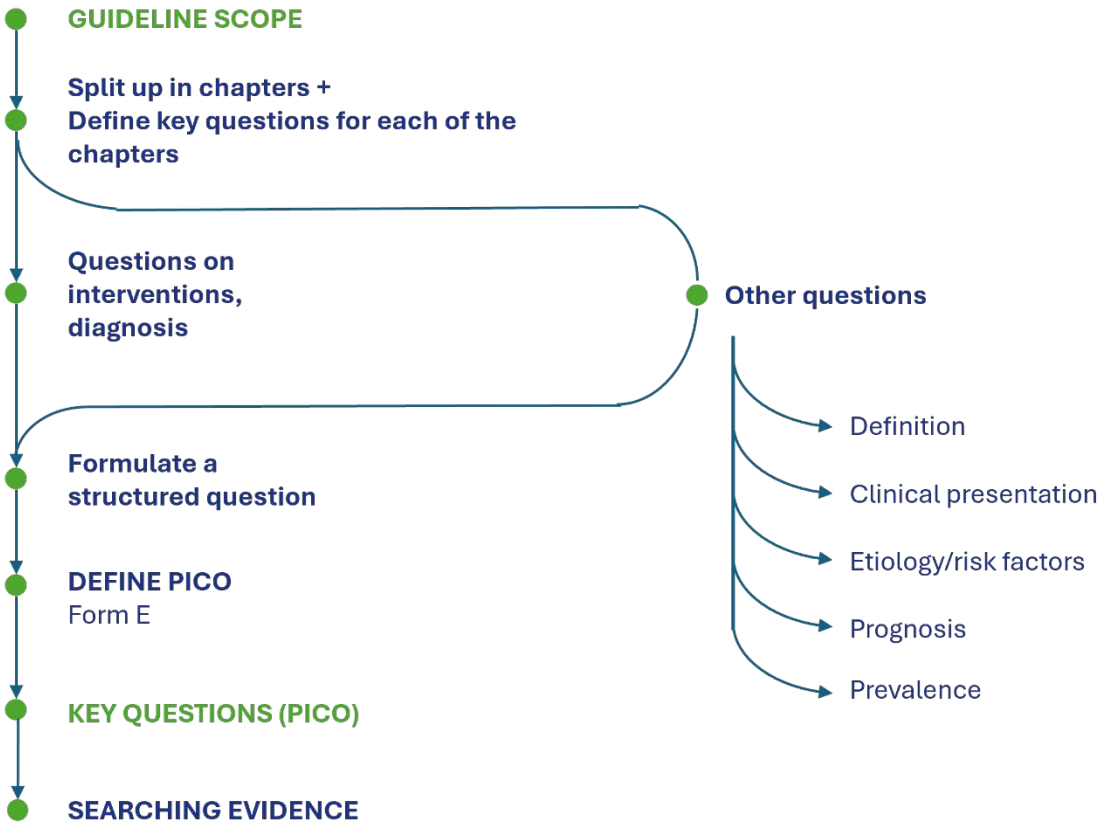
In general, after consensus by the GDG, the key questions are final and modifications should be minimised. However, once the evidence has been searched, the key questions may need refining. In any case, the entire GDG should be informed of and agree with any changes to the key questions.

Changes to the key questions could include:

- Reassessment of the importance of the outcomes
- Addition of an outcome: for instance, the importance of an outcome (e.g. a serious adverse effect) may only become known after the evidence synthesis
- Addition of an intervention that is relevant for the flow and consistence of the guideline
- Specific search on an important subgroup, not defined as such in the PICO questions
- Merging of key questions based on significant overlap of the evidence synthesis.
- Splitting up of key questions

As changes to the key questions imply additional work for the research specialist and the GDG members, these should be avoided as much as possible.

Summary



Tips

→ Generate and document the key questions (e.g. clinical, health, policy) to be answered in the guideline using a standard format (e.g. PICO) and determine the criteria by which the questions generated will be prioritised if it is not feasible to answer all questions (e.g. survey guideline panel members, survey stakeholders).

→ Select no more than 15-20 key questions.

→ Define key questions in such a way that answering the question gives the opportunity to make a recommendation.

→ Think of formulating key questions in addition to health benefits – and on side effects and risks.

→ List all key questions in the guideline, at the start of each guideline section.

→ Set timelines for the whole guideline development process.

→ Consult appropriate stakeholders to ensure all relevant topics for the guideline have been identified and will meet the needs of the target audience(s).

Available forms/checklists:

Ⓔ Template PICO Question

5. SEARCHING EVIDENCE

The identification and selection of evidence is an essential step towards answering the key questions. Secondly, to perform an evidence search the key questions should be translated into key words or search terms. The evidence search itself should be gathered in a systematic process to avoid or minimise bias. Finally, from the identified literature the relevant evidence should be selected for summary and evaluation. The search terms, PRISMA flowchart of literature search and the list of excluded studies is documented in Annex 7 of the guideline.

Literature search

According to the National Academy of Medicine (NAM), literature searching is the key step in developing valid guidelines, because incomplete or biased literature evaluation can lead to inappropriate recommendations. The search for relevant research should be comprehensive, based on explicit criteria, and the validity of the results should be judged in a rigorous and reproducible fashion (Cook et al., 1997).

ESHRE applies a stepwise methodology, common to guideline development, focusing on the best available evidence to address each key question (Scottish Intercollegiate Guidelines Network (SIGN), 2015). A set of standard search filters is used for identification in the following order:

- Systematic reviews/meta-analyses of randomised controlled trials
- Randomised controlled trials
- Non-randomised studies / observational studies
- Case reports/opinion documents.

The literature search will be performed in this stepwise approach, but all studies will be available for title and abstract screening by the assigned GDG member. Where adequate published systematic reviews exist, the review and additional studies from the time period since the review was conducted should be selected. If no systematic review exists, the next type of studies to be assessed are RCTs (at least for intervention questions), followed by non-randomised and observational studies. For observational studies, preference is given to prospective cohort studies. When including retrospective studies, it is important to make sure that appropriate statistical measures have been used to control for confounding factors. Retrospective studies that have not been corrected for confounding factors should not be used.

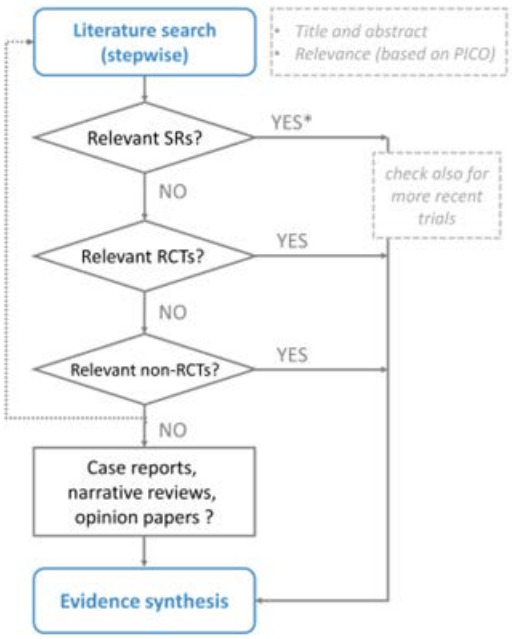


Figure 2. Stepwise literature search and evidence selection workflow

The process for evidence identification should be repeatable and transparent. The search strategy, including search terms, should therefore be documented and stored. This also simplifies running the search strategies to check the validity of a guideline. The results of the literature searches are documented and published with the guideline as Annex 7.

The ESHRE research specialist will conduct the literature searches, based on a list of search terms for each of the PICO questions defined by the GDG members. Literature searching includes at least MEDLINE/PubMed and the Cochrane Library, but additional sources can be covered (e.g. NHS Economic Evaluations Database (NEED), PsycInfo and Embase) specific to the topic under review.

The searches are limited to:

- Peer-reviewed published literature
    - *the use of conference abstracts should be avoided unless there are good reasons for including them (and always combined with a search for the full paper)*
    - *unpublished clinical trials should be avoided to support any recommendation.*
- English language
- Human subjects
- Defined time frame; searches in a guideline update are limited to the period following the last publication of the guideline; if a suitable systematic review is identified, an update search is limited to the time period following the reported search cut-off date.

Conference abstracts and unpublished results from clinical trials seldom contain enough information to allow confident judgements about the quality and results of a study. They are therefore not routinely included, unless there are good reasons for doing so. Generally, only peer-reviewed studies in the English language are considered for inclusion in ESHRE guidelines, unless decided otherwise by the GDG. In this case, justification for doing so should be clearly documented in the methodology section of the guideline.

The literature searches are updated at least once during the development process, to make sure the literature supporting the recommendations is up-to-date. Ideally, the guideline is published within one year after final evidence search.

Although the research specialist performs a preliminary level of selection based on title and abstract, the clinical expertise of GDG members is necessary to carry out the definitive selection of the search output.

Different questions may be best answered by different databases, or may rely on different levels of evidence. Following evaluation of the first search results the key questions may be redefined and subsequent searches focused on the most appropriate sources and study types. As a result, the assembly of evidence is a stepwise and iterative process.

Quality assurance of search strategies

The research specialist will check the quality and accuracy of search strategies during the development of the guideline. Although it will not usually be possible to check all strategies for every search, the following approaches can be used to ensure that the key studies are retrieved.

- Ask GDG members to identify key clinical studies that are already published, in order to gather useful search terms
- Check search strategies used in existing published systematic reviews
- Run searches with and without certain search terms and assess the differences between the results obtained.
- Check the bibliographies of includes studies to ensure that all relevant papers have been retrieved by the search strategy used.
- During pre-selection of the evidence for a specific question, identify the papers that could be relevant for other questions in the guideline and check if they have been included in those questions.
- If relevant papers have not been retrieved by the search strategy, investigate and amend the strategy

Selection of evidence

Papers are initially pre-selected according to title and abstract by the ESHRE research specialist and the final selection is made by the GDG member.

First, the titles of the retrieved citations are scanned and those that fall outside the topic of the guideline are eliminated. Next, a quick check of the remaining abstracts identifies further papers not relevant to the key questions, and these are also excluded. The remaining abstracts are investigated if they fulfil the selection criteria agreed by the GDG. If no or incomplete information is available in the abstract, the reference is selected and in the next step, the full text is assessed for relevance and quality to ensure its validity and applicability. The study selection process is clearly documented and details the applied inclusion criteria.

Role of qualitative research

At present there is no established mechanism for incorporating qualitative studies in evidence-based guideline development. Nevertheless, the use of qualitative studies can help identify issues of concern to patients. A qualitative approach to complement trial data in the collection of information on patient preferences and the values placed on outcomes would perhaps help bridge the gap between scientific evidence and clinical practice. In case qualitative studies are used to support recommendations, an appropriate quality assessment checklist should be used to validate the quality of the studies.

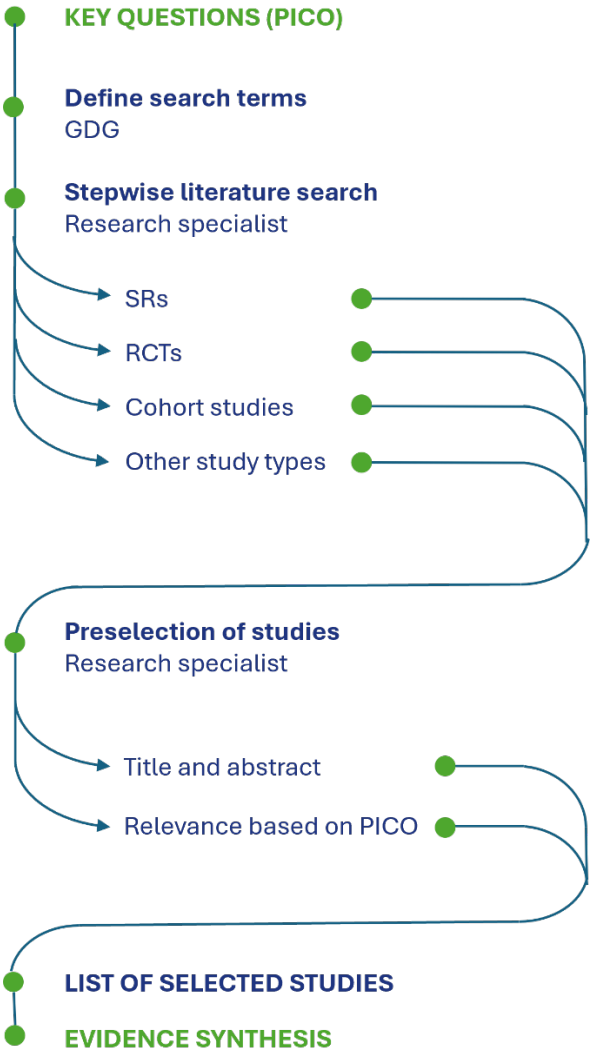
Narrative or descriptive review are generally not selected during evidence synthesis, but they may be helpful as background information. These papers represent an interpretation of evidence in the context of experts' experiences and knowledge. Expert opinion is not evidence per se and should not be used as evidence; rather, experience or observations that support expert opinions should be described, identified and, if possible, appraised in a systematic and transparent way.

Updating of the literature searches during the guideline development process

Searches undertaken to identify evidence for each review question may be re-run before consultation (or before publication). Searches are re-run especially on topics where the evidence changes quickly, there is reason to believe that substantial new evidence exists, or the development time is longer than usual.

If evidence is identified after the last cut-off date for searching, but before publication, a judgement on its impact is made by the development team and staff. In exceptional circumstances, this evidence can be considered if its impact is judged as potentially substantial.

Summary



Tips

→ Follow systematic review methods (either full systematic reviews or rapid systematic reviews depending on the topic and organisatin's framework) or provide a rationale for why this is not done.

→ Develop a protocol for locating, selecting, and synthesising the evidence (e.g. conduct a search for existing systematic reviews, new systematic review and grey literature search) and determine the types of evidence to include (e.g. data-bases searched, types of studies, inclusion and exclusion criteria, searching for specific studies on adverse effects or deciding to abstract information on adverse effects from studies on benefit).

→ Identified evidence gaps can be used for future research goals.

→ Document and store the search strategies used.

→ Record how patients' perspectives are included within the evidence search.

6. EVIDENCE SYNTHESIS

Studies identified during the stepwise literature search should be reviewed to identify the most appropriate data for answering the key questions and ensure that recommendations are based on the best available evidence. This process should be explicit and transparent and should be carried out through a systematic review process. This involves selecting relevant studies (step 6), assessing relevance and quality, summarising the results, and grading the evidence.

Relevance and Quality check

Relevance and quality assessment (template Ⓕ) of the selected evidence is necessary to ensure that recommendations are based on the highest quality evidence available. Quality assessment is performed on each individual study. However, if the study is a meta-analysis or systematic review quality assessment should be performed on the meta-analysis or review itself and not on the studies included. Depending on the type of study, different checklist should be used (checklists Ⓖ). For systematic reviews and meta-analysis, the AMSTAR 2 quality assessment checklist is recommended.

A study should be rejected if its quality is assessed as low. Hereto, checklists can be used, such as the Cochrane Risk of Bias 2 (RoB2) tool. If no better evidence can be found, the study might be considered as low level evidence, comparable with expert opinions.

In addition to the inherent quality of a paper, the applicability of findings (relevance) should also be assessed.

Applicability or relevance is related to the definition of the components (PICO) of the formulated key questions. Comparison of the available articles with the defined PICO components guides the selection of papers with the relevant evidence.

The validity of a study is the extent to which its design and conduct are likely to prevent systematic errors, or bias. There are four potential sources of systematic bias in healthcare trials:

- Selection bias – randomisation and allocation concealment
- Performance bias – blinding
- Attrition bias – handling participant loss and
- Detection bias – outcome assessment.

One of the most important factors leading to bias and distorted treatment comparisons is patient assembly (selection bias). An appropriate method for preventing foreknowledge of treatment assignment is crucial for any study. Thus, studies are for the selection bias judged on the quality of the used allocation concealment methodology.

Performance bias refers to systematic differences in the provision of care to the participants in the intervention and control group. Those providing and receiving care can be 'blinded' to protect against unintended differences in care.

Attrition bias, also known as exclusion bias, alludes to systematic differences in the approach to handling the loss of participants (e.g. withdrawals, dropouts, protocol deviations) in the two study groups. This may have great potential for biasing results.

Detection bias is a systematic difference between two study groups in outcome assessment. Trials that blind those assessing outcomes are logically less likely to be biased than trials that do not.

Risk of bias assessment categorises studies as low, moderate or high-risk bias based on the 4 sources of potential bias. The assessment can then be used as a (1) a threshold for study inclusion (e.g. for studies judged at high risk

of bias, this assessment constitutes grounds for study exclusion), or (2) a possible explanation for found differences in study results.

Risk of bias assessment provides a structured evaluation of the possible sources of bias. However, it is important to go back to the question and assess how important the study flaws are in the interpretation of the overall results.

The study selection procedure and results of the risk of bias assessment and relevance should be documented and will be published as an annex to the guideline. At this point the available evidence is ready for summary.

Integrity check

It is not within ESHRE's remit to conduct a formal investigation or to draw formal conclusions regarding the misconduct of an individual or group of individuals or to determine whether a published article should be retracted. However, papers that are withdrawn, have a published editorial note of concern or a published expression of concern will be excluded from the guideline.

When updating the guideline, the GDG will actively verify the status of all the referenced studies.

Evidence tables

Evidence tables help to identify similarities and differences between studies. Data for inclusion within an evidence table should be extracted according to a standard template (checklists ©). Here, key characteristics of the study population (e.g. sample size, age), are important. For intervention studies, this includes: intervention (e.g. follow-up period, kind of intervention), comparison (e.g. IUI versus timed intercourse) and outcome measures (e.g. summary of the absolute and relative effect size, benefits versus harms) are important. For diagnostic studies, this includes: diagnostic test accuracy, reference test, reproducibility and effect size. The evidence table was updated to the recommendations of the evidence table working group of the Guidelines International Network (http://www.g-i-n.net/). The completed evidence tables will be published as an appendix to the guideline (Template evidence table Ⓗ).

GRADE Evidence profiles and Summary of Findings Tables

A GRADE evidence profile allows presentation of key information about all relevant outcomes for a given healthcare question. It presents information about the body of evidence (e.g. number of studies), the judgments about the underlying quality of evidence, key statistical results, and the quality of evidence rating for each outcome.

A GRADE evidence profile is particularly useful for presentation of evidence supporting a recommendation in clinical practice guidelines. It includes:

- A list of outcomes (those considered critical and important; classification of the importance can be added)
- The number of studies and study design(s)
- Judgements about each of the quality of evidence factors assessed; risk of bias, inconsistency, indirectness, imprecision, other considerations (including publication bias and factors that increase the quality of evidence)
- The assumed risk; a measure of the typical burden of the outcomes, i.e. illustrative risk or also called baseline risk, baseline score, or control group risk
- The corresponding risk; a measure of the burden of the outcomes after the intervention is applied, i.e. the risk of an outcome in treated/exposed people based on the relative magnitude of an effect and assumed (baseline) risk
- The relative effect; for dichotomous outcomes the table will usually provide risk ratio, odds ratio, or hazard ratio

- The absolute effect; for dichotomous outcomes the number of fewer or more events in treated/exposed group as compared to the control group
- Rating of the overall quality of evidence for each outcome (which may vary by outcome)
- Footnotes, if needed, to provide explanations about information in the table such as elaboration on judgements about the quality of evidence

GRADE evidence profiles are used for discussion of recommendations during guideline meetings. Summary of findings tables provide similar information in a more accessible format. Where relevant, and at least for all interventions questions with more than 1 RCT available, summary of findings tables will be drafted by the research specialist and added in the guideline (Annex 6). The corresponding GRADE evidence profiles will be used for discussion and are available upon request.

Quality of evidence for each outcome

The quality of evidence reflects the extent to which our confidence in an estimate of the effect is adequate to support a particular recommendation. It gives the reader a quick impression of the quality of the supporting evidence, which is not necessarily related to the importance of the recommendation.

Guideline panels must make judgments about the quality of evidence relative to the specific context for which they are using the evidence.

The GRADE approach involves separate grading of quality of evidence for each patient-important outcome (across studies) followed by determining an overall quality of evidence across outcomes. Although the quality of evidence represents a continuum, the GRADE approach results in an assessment of the quality of a body of evidence in one of four grades:

Table 6.1: Quality of Evidence Grades (Balshem et al., 2011)

GRADE	DEFINITION
High (⊕⊕⊕⊕)	We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate (⊕⊕⊕⊖)	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low (⊕⊕⊖⊖)	Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.
Very Low (⊕⊖⊖⊖)	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Factors determining the quality of evidence

The GRADE approach to rating the quality of evidence begins with the study design (trials or observational studies) and then addresses five reasons to possibly rate down the quality of evidence and three to possibly rate up the quality (Balshem et al., 2011).

For intervention studies, randomised trials provide, in general, far stronger evidence than observational studies, and rigorous observational studies provide stronger evidence than uncontrolled case series. As such, RCTs without important limitations provide high-quality evidence, while observational studies without special strengths or important limitations provide low-quality evidence.

In case of RCTs, 5 factors should be assessed to detect limitations and reduce the quality of the evidence (for a certain outcome). In observational studies, 3 factors should be assessed to detect strengths and increase the quality of the evidence (for a certain outcome). If one or more of these factors is met (and there is no reason for downgrading), it is possible to rate up the quality.

Table 6.2: Factors for upgrading and downgrading quality of evidence (Balshem et al., 2011)

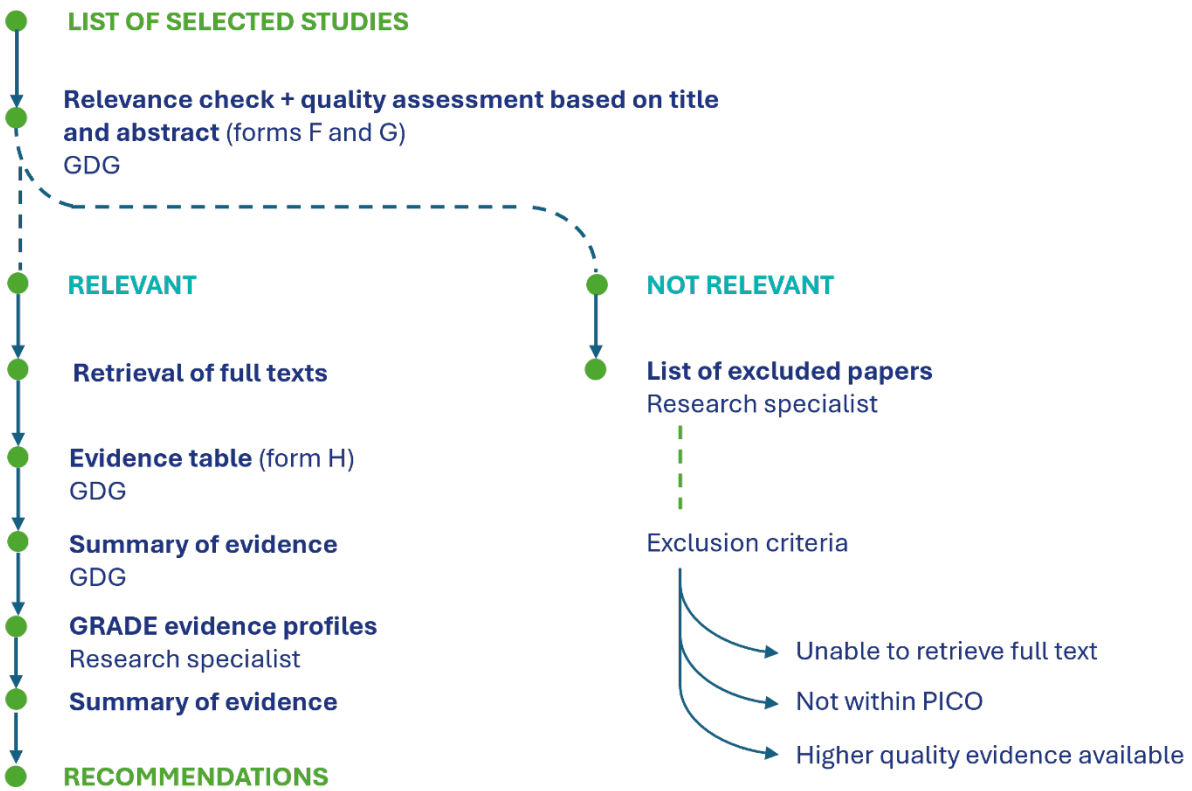
SOURCE OF THE BODY OF EVIDENCE	INITIAL QUALITY OF THE BODY OF EVIDENCE	FACTORS THAT MAY DECREASE THE QUALITY	FACTORS THAT MAY INCREASE THE QUALITY	QUALITY OF A BODY OF EVIDENCE*
Randomised trials	**High**	1. Risk of bias 2. Inconsistency 3. Indirectness 4. Imprecision 5. Publication bias	1. Large effect 2. Dose–response 3. All plausible residual confounding would reduce the demonstrated effect or would suggest a spurious effect if no effect was observed	**High** (⊕⊕⊕⊕)
				Moderate (⊕⊕⊕⊖)
Observational studies	**Low**			**Low** (⊕⊕⊖⊖)
				Very low (⊕⊖⊖⊖)

In the end, the overall quality of evidence for an intervention across outcomes is the lowest quality of evidence for the critical outcomes, as the overall confidence in effect estimates cannot be higher than the lowest confidence in effect estimates for any outcome that is critical for a decision.

Observational studies, by default, start at low-quality and may be upgraded to moderate quality only when the three GRADE upgrading criteria have been assessed (large effect, a dose-response relationship, residual confounding factors favouring null).

Case series and case reports are observational studies that investigate only patients exposed to the intervention. Source of control group results is implicit or unclear, thus, they will usually warrant downgrading from low to very low quality evidence.

Summary



Tips

→ Document and publish the search and selection of evidence, judging eligibility, range of evidence included, and search strategies used to ensure the methods are explicit and transparent.

→ Summarise the evidence using a concise summary (e.g. evi- dence table, evidence profile or summary of findings table) of the best available evidence for each important outcome, including diagnostic test accuracy, anticipated benefits, harms, resources (costs), the quality of evidence rating, and a summary of the relative and absolute results/estimate of effect for each outcome.

→ Assess the quality of evidence for each important outcome.

→ Assess and report the overall quality of evidence (e.g. lowest quality of evidence from outcomes rated as most important or critical, or highest quality of evidence when all outcomes point in the same direction).

→ Document the judgements made in appraising the quality of evidence to ensure they are transparent and explicit.

→ Record the set of evidence selection criteria.

→ Record the strengths and limitations of the evidence.

Available forms/checklists:

Ⓕ Template Relevance and Quality check
Ⓖ Quality assessment checklists
Ⓗ Template evidence table

Factors for downgrading the quality of evidence

1. Risk of bias

The risk of bias of the included studies should be assessed in relation to the effect on the outcome (Guyatt et al., 2011d). In assessing the studies, the weight of the studies in the meta-analysis should be considered as small studies with high risk of bias may not necessarily impact on the estimate of effect if combined with a very large study at low risk of bias.

Guidance to assess risk of bias and corresponding downgrading for limitations in study design:

RISK OF BIAS	ACROSS STUDIES	INTERPRETATION	CONSIDERATIONS	GRADE
Low risk of bias	Most information is from studies at low risk of bias	Plausible bias unlikely to seriously alter the results	No apparent limitation	No serious limitations, do not downgrade
Unclear risk of bias	Most information is from studies at low or unclear risk of bias	Plausible bias that raises some doubt about the results	Potential limitations are unlikely to lower confidence in the estimate of effect	No serious limitations, do not downgrade
			Potential limitations are likely to lower confidence in the estimate of effect	Serious limitations, downgrade 1 level
High risk of bias	The proportion of information from studies at high risk of bias is sufficient to affect the interpretation of results	Plausible bias that seriously weakens the confidence in the results	Crucial limitation for one criterion or some limitations for multiple criteria, sufficient to lower confidence in the estimate of effect	Serious limitations, downgrade 1 level
			Crucial limitation for one or more criteria, sufficient to substantially lower confidence in the estimate of effect	Serious limitations, downgrade 2 levels

2. Inconsistency of results

Estimates of treatment effect across studies can differ because of clinical heterogeneity (P, I, O: for instance larger effect with higher dose, or shorter time of follow-up), or methodological heterogeneity (differences in study design) (Guyatt et al., 2011c).

In case of (unexplained) inconsistency, the quality can be downgraded, or subgroups can be presented.

Inconsistency can be detected by assessing confidence intervals and direction of effect for the included studies in a forest plot, or by means of the outcome of statistical tests for heterogeneity (I^2 statistic; >60% = substantial, p-value)

3. Indirectness of evidence

Downgrading for indirectness can be considered if the evidence from the studies is different from the PICO question (Guyatt et al., 2011b). Examples could be:

- Indirect comparison between 2 interventions (A vs placebo and B vs placebo instead of drug A vs drug B)
- Population : evidence in menopausal women instead of POI
- Comparator
- Outcomes: surrogate outcomes: bone density instead of fractures

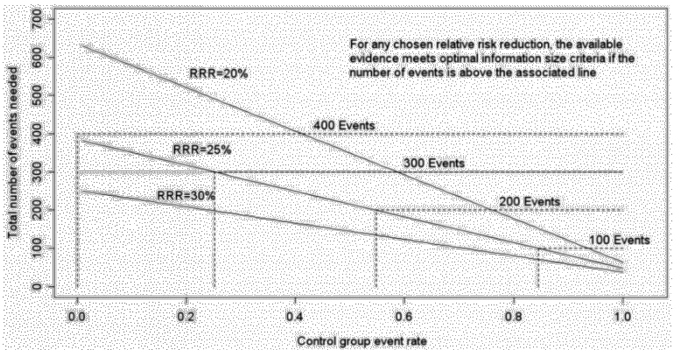
4. Imprecision

In general, results are imprecise when studies include relatively few patients and few events and thus have a wide confidence interval (CI) around the estimate of the effect (Guyatt et al., 2011a). In this case, one may judge

the quality of the evidence lower than it otherwise would be considered because of resulting uncertainty about the results.

In general, downgrading for imprecision can be considered in cases of small sample size (optimal information size not met), and/or wide confidence intervals.

The optimal information size is the threshold of number of events that needs to be included. The criteria for the optimal information size can be estimate by plotting the background risk against a chosen relative risk reduction in the following graph.



If the total number of patients included in a systematic review is less than the number of patients generated by a conventional sample size calculation for a single adequately powered trial, consider rating down for imprecision.

Regarding the wide confidence intervals, downgrading is appropriate when the confidence interval crosses the clinical decision threshold between recommending and not recommending treatment. This clinical decision threshold should be set based on the intervention. For example, for an intervention with limited adverse events, inconvenience and cost, the threshold for an absolute reduction in pregnancy loss may be set at 0.5%, which means that even a small benefit of the intervention would lead to recommending it. Alternatively, for an intervention with significant toxicity, the clinical decision threshold could be set to at least 1%. A theoretical intervention with an absolute reduction of pregnancy loss of 1.3% (95% CI 0.6% to 2.0%), would be rated down for imprecision with the clinical decision threshold of 1%, but not with a threshold of 0.5%.

5. Publication bias

Publication bias is a systematic under-estimation or an over-estimation of the underlying beneficial or harmful effect due to the selective publication of studies (Guyatt et al., 2011d). Confidence in the combined estimates of effects from a systematic review can be reduced when publication bias is suspected, even when the included studies themselves have a low risk of bias. Funnel plots can be used to detect publication bias.

Factors for upgrading the quality of evidence (Guyatt et al., 2011e)

1. Large magnitude of effect

When the body of evidence from observational studies yield large or very large estimates of the magnitude of an intervention effect, then we may be more confident about the results. Decisions to rate up quality of evidence because of large (RR>2 or RR<0.5) or very large effects (RR >5 or RR <0.2) should consider not only the point estimate but also the precision (width of the CI) around that effect. Furthermore, upgrading should only be considered in absence of any problems with risk of bias.

2. Dose-response gradient

The presence of a dose-response gradient has long been recognised as an important criterion for believing a cause-effect relationship. The presence of a dose-response gradient may increase our confidence in the findings of observational studies and thereby increase the quality of evidence.

3. Opposing plausible residual confounding

The reason that in most instances we consider observational studies as providing only low-quality evidence is that unmeasured or unknown determinants of outcome unaccounted for in the adjusted analysis are likely to be distributed unequally between intervention and control groups, referred to as “residual confounding” or “residual biases.” On occasion, all plausible confounders (biases) from observational studies unaccounted for in the adjusted analysis (i.e. residual confounders) of a rigorous observational study would result in an underestimate of an apparent treatment effect. If, for instance, only sicker patients receive an experimental intervention or exposure, yet they still fare better, it is likely that the actual intervention or exposure effect is even larger than the data suggest. A parallel situation exists when observational studies have failed to demonstrate an association.

7. DEVELOPING RECOMMENDATIONS

Once the selection and summary of evidence is complete, the available evidence must be combined and condensed into recommendations.

Recommendations are statements mostly proposing a course of action. An ESHRE guideline recommendation should be a stand-alone text written in a complete sentence.

Based on the available evidence, each GDG member prepares specific recommendations and presents them to the other GDG members at the GDG meeting. In addition to the evidence summary, the full systematic review(s) and the original studies and other sources of evidence will be available for the entire GDG during the process and prior to the meetings to inform deliberations (through a collaborative website and/or via electronic communication).

When the GDG has reached consensus on the recommendations, the draft version of the guideline can be written.

Strong or weak recommendations

The strength of a recommendation reflects the extent to which a guideline panel is confident that desirable effects of an intervention outweigh undesirable effects, or vice versa, across the range of patients for whom the recommendation is intended.

According to GRADE, recommendations are classified as “strong” or “conditional” (weak). The strength of a recommendation may not be directly correlated with its priority for implementation (Andrews et al., 2013a).

When the GDG formulates a strong recommendation, they must be certain about the various factors that influence the strength of a recommendation. The GDG should also have the relevant information at hand that supports a clear balance towards either the desirable effects of an intervention (to recommend an action) or undesirable effects (to recommend against an action). When the GDG is uncertain whether the balance is clear or when the relevant information about the various factors that influence the strength of a recommendation is not available, a guideline panel should be more cautious and, in most instances, it would opt to make a weak recommendation. As an alternative to the term “weak” recommendations, the terms conditional (depending on patient values, resources available or setting) or discretionary (based on opinion of patient or practitioner) can be used.

Table 7.1: The implications of the two grades of strength of recommendations in the GRADE approach (Andrews et al., 2013b), with the addition of the implications of a “research only” recommendation and a GPP:

Target group	Strong recommendations*	Conditional (weak) recommendations	Research only recommendations	Good practice points (GPP)**
Patients	Most people in your situation would want the recommended course of action and only a small proportion would not.	Most of the people in your situation would want the recommended course of action, but many would not.	The test or intervention should only be considered by patients and clinicians within the setting of a research trial for which appropriate approvals and safety precautions have been established	Clinicians, patients and policy makers are informed of the advice on the GDG regarding a certain recommendation.
Clinicians	Most patients should receive the recommended course of action.	Recognise that different choices will be appropriate for different patients and that you must make greater effort with helping each patient to arrive at a management decision consistent with his or her values		

		and preferences. Decision aids and shared decision making are particularly useful.		
Policy makers	The recommendation can be adopted as a policy in most situations.	Policy making will require substantial debate and involvement of many stakeholders.	NA	
Researchers	The recommendation is supported by credible evidence or other compelling considerations, making it unlikely that further research would change the recommendation. In some cases, a strong recommendation may be issued despite low or very low certainty of evidence.	The recommendation may be strengthened by additional research. Evaluating the conditions and criteria underlying the conditional recommendation can help identify relevant research gaps.		

** Strong recommendations based on high quality evidence will apply to most patients for whom these recommendations are made, but they may not apply to all patients in all conditions; no recommendation can consider all the often-compelling unique features of individual patients and clinical circumstances.*

*** A good practice point or GPP is written by the GDG to support the recommendations. Advice can for instance be provided on how to establish shared decision making, and on factors to be considered for a specific test or intervention.*

Table 7.1 will be provided in the methodology section of the guidelines to provide clear direction on the implication of the strength of recommendation for clinicians, patients, policy makers, and any other target audience groups.

The decision on a strong or a weak recommendation depends on 5 key factors (Andrews et al., 2013b, The GRADE Working Group, 2013). Judgement on these factors will be documented in a framework and summarised (narratively), with information on the explicit link between the recommendation and evidence supporting the recommendation in a justification statement in the guideline. If informal consensus among GDG members cannot be reached, formal voting will be undertaken, and the outcome will be documented accordingly. In this situation, the addition of a sentence of dissent to the guideline can be considered. However, explicit results of the voting process will not be included in the guideline.

In cases of insufficient evidence, a recommendation based on expert consensus (GPP) or a research-only recommendation may be considered. A research-only recommendation is appropriate when the required studies are feasible or already planned, and there is a reasonable expectation that their results will inform future ESHRE guidelines. Refraining from making a recommendation remains an option but should be used sparingly, as the guideline scoping process has already established a clear need for guidance.

Table 7.2: Key factors for deciding on a strong or a weak recommendation (Andrews et al., 2013b, The GRADE Working Group, 2013):

FACTORS	COMMENTS
Balance between desirable and undesirable outcomes (trade-offs) considering: **- best estimates of the magnitude of effects on desirable and undesirable outcomes** **- importance of outcomes (estimated typical values and preferences)**	The larger the differences between the desirable and undesirable consequences, the more likely a strong recommendation is warranted. The smaller the net benefit and the lower certainty for that benefit, the more likely a weak recommendation is warranted.
Confidence in the magnitude of estimates of effect of the interventions on important outcomes (overall quality of evidence for outcomes)	The higher the quality of evidence, the more likely a strong recommendation is warranted.
Confidence in values and preferences of patients and their variability	The greater the variability in values and preferences, or uncertainty about typical values and preferences, the more likely a weak recommendation is warranted.
Resource use	The higher the costs of an intervention (the more resources consumed), the less likely a strong recommendation is warranted.
Health system perspective, including **- equity (what would be the impact on health inequities?),** **- acceptability (is the option acceptable to key stakeholders?)** **- feasibility (is the option feasible to implement?)**	A higher impact on equity, acceptability and feasibility makes a strong recommendation more likely.

Strong recommendations against an intervention may be warranted when potential harms outweigh anticipated benefits, when discontinuation is unlikely to cause harm, when high-quality evidence demonstrates lack of efficacy or effectiveness, or when evidence of benefit is absent, of very low quality, or highly uncertain.

Conversely, a strong recommendation for a specific subgroup may be appropriate when clear benefit is demonstrated in that subgroup, even if the overall population is unlikely to benefit or may experience harm.

The methods in which additional information is to be incorporated with the synthesised evidence is documented in the annex of the guideline to ensure transparency (e.g. formal consensus on patient values, consensus on equity issues, formal economic analysis, consideration of disaggregated resource use data in a qualitative manner).

Regarding resource use, the guideline will not include a formal analysis of cost effectiveness of recommended practice versus current or established practice. The economic feasibility of recommendations will also not be covered. The clinical and organisational impact of recommendations on costs will be considered in the GDG meetings and, if relevant, described in the justification section of the guideline.

Wording of recommendations

ESHRE guideline recommendations could stand alone and contain enough detailed information to be understandable without references to supporting material. Recommendations are written in complete sentences, and they should answer the key questions. In addition, the wording must be:

- unambiguous
- clearly defined

- actionable
- easy to translate into clinical practice and
- agreed by the complete GDG.

Indicate in the recommendation statements the population for which the recommendation is intended, the intervention being recommended, and the alternative approach(es) or intervention(s). A help to guarantee the formulation of such clear recommendations is the five 'W' rule: each recommendation should be a description about **who** does **what** for **whom**, **when** and in **which way**.

The quality of the evidence and the strength of the recommendation should also be described in proximity to the recommendation statement.

Possible benefits and harms should be quantified as much as possible. Any exceptions to the recommendations should be listed whenever possible.

Despite the lack of studies supporting this, standardised wording is usually defined reflecting the strength of the recommendation. Standardised wording to use for recommendations provides structure for the GDG members and aids to ensure clarity and to maintain consistency throughout the guideline, avoiding wording that may be vague and nonspecific (see **table 7.3**).

Table 7.3: Recommended phrasing for recommendations in the ESHRE guidelines.

RECOMMENDED PHRASING	
Strong recommendation	Clinicians should/should not * It is recommended/it is not recommended * It is indicated/it is not indicated Do/Do not *
Weak recommendation	It is conditionally recommended* It is probably recommended* It is suggested* Clinicians might* Clinicians could consider Clinicians may/might consider
Good practice point (GPP)	The GDG recommends

** suggested by the GRADE working group*

Good practice points (GPPs) are not an alternative to evidence-based recommendations; any evidence relating to a key question excludes the possibility of using a GPP to make a recommendation. GPPs might be used to emphasise the importance of patient participation in decision making about specific procedures, provide advice on the management of specific surgical procedures for which there is an evidence based recommendation, or advise caution where there is perceived risk of harm but no available direct evidence of such harms (Scottish Intercollegiate Guidelines Network (SIGN), 2015).

If the GDG group feels strongly that they want to make a recommendation even though there is **no significant direct evidence**, however indirect evidence is available, it should be issued as a weak recommendation rather than a GPP. For such recommendations, the evidence may include expert opinions (from outside the GDG). If no such evidence exists, formal methods should be used to develop a consensus-based recommendation which will be clearly identified as such within the guideline by a statement accompanying the recommendation. The methods used to reach consensus should be described in the methodology annex of the guideline.

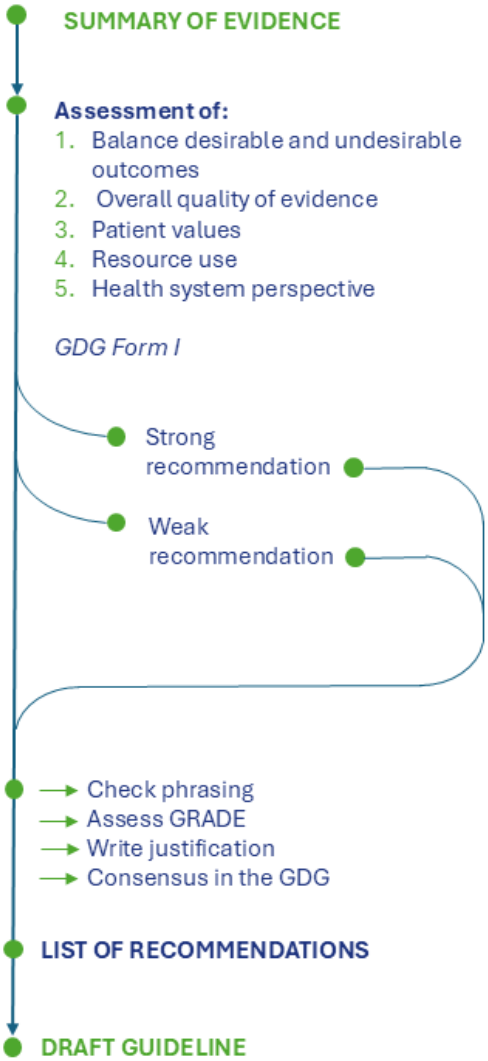
For newer diagnostic tests and interventions, the GDG is encouraged to clearly state whether the test/intervention is recommended in routine clinical practice or whether it can be used only in a research context. Further data could allow for a more comprehensive recommendation in the update of the guideline.

For some tests and interventions, the GDG may decide not to make a recommendation at all.

Recommendations for future research

In discussing the available evidence and its shortcomings, the GDG may feel it is important to formulate recommendations for future research. Where possible, research recommendations should be specific, detailing the necessary type of studies (RCTs, large multicentre studies), but also the (specific) interventions, or patient subgroups where further research is needed. The aim of formulating research recommendations is to stimulate research with a direct impact on future clinical recommendations. For this purpose, the GDG will also be asked to reach consensus on the research recommendations considered to have the highest priority.

Summary



Tips

→ If applicable, make provisions for formulating recommendations in situations where there is insufficient evidence or very low quality evidence (e.g. conditional recommendation with judgements laid out transparently, no recommendation if the guideline panel feels there is substantial risk that their decision may be wrong, recommend that the intervention be used in the context of research complemented by guidance for what are the best management options until further research becomes available).

→ Provide suggestions about whether the recommendations are appropriate to serve as performance measures/quality criteria (e.g. management options associated with strong recommendations based on high- or moderate-quality evidence are particularly good candidates for quality criteria, when a recommendation is weak, discussing with patients the relative merits of the alternative management strategies and appropriate documentation of this interaction may become a quality criterion).

→ Record or refer to the methodology used for recommendations' formulation.
- If no consensus is reached, describe the different views and options.
- Record benefits and harms considerations.

→ Recommendations should be specific and unambiguous.

Available forms/checklists:

ⓘ Framework for justification of recommendations

8. WRITING THE GUIDELINE DRAFT

Principles for writing

Once key questions are answered and there is consensus about the guideline's recommendations, the first draft version can be written. However, writing in committee requires prior agreement about the consistent use of terminology and writing style. ESHRE guidelines should be written in English and within a European scope. Furthermore, they should be comprehensive and flexible to allow adaptation to diverse settings and circumstances of clinical practice.

The use of paragraphs and headings are recommended to facilitate readers' navigation. Moreover, the use of tables, illustrations, figures and algorithms is encouraged. For guideline uniformity an ESHRE guideline is written according to an established structure.

The research specialist is responsible for merging the input of the different GDG members and adapting the content where needed to result in a consistent and well-structured guideline. Where appropriate, the co-chairs or other GDG members may review the guideline to ensure consistency and clarity.

Guideline structure

In general, an ESHRE guideline consists of 3 main parts: the introduction section, the key question-related part and the annexes.

The introduction section is subdivided into the introduction to the guideline, and a clinical introduction.

The introduction to the guideline includes the scope of the document based on the scoping checklist, information on target users, and details on how the guideline was developed (referring to an annex with the full methodology). If relevant, previous versions of the guideline are listed in this section.

Depending on the topic of the guideline, the clinical introduction contains disease definition and terminology, prevalence, variation in practice, provision of suboptimal care, personal and societal costs.

For the key question related part, the GDG determines a logical order for reporting the total collection of key questions with their recommendations. One such method is to follow the patients' pathway, starting with the clinical evaluation (e.g. symptoms, risk factors) followed by diagnostics, treatment options, follow-up, complications, information provision. Per key question, the following items are reported:

- Key question
- Evidence: a descriptive summary of the selected clinical evidence, with GRADE summary of findings tables (or a reference to the tables in annex).
- Recommendations: one or more recommendations, in boxes with appropriate GRADE.
- Justification: a summary of the relevant evidence (with most important limitations) and the considerations taken into account when determining the strength of the recommendations. In the case of non-consensus, practice statements about the different schools of thought should be recorded. Furthermore, the explanatory text gives room for considerations from ethical or legal perspectives.
- References.

If there are no recommendations for a certain section, a conclusion and justification can be written.

The annexes contain:

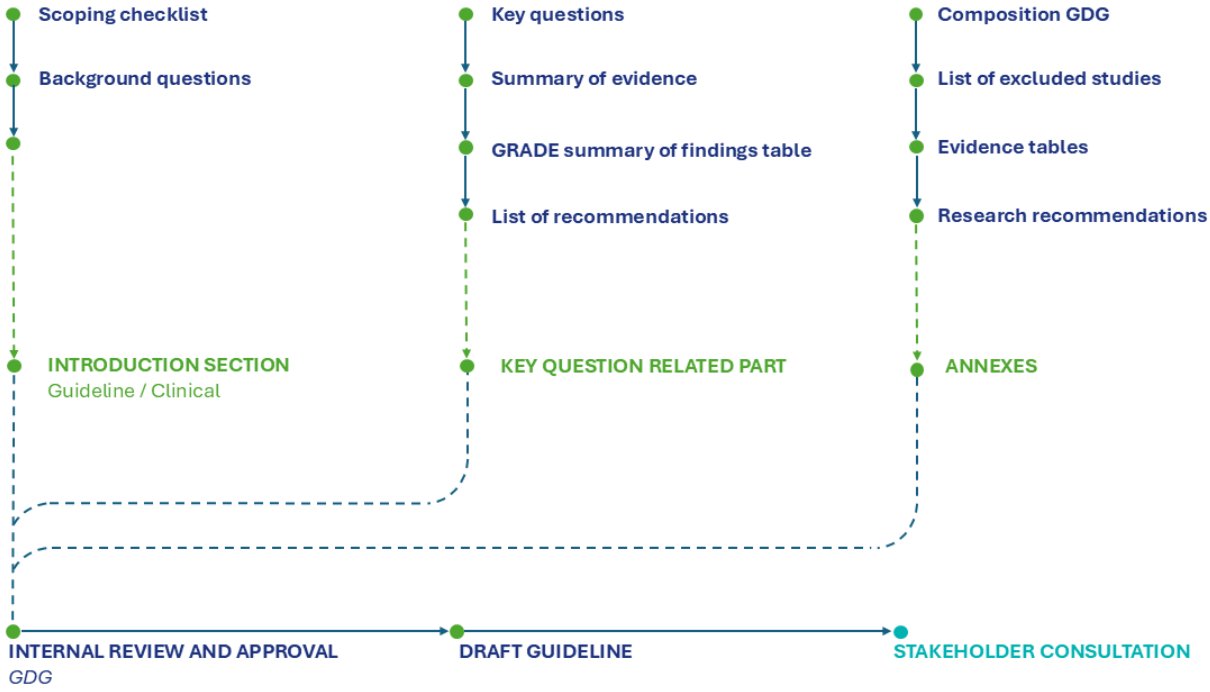
- List of abbreviations
- Glossary
- Summary of findings tables

- Evidence tables
- Details on the literature study: flowcharts, list of excluded studies
- GDG composition, with list of declared COI
- COI management and actions taken
- Research recommendations: describe gaps in scientific knowledge for future investigation
- Methodology (Appendix 5), that describes the guideline development in detail:
    - Guideline development process in 12 steps, according to the manual: funding, tailored information on the scoping procedure, details on the literature searches (searched databases, timeframe, inclusion and exclusion criteria), methodology of writing recommendations, and the used guidelines' manual version.
    - Information on the quality of evidence, grades of recommendations and phrasing, i.e. basic information needed to understand and interpret the recommendations.
    - Strategy for review of the guideline draft: invited reviewers, review deadlines, processing of comments.
    - Guideline implementation strategy, tailored to the guideline.
- Schedule for updating the guideline
- List of reviewers in stakeholder consultation
- Guideline versions and dissemination (e.g. existence of additional tools), if relevant.
- Relationship with other existing guidelines or ESHRE documents, if relevant.
- Key priorities for implementation, if relevant.

Any other relevant supporting material can be published as an annex to the guideline (survey results, tools, patient scenarios, etc.). A legal disclaimer is also added at the back of the cover of the guidelines.

A summary of all recommendations (condensed version) will be published in one of the ESHRE journals. This version will contain the most important content of the full guideline, at least all key questions and recommendations. Only GDG members will be listed as authors of this summary paper (fulfilling all ICMJE authorship criteria (see ICMJE recommendations)), in alphabetical order. The contribution of external invited reviewers will be acknowledged in the acknowledgements section.

Summary



Tips

→ Check if recommendations answer the key questions.

→ Use the AGREE Instrument (www.agreetrust.org) as a checklist.

→ Develop or adopt a standardised format for reporting the guideline, with specific structure, headings, and content.

→ Decide on the format(s) to be prepared for the guideline product(s) (e.g. full guideline, full guideline with technical report/systematic reviews, brief guideline for clinicians or policymakers, consumer version for patients) that will correspond to the dissemination plan.

→ Seek approval from all members of the GDG for the final document(s).

→ Link the recommendations explicitly to the supporting evidence.

→ Present the different management options clearly.

→ Present if appropriate expected exceptions for recommendation application.

→ Facilitate recommendation identification (e.g. bullets, numbering, boxes).

→ Discuss potential barriers in applying the recommendations.

→ Consider potential cost implications of applying the recommendations.

9. STAKEHOLDER CONSULTATION

The final stages of guideline development involve review by future users and approval by the parties involved. Within this phase the adequacy of the guideline document is evaluated, especially for its methodological quality, its clinical content and its applicability.

Review procedure

The review phase starts with a review of the guideline draft and annexes by several stakeholders. Their consultations concern in particular the guidelines' comprehensiveness, the accuracy of evidence interpretation and the acceptance of recommendations.

Following the final meeting of the GDG, a pre-registration period of 2-4 weeks will open for a sample of the target group, all members of the involved SIGs, patients' representatives (lay reviewers), and representatives of national organisations. This will be announced via email and on social media. During pre-registration, reviewers must declare their COI (Reviewer disclosure form Ⓙ). After this pre-registration period, the draft guideline will be published online and opened for stakeholder review. Pre-registered reviewers will be invited to submit their comments within six weeks. Following this procedure, results in an addition to the reviewers' list will be mentioned as an appendix to the guideline (Reviewer comments form Ⓚ). Potential stakeholders should note that ESHRE can accept only one set of comments per reviewer. Comments submitted by individuals who have not registered as reviewer, submitted outside the official review form, or accompanied by attachments such as research articles, letters, or leaflets, will not be considered as part of the stakeholder review. ESHRE also reserves the right not to respond to comments that are hostile, inappropriate, or otherwise unsuitable for constructive review. All comments should be solely aimed at improving the quality and clarity of the guidelines.

For adapted (parts of) or updated clinical guidelines, the ESHRE draft guideline should be also sent for review to the original developers.

The comments received from reviewers are tabulated and discussed in the stakeholder review report. The GDG will respond to each of the comments but does not commit to altering the recommendations in all instances. If no change is made, the reasons for this will be recorded in the report.

Any alterations to the recommendations must be made with the agreement of the whole GDG and noted in the report. The report with all the stakeholder comments with their responses will be published on the ESHRE website along with a summary on how the stakeholder input has influenced the final guideline.

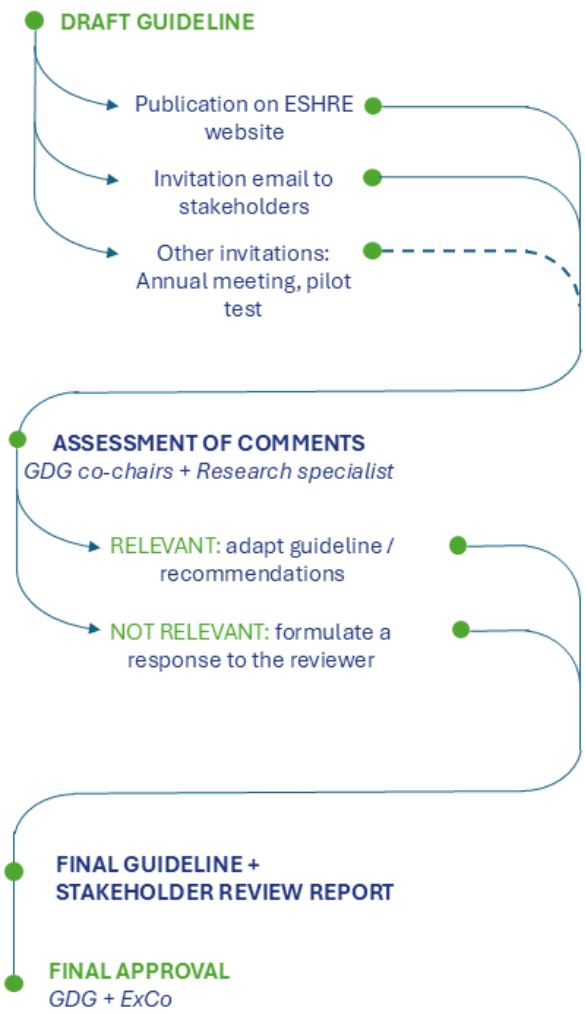
New sections or topics should not be added after stakeholder review unless essential changes are identified, in which case an additional stakeholder review must be conducted. Similarly, removal of evidence after stakeholder review should be avoided; if unavoidable, it must be clearly justified and transparently documented.

Additional options for stakeholder consultation

An additional open meeting at the ESHRE annual meeting is also an option for review. This provides the opportunity to present preliminary conclusions and draft recommendations to a wider audience and to hear valuable suggestions for additional evidence or alternative evidence interpretation. Because participation in such a meeting generates a sense of ownership across geographical and disciplinary boundaries, the organisation of such a meeting might accelerate the internal consensus process, the review procedure and final implementation.

A draft guideline can also be pilot tested before a wider launch. This step can detect problems in formatting, usability and acceptance.

Summary



Tips

→ Use the reviewing and piloting phase as an opportunity to advertise the existence of a new guideline.

→ Set a policy and process for handling consumer and stakeholder feedback and dealing with different perspectives (e.g. ensure that diverse perspectives are considered in making decisions, provide transparent rationale for judgements made, provide an appeal process for stakeholders, publish consultation comments and the guideline development panel's responses).

→ Record the stakeholder review reporting methodology, document the internal and external peer review process and, if applicable, publish consultation comments and the GDG responses.

→ Document the enrolment and selection of consumers and stakeholders for the guideline panel and the involvement and consultation with all other consumers and stakeholders to ensure explicit and transparent methods.

Available forms/checklists:

Ⓙ Reviewer disclosure form

Ⓚ Reviewer comments form

10. APPROVAL

Final version & authorisation

After stakeholder consultation, the entire GDG should formally approve the final version of the guideline.

After completion of all revisions, English language reviewers and proofreaders (and possibly lawyers) can be called upon when necessary.

The final step is to submit the guideline for formal approval by the Guidelines Committee and the ESHRE ExCO.

Summary

11. PUBLICATION, DISSEMINATION, TRANSLATION AND IMPLEMENTATION

Dissemination of ESHRE guidelines is considered as a continuation of the work of the GDG and involves making guidelines accessible, advertising their availability, and distributing them widely. Guidelines are (most) effective if their dissemination and implementation are carefully considered and vigorously pursued. If not, the time, energy and costs devoted to their development are wasted and potential improvements in reproductive healthcare may be missed.

A range of dissemination strategies can be effective, for instance the:

- Use of short summaries, such as one-page summaries and flowcharts
- Promotion of guideline's development/existence
- Publication in professional journal(s)
- Publication on the internet and links on related websites
- Explanation via webinars (organised by the relevant SIG)

Currently there is too little evidence to support decisions about which guideline dissemination strategy is efficient under which circumstances. In general, the use of multi-faceted dissemination strategies is recommended.

The standard dissemination procedure for all ESHRE guidelines comprises publishing and announcement.

Publications

The document will be published in at least two formats:

- Publication of the full guideline on the ESHRE website (https://www.eshre.eu/Guidelines-and-Legal)
- Publication of the summary guideline (including all recommendations) in one of the ESHRE journals. After publication, a link to the paper is added to the guideline page. Only members of the current GDG can be listed as authors on the summary paper, in alphabetical order.

Additional options

Distribution of guidelines alone has been shown to be ineffective in achieving change in practice; guidelines are more likely to be effective if they are disseminated by a strategy based on barrier research, by an active educational intervention, or by patient-specific reminders. However, the potential clinical benefits and resources required to introduce guidelines and the likely benefits and costs because of any provider's behaviour change need to be carefully considered before developing additional tools. The efficiency of a dissemination strategy is best evaluated in the presence of different barriers and effect modifiers.

Two more options include the development of algorithms and the creation of patient information leaflets:

- An algorithm is a flow chart of the clinical decision pathway described in the guideline, in which process steps and decision points are linked by arrows. Whenever relevant and useful, these flowcharts can be digitalised in decision-support tools, available through web-based platforms or smartphone/tablet applications.

- Patient information leaflets summarise the recommendations mentioned in the ESHRE guideline in layman's terms. It aims to help patients understand the guideline's recommendations and facilitates decision-making. Moreover, guideline leaflets may be used by hospitals or patient organisations for developing their own information leaflets. Patient versions of guidelines will be developed in collaboration with the patient representative involved in the guideline, and its accuracy and correctness will be checked by the co-chairs, or a delegated GDG member. If possible, a review of the patient version will be organised by inviting all relevant patient organisations to send comments. The final version will be distributed among all relevant patient organisations with an invitation to endorse it, and if necessary, translate it. The translated version should be reviewed by a GDG member from the respective country or by that country's national representative. It should also contain a disclaimer, provided by ESHRE, stating that the English version remains the reference document.

Dissemination

All relevant ESHRE communication channels will be used to announce the release of a new guideline:

- A newsflash on the ESHRE website's homepage
- A mailing to the members of the relevant SIGs, or all ESHRE members
- An announcement in "Focus on Reproduction"
- Promotion at the annual ESHRE meeting via different media. Optionally, participants will be informed about the development and release of new guidelines during a specific guideline session.
- A mailing to all related National Societies to inform them about the guideline release. They are asked to encourage local implementation by, for instance, translations or condensed versions, but they are also offered a website link to the original document.
- All appropriate remaining stakeholders for instance, European policy makers, patient societies and industry representatives will be separately informed.

Guideline translation and endorsement

An important factor in facilitating guideline implementation is endorsement by professional groups. Endorsement of ESHRE guidelines is systematically sought from relevant National Societies by informing their presidents.

ESHRE and the guideline group members put significant effort into developing evidence-based guidelines. Furthermore, we try to involve different nationalities in the guideline groups, and to organise a broad stakeholder review. By doing so, we provide guidelines designed for application to a broad population which can ideally be endorsed by National Societies, and if wished upon, translated verbatim.

National Societies and organisations can request permission to translate (verbatim) one of the guidelines, or any specific content related to the guidelines, in their language. For an official ESHRE approved translation, a straightforward 4-step procedure of approval, translation, validation and publication is outlined in a policy (see below).

For reasons of consistency only one translation of a certain ESHRE guideline in any given language is accepted by ESHRE. ESHRE reserves the exclusive right to publish the first edition of all ESHRE guidelines and post its translation on the ESHRE website. National Societies must secure copyright protection in their own country.

When a verbatim translation of a guideline is insufficient for national uptake, ESHRE will allow National Societies to use the ESHRE guidelines as the basis of their national guidelines. To ensure transparency, the methodology should clearly refer to the ESHRE Guideline (including the weblink https://www.eshre.eu/Guidelines-and-Legal) and state how the ESHRE guideline was used, including which text blocks / search strings have been used from the ESHRE guideline and for which topics the recommendations differ significantly between the documents. The resulting document will be considered a national guideline, not an ESHRE guideline.

ESHRE gives National Societies and organisations the optional right to publish the translated guideline in their own national journals. All costs of carrying out these rights and of translating the guideline are for the National Societies.

The above information applies only to documents to which ESHRE holds the copyright. For translation of the summary guideline published in one of the ESHRE journals, permission from Oxford University Press (OUP) should be requested.

Background information

Policy for translation of ESHRE® Documents

Please note that this policy sets out general rules regarding the translation of ESHRE® Documents (as defined below). Depending on the type of ESHRE® Document, specific provisions might also be applicable (as is for example the case for the ESHRE® guidelines). In case of a conflict between the provisions of this policy and specific provisions, the latter shall prevail.

Definitions:

ESHRE®:	The international non-profit organisation "European Society of Human Reproduction and Embryology" with its registered office at BXL7 Building 1 Nijverheidslaan 3 (1st floor) |1853 Strombeek-Bever, Belgium, VAT BE-0430.069.888, RLE Brussels;
ESHRE® Document(s):	Any document, produced and published by ESHRE®, to which ESHRE® exclusively possesses all rights of ownership. The English version always entails the original version of the document.;
National Society:	An association or organised group of professionals/patients with a significant background/experience in the field of (in)fertility.

Translation of ESHRE® Documents:

In summary, the following four steps must be followed in case of translation of an ESHRE® Document:

1. Request written permission of ESHRE® before endeavouring translation
2. Make an exact translation and ensure that the ESHRE® copyright statement and the ESHRE® disclaimer are foreseen on the document, as well as full reference to the ESHRE® Document
3. Request written validation of the translation from ESHRE®
4. Ensure that the translation is up-to-date and corresponds to the latest version of the ESHRE® Document

1. Prior permission to translate

A National Society shall have the right – at its own cost – to translate ESHRE® Documents and publish the translations thereof in its own country upon (i) prior written approval of ESHRE® and (ii) full endorsement of the corresponding parent ESHRE® Document.
For reasons of consistency, ESHRE® shall accept only one translation per ESHRE® Document in any given language. At all times, ESHRE® retains full (copy)rights whatsoever on every ESHRE® Document and its translations.

2. Obligations for the translators and the National Society

General
All costs and expenses relating to the translation of an ESHRE® Document (including the cost of compensating translators) shall be borne by the National Society exclusively.

The National Society ensures that every translator transfers all rights whatsoever (which the latter might possibly possess with respect to the performed translation) to ESHRE®.

The National Society shall be responsible for the exact translation of the ESHRE® Document by the translator it appeals on. Each translation shall contain all textual, pictorial and diagrammatic material, as foreseen in the

ESHRE® Document, without any alterations. Footnotes or annexes may be added to highlight national and/or regional practices. In no event, amendments to the original text shall be allowed.
Further, the National Society (and the translator it appeals to) undertakes to:

(i) give full credit to ESHRE® for the ESHRE® Document by including on the title page of the translated document:
  (a) the ESHRE® copyright statement (as mentioned below),
  (b) the ESHRE® logo,
  (c) full reference to the original publication of the ESHRE® Document on ESHRE's® website and in ESHRE's® official journals ('Human Reproduction');
(ii) foresee the appropriate ESHRE® disclaimer, as mentioned below, in the translated document;
(iii) mention in the title of the translated document the name of each ESHRE® working group member who is (co-) author of the ESHRE® Document; and
(iv) clarify in the (sub)title of the translated document that it entails a translation of an ESHRE® Document, whereby the full title of the parent ESHRE® Document needs to be mentioned.

Whenever possible, a back-to-back translation is recommended.

The National Society that produces a translation of an ESHRE® Document may foresee the translated document of its own logo(s) and additional information about its society. The names of the translators, reviewers and/or other people involved in the translation of the ESHRE® Document, can also be foreseen on the translated document, if it has been made clear they were solely involved in the translation of the ESHRE® Document and thus took no part in the production and publication of the ESHRE® Document.

Translation sponsored by companies

In case a National Society obtains sponsoring from commercial organisations to finance the translations of ESHRE® Documents, it shall be strictly prohibited to foresee in any kind of product advertising on the translated document.

However, corporate logos of the sponsoring company(ies) in question can be displayed with the following statement*: 'The translation of this ESHRE® document was made possible through an educational grant from [name sponsor]. [Name sponsor] acknowledges explicitly that it was not involved in the actual production and publication of the parent ESHRE® document, hence influenced in no way the content thereof.'*

3. Validation of the translation

All documents translated in line with the above can only be published upon prior written validation of ESHRE®. Such validation shall:

(i) Be organised by the ESHRE® central office;
(ii) be performed by a native speaker from the ESHRE® working group or the committee of national representatives; and
(iii) only relate to the translation itself and in no case entail a review of the content, meaning that ESHRE® shall not verify if the scientific value of the parent ESHRE® Document has been preserved in the translated document.

ESHRE® strives to inform the National Society on the outcome of the performed validation within four weeks upon receipt of the translation by ESHRE®.
Validated translations of ESHRE® Documents will be published by ESHRE® on its website, upon prior written approval of the respective National Society.

4. Keep the translation up to date

It is the responsibility of the National Society to ensure that the translated document is kept up-to-date and corresponds to the latest version of the parent ESHRE® Document.

ESHRE® strives to inform the National Society of any updates on the parent ESHRE® Document, and this within due time.

ESHRE® copyright statement

"Copyright © European Society of Human Reproduction and Embryology ('ESHRE'®) – All rights reserved"

ESHRE® disclaimer

"This publication entails a translation of an original ESHRE® document – as fully referred to on the title page of this document – whereby such translation was performed in line with the provisions of the 'Policy for the translation of ESHRE® Documents', which is available on the ESHRE® website (www.eshre.eu).

The translation of the original ESHRE® document is made by and under supervision of [name of the National Society], which is solely responsible for the content of this translation. Prior validation by ESHRE® of this translation does not affect such responsibility.

If any questions arise related to the accuracy of the information contained in the translation and/or its scientific value, please refer to the original ESHRE® document. Any discrepancies or differences created in the translation are not binding to ESHRE® and shall have no legal effect for compliance or enforcement purposes. The English version, being the language in which the original ESHRE® document is published, shall always prevail."

Implementation and evaluation

Guidelines do not implement themselves. Local ownership of the implementation process is crucial for changing practice. ESHRE is responsible for the development of European guidelines and their implementability, but not directly for their implementation into local practice. Nevertheless, the identification of barriers to guidelines' acceptance is one of the first steps of an implementation process and has ideally been part of the guideline developmental phase. Instruments like 'The Guideline Implementability Appraisal instrument' can be helpful for identifying obstacles to guideline implementation (Shiffman et al., 2005) (http://nutmeg.med.yale.edu/glia/doc/GLIA_v2.pdf).

There are different types of barriers to guideline implementation:

- internal to the guideline itself
- factors related to the individual care providers (e.g. attitude and skills)
- factors related to the (social) setting (e.g. patients' and colleagues' characteristics)
- external factors related to the system (e.g. reimbursement).

After the determination of factors affecting guideline adoption, the currently recommended approach is to plan a targeted intervention. However, there is no specific guidance available for translating identified barriers into tailor-made implementation interventions. Each implementation strategy is effective under certain circumstances, and a multifaceted approach is more likely to succeed than a single approach. Evaluation of such complex interventions is therefore important and mostly undertaken by investigators with research funding.

Focusing on individual recommendations rather than on the guideline as a whole, makes the implementation initiative more manageable. Criteria reflecting one or more of the six quality domains defined by the Institute of Medicine (safety, effectiveness, patient-centeredness, timely, efficiency and equitability) can help to prioritise guideline's recommendations for this purpose.

At an appropriate time after dissemination and implementation an evaluation is necessary for insight into the impact of the guideline. Such an evaluation consists of several components, namely an assessment of:

- Guideline dissemination

- Change in practice performance
- Change in health outcomes
- Change in consumer's knowledge and understanding
- Economic consequences

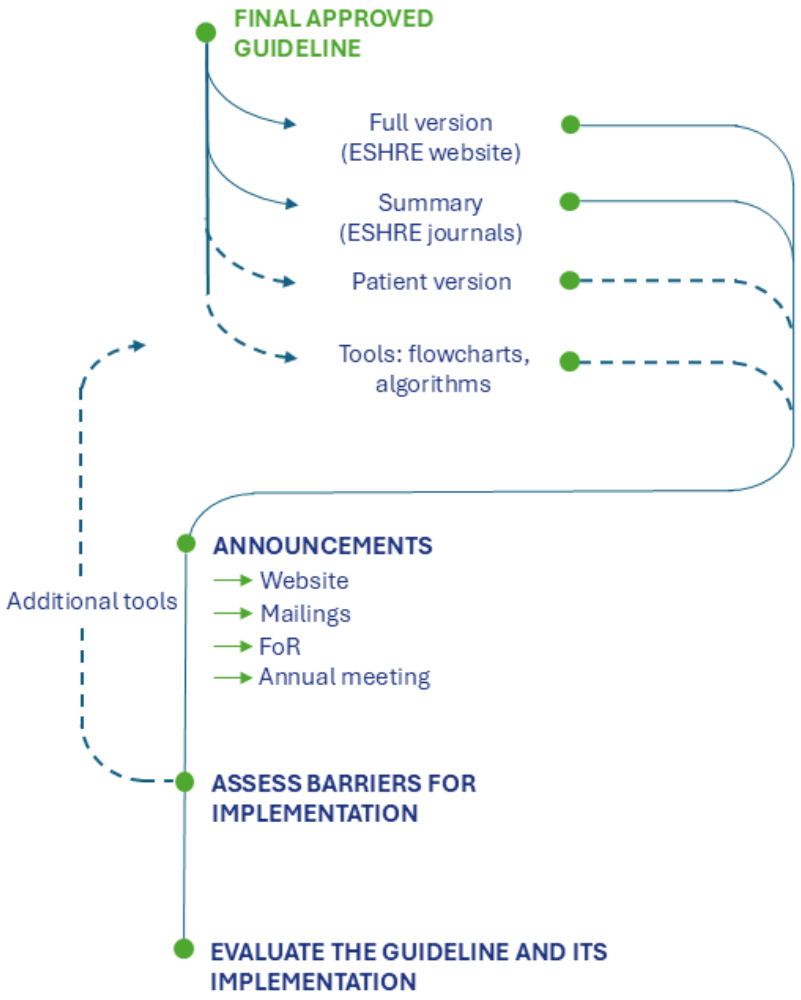
Practice performance is usually measured by a clinical audit and indicators. The frequently used definition for an indicator is "a measurable element of practice performance for which there is evidence or consensus that it can be used to assess the quality and hence change in the quality of care provided". Based on the manual by the Agency for Healthcare Research and Quality (AHRQ) and additional literature on quality indicators in infertility, a set of quality indicators for each ESHRE guideline can be developed in a 3-step process:

- The GDG members rank the recommendations on priority for implementation to obtain key recommendations.
- The GDG members propose quality indicators for each key recommendation.
- The GDG members determine the importance and the preparedness to measure for each quality indicator, in a stepwise process.

For each step specific structured questionnaires should be developed.

The resulting set of key quality indicators will be used to evaluate the quality of care and the impact of the guideline on the quality of care within Europe.

Summary



Tips

→ Develop or adapt tools, support, and derivative products to provide guidance on how the recommendations can be implemented into practice (e.g. mobile applications, integration with clinical decision support systems, make the guideline adaptable as an educational resource for target audiences to support educational outreach).

→ Consider how the guideline can be adapted and provide specific instructions for how target end users can adapt the guidelines to other contexts in a systematic and transparent way (e.g. modifying a recommendation based on local resources and baseline risk, or by addressing implications that deviate from the judgements made by the guideline panel).

→ Conduct an internal evaluation (i.e. self-assessment) of the guideline development process, including the guideline panel meeting(s) held to formulate recommendations, by asking guideline group members for feedback.

→ Consider pilot testing the guideline with the target end users (e.g. with members of target audience and stakeholders who participated in the GDG).

→ Provide criteria and tools for target end users to monitor and audit the implementation and use of the guideline recommendations (e.g. identify outcomes that should change with implementation and suggest methods for measuring the outcomes).

→ Provide support and tools for prospective evaluation of the guideline to determine its effectiveness after implementation (e.g. using randomised evaluations where possible, using before-after evaluations cautiously due to uncertainties regarding the effects of implementation).

→ Consider the potential involvement of the GDG in prospective evaluation(s) of the guideline (e.g. partnering with organisations that implement the guideline to plan evaluation studies).

→ Plan to collect feedback and evaluations from users to identify how to improve the intrinsic implementability of the recommendations in subsequent versions of the guideline.

→ Support the guideline with application tools and record those within the guideline

12. UPDATING THE GUIDELINE

Updating of guidelines is an essential part of guideline development, to ensure guidelines remain current and their quality is maintained (Agbassi et al., 2014). New evidence is emerging constantly, and recommendations may be compromised when they are no longer reflective of the current clinical evidence. An analysis of NICE guidelines showed that the median life span of the clinical guidelines was 60 months (95% CI 51 - 69) (Alderson et al., 2014). The authors also suggested a system of monitoring to detect guidelines that are quickly outdated. As new evidence is published at different rates in different fields, a full revision of guidelines after a fixed time period is not always appropriate. Recent publications on guideline updating propose systems of monitoring and assessment of the need for a review, before undertaking a formal update (Agbassi et al., 2014).

Guideline monitoring

Guideline monitoring includes guideline assessment, streamlined systematic review and formal update. Guideline assessment aims to identify documents that are no longer applicable to current practice and documents with recommendations that may have been, have the potential to be, or are expected to be invalidated by new evidence. Guideline assessment can be performed by the document assessment questionnaire (6 questions, see tool Ⓛ).

The application of the document assessment questionnaire classifies each guideline into one of the following four groups: (1) endorse, (2) defer, (3) review, or (4) archive the guideline.

Only guidelines categorised as review from the assessment are eligible for the next step.

Guideline update process

Initiating the guideline update process

If a need for review is identified for one or more PICO questions (partial review), or the full guideline (complete review), the research specialist will notify the responsible SIG coordinator, who will complete the application form. Approval for the update must then be requested from the Guidelines Committee, which will prepare .an assessment report for final approval by the ExCO. A complete review, approved by the ESHRE ExCO, will follow the standard process described in this manual.

Forming guideline development group

GDG members, including the co-chairs, may serve a maximum of two consecutive terms in the development of a single guideline. After a minimum interval of one guideline version, a former GDG member may rejoin the development group for the same guideline.

To ensure both continuity and renewal, some members will be replaced after one term, allowing new experts to join the GDG while maintaining institutional knowledge within the group.

Members of the previous guideline group will be invited to submit an application to be part of the guideline group for the update of the guideline. In addition, an open call will be published on the ESHRE website and social media platforms to invite ESHRE members to submit an application. The Guidelines Committee will review all applications and COI of the aspiring GDG members, and make a proposal for the GDG to be approved by the ESHRE ExCO.

Scoping the guideline

For updates, the guideline scope needs to indicate which sections will be updated and any changes from the current guideline. The revised scope and key questions will be published for stakeholder comments, in survey form, for 2 weeks.

Formulating key questions

When updating an evidence-based guideline, the key questions are reviewed by the GDG and may need refining. In any case, the entire GDG should be informed of and agree with any changes to the key questions.

Changes to the key questions could include:

- Reassessment of the importance of the outcomes
- Addition of an outcome: for instance, the importance of an outcome (e.g. a serious adverse effect) may only become known after the evidence synthesis
- Addition of an intervention that is relevant for the flow and consistence of the guideline
- Specific search on an important subgroup, not defined as such in the PICO questions
- Merging of key questions based on significant overlap of the evidence synthesis.

The revised scope and key questions will be published for stakeholder comments, in survey form, for 2 weeks.

Searching evidence

When high-quality review-level evidence is available on a given topic, previously used search terms may be reused. This approach allows robust existing evidence to be retained while incorporating newly published evidence into the updated guideline. Hereto, the original review searches are re-run and expanded to account for any differences in scope and inclusion criteria between the original review and the update.

Evidence synthesis

In many cases, the evidence reviews will be an update of a previous review, to include more recently published evidence. In these cases, a judgement should be made on what elements of the previous review can be reused, and which need to be redone, based on the level of similarity between the original and new review questions, protocols and methods.

The clinical expert (or the entire GDG) reviews the new evidence to determine:

(1) if it supports or contradicts current recommendations;
(2) if the current recommendations cover all relevant subjects addressed by the new evidence, and
(3) if strong evidence that may change the current recommendations is expected to be published in the near future.

Based on the assessment for each PICO question, it is categorised as endorse, or update. Endorse means that the newly identified evidence supports the current recommendations with only minor changes or new qualifying statements. If no changes are made to the recommendations, this will be reflected by adding the year of the original version of the recommendation in brackets (for example [2019]). Update means that the new evidence requires changes to the existing recommendations, reflected by [updated] added at the end of the recommendation; or the formulation of new recommendations, indicated by adding the year of the update at the end of the recommendation (for example [2026]). In some instances, the GDG can decide to re-formulate a recommendation, even in the absence of new evidence, to promote a better understanding of the recommendation. This will be reflected by adding [reworded] to the recommendation.

Changes to the included evidence will be documented either in the body of the guideline with footnotes or in the annex of the literature search results, which provides more information on the literature searches.

Stakeholder consultation

Updated guidelines are also subject to consultation and will follow the usual validation process.

Approval

After stakeholder consultation, the entire GDG should formally approve the final version of the updated guideline before submitting it to the Guidelines Committee and the ESHRE ExCO for approval.

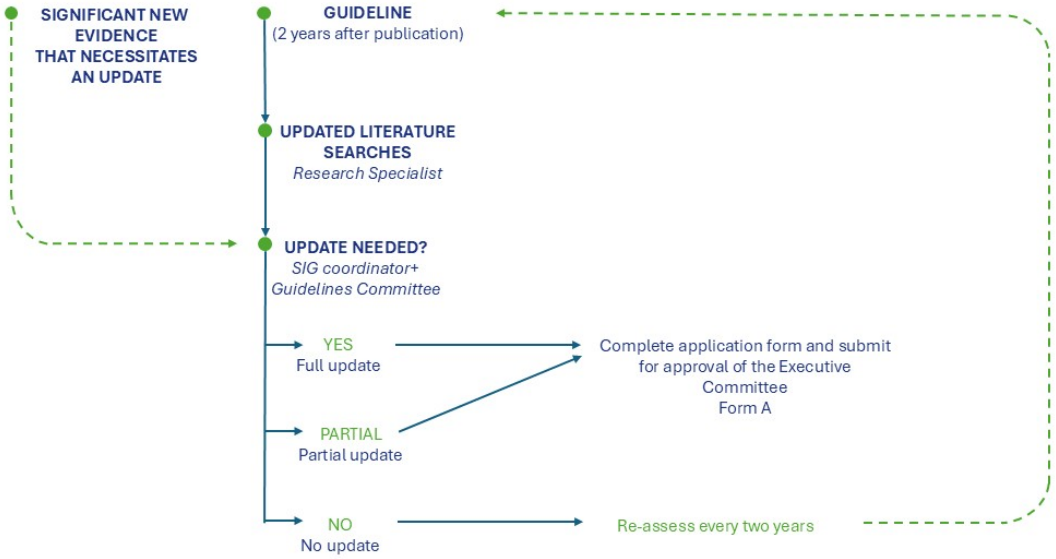
Publication

The latest approved version of the guideline will be available on the ESHRE website (https://www.eshre.eu/Guidelines-and-Legal). A change log, documenting all changes except for typographical corrections, along with version history, will be provided on the final page of the guideline. Older versions can be requested by contacting guidelines@eshre.eu.

The updated guideline will also be published as a summary paper (including all recommendations) in one of the ESHRE journals. Authorship will be granted only to the current GDG members (fulfilling all ICMJE authorship criteria (see ICMJE recommendations)). The contribution of authors of the original guideline who are involved in the final revision of the updated guideline as invited experts will be acknowledged in the acknolegements section, together with the contributions of external invited reviewers.

Upon publication of an updated guideline, the previous version is withdrawn.

Summary



Tips

- ➔ Plan the logistics for updating the guideline in the future
- ➔ Refer to the procedure for guideline updating.

Available forms/checklists:

Ⓐ Application form

Ⓛ Document assessment tool

Reference list

Agbassi C, Messersmith H, McNair S, Brouwers M. Priority-based initiative for updating existing evidence-based clinical practice guidelines: the results of two iterations. *Journal of clinical epidemiology* 2014;67: 1335-1342.

AGREE Reporting Checklist. 2016.

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: 52-55.

Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, Nasser M, Meerpohl J, Post PN, Kunz R *et al.* GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *Journal of clinical epidemiology* 2013a;66: 719-725.

Andrews JC, Schunemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, Rind D, Montori VM, Brito JP, Norris S *et al.* GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. *Journal of clinical epidemiology* 2013b;66: 726-735.

Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S *et al.* GRADE guidelines: 3. Rating the quality of evidence. *Journal of clinical epidemiology* 2011;64: 401-406.

Brouwers MC, Kerkvliet K, Spithoff K, Consortium ANS. The AGREE Reporting Checklist: a tool to improve reporting of clinical practice guidelines. *bmj* 2016;352: i1152.

The Cochrane Collaboration (2011) Cochrane Handbook for Systematic Reviews of Interventions. 2011.

Cook DJ, Greengold NL, Ellrodt AG, Weingarten SR. The relation between systematic reviews and practice guidelines. *Annals of internal medicine* 1997;127: 210-216.

Darzi A, Abou-Jaoude EA, Agarwal A, Lakis C, Wiercioch W, Santesso N, Brax H, El-Jardali F, Schunemann HJ, Akl EA. A methodological survey identified eight proposed frameworks for the adaptation of health related guidelines. *Journal of clinical epidemiology* 2017;86: 3-10.

den Breejen EM, Hermens RP, Galama WH, Willemsen WN, Kremer JA, Nelen WL. Added value of involving patients in the first step of multidisciplinary guideline development: a qualitative interview study among infertile patients. *International journal for quality in health care : journal of the International Society for Quality in Health Care / ISQua* 2016;28: 299-305.

Fervers B, Burgers JS, Haugh MC, Latreille J, Mlika-Cabanne N, Paquet L, Coulombe M, Poirier M, Burnand B. Adaptation of clinical guidelines: literature review and proposition for a framework and procedure. *International journal for quality in health care : journal of the International Society for Quality in Health Care / ISQua* 2006;18: 167-176.

GIN-McMaster Guideline Development Checklist.

Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, Devereaux PJ, Montori VM, Freyschuss B, Vist G *et al.* GRADE guidelines 6. Rating the quality of evidence--imprecision. *Journal of clinical epidemiology* 2011a;64: 1283-1293.

Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Falck-Ytter Y, Jaeschke R, Vist G *et al.* GRADE guidelines: 8. Rating the quality of evidence--indirectness. *Journal of clinical epidemiology* 2011b;64: 1303-1310.

Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Glasziou P, Jaeschke R, Akl EA *et al.* GRADE guidelines: 7. Rating the quality of evidence--inconsistency. *Journal of clinical epidemiology* 2011c;64: 1294-1302.

Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, Alonso-Coello P, Djulbegovic B, Atkins D, Falck-Ytter Y *et al.* GRADE guidelines: 5. Rating the quality of evidence--publication bias. *Journal of clinical epidemiology* 2011d;64: 1277-1282.

Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, Atkins D, Kunz R, Brozek J, Montori V *et al.* GRADE guidelines: 9. Rating up the quality of evidence. *Journal of clinical epidemiology* 2011e;64: 1311-1316.

Hilbink MA, Ouwens MM, Burgers JS, Kool RB. A new impetus for guideline development and implementation: construction and evaluation of a toolbox. *Implementation science : IS* 2014;9: 34.

Institute of Medicine Committee on Clinical Practice Guidelines. In Field MJ and Lohr KN (eds) *Guidelines for Clinical Practice: From Development to Use*. 1992. National Academies Press (US)

Copyright 1992 by the National Academy of Sciences., Washington (DC).

Moses RE, Feld AD. Legal risks of clinical practice guidelines. *The American journal of gastroenterology* 2008;103: 7-11.

Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J, Marteau T. Consensus development methods, and their use in clinical guideline development. *Health technology assessment (Winchester, England)* 1998;2: i-iv, 1-88.

Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. *ACP journal club* 1995;123: A12-13.

Schünemann HJ, Osborne M, Moss J, Manthous C, Wagner G, Sicilian L, Ohar J, McDermott S, Lucas L, Jaeschke R. An official American Thoracic Society Policy statement: managing conflict of interest in professional societies. *Am J Respir Crit Care Med* 2009;180: 564-580.

Schünemann HJ, Wiercioch W, Etxeandia I, Falavigna M, Santesso N, Mustafa R, Ventresca M, Brignardello-Petersen R, Laisaar K-T, Kowalski S *et al.* Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise. *Canadian Medical Association Journal* 2014;186: E123-E142.

Scottish Intercollegiate Guidelines Network (SIGN). *SIGN 50 : A guideline developer's handbook.* , 2015 (SIGN publication no. 50). [November 2015]. Available from URL: http://www.sign.ac.uk, Edinburgh.

Shiffman RN, Dixon J, Brandt C, Essaihi A, Hsiao A, Michel G, O'Connell R. The GuideLine Implementability Appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. *BMC Medical Informatics and Decision Making* 2005;5: 23.

The GRADE Working Group. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. . In Schünemann H BJ, Guyatt G, Oxman A, (ed). 2013.

Whiting P, Rutjes AW, Dinnes J, Reitsma J, Bossuyt PM, Kleijnen J. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health technology assessment (Winchester, England)* 2004;8: iii, 1-234.

Forms

- Ⓐ Application form
- Ⓑ Disclose form
- Ⓒ Confidentiality form
- Ⓓ Scoping checklist
- Ⓔ PICO checklist
- Ⓕ Relevance and quality check
- Ⓖ Quality assessment checklist
- Ⓗ Evidence tables
- Ⓘ Framework for recommendations
- Ⓙ Reviewer comments form

A. APPLICATION FORM

This application form can be used to request support of ESHRE for a project, including Guidelines or Good practice recommendations. The completed application form should be sent to guidelines@eshre.eu.

Applications will be reviewed by the Executive Committee. The aim is to provide feedback within 6 weeks after submission of the form. Further information is available in the ESHRE Internal rules (provided with this document).

Summary

Proposed title:		
Type of application:	☐	**Guideline**
	☐	**Good practice recommendations paper**
	☐	**Position paper**
	☐	**Other (ESHRE) project** *[please specify]*
Contact person: (name and email)		

Section 1: Information on the Applicants

ESHRE Special Interest Group(s) /Committee/working group involved (if any):		
Project lead (or chair)[4,5]:		
[name]	*[country]*	*[Affiliation]*
Project team: **(one row per team member)**		
[name]	*[country]*	*[Affiliation]*

[1] For guidelines and recommendations, the composition of the group and the appointment of the co-chairs of the group will be the responsibility of the Guidelines Committee.
[5] The project lead or chair is also the main contact person. On the resulting publication, s/he will be the corresponding author, and the first or last author (upon preference). The role can be split among 2 co-chairs.

Section 2: Information on the Project

Background and rationale[6] (500 words max)

Aims and objectives (300 words max)

Expected benefits of the project (500 words max)

Link to other ESHRE projects/documents (300 words max)

Expected outcomes in terms of documents, publications, etc. (500 words max)

Annexes

[Annexes can be added detailing reference lists, protocols, examples. If any annexes are added, please add here a short description on what is included]

Section 3: Methodology, timeline and resources

Proposed methodology[7] (500 words max)

Timeline (300 words max)

[6] In case of an application for a guideline or good practice recommendation paper, please include information on the relevance of the topic, and actual variation in practice.

[7] In case of guidelines and good practice recommendations papers, the methodology will follow the procedures described in the respective manuals.

Resources[8]	
• **Technical/administrative/project management support**	*Yes/No* *[If yes, please explain the tasks for which support is needed]*
• **Support for meetings**	*Yes/No* *[If yes, please explain]*
• **Financial support**	*Yes/No* *[If yes, please describe the budget and for what the money will be used]*
• **Other support**	*Yes/No* *[If yes, please explain]*

Section 4: Additional information

Any further information related to the project

DECISION (to be completed by the Executive Committee upon review and discussion)

Conclusion:	*Approved/not approved/returned for revision*
Feedback:	
Resources awarded:	
Conditions[9]:	

[8] In case of guidelines and good practice recommendations papers, the resources required will be defined by the central office.

[9] In case the significant resources from ESHRE are requested, the Executive Committee can award the support under conditions in terms of publication in an ESHRE journal, or specific acknowledgement.

B. DISCLOSURE FORM

All ESHRE Guideline Development Group members are expected to provide completed and signed disclosure statements about all financial, personal, or professional relationships with industry, individuals or organisations to avoid the perception of a conflict of interest. Our work relies on trust, independence and transparency. Having an interest does not automatically prevent participation. Declaring interests allows us to manage them proportionately and openly.

Guideline title	

Before completing this form, please carefully read the ESHRE manual for Guideline Development including the instructions for completing this form. The information provided will be treated confidentially by the ESHRE Guidelines Committee for purposes of guideline development and not shared with any third-party without your explicit consent. This form should be completed during application and submitted to the ESHRE central office (guidelines@eshre.eu). Updates should be made annually, before the in-person meeting and if changes occur during the guideline development process.

Having an interest does not automatically disqualify you or limit your participation in the guideline development. Your answers will be reviewed by the Guidelines Committee to determine whether you have a conflict of interest relevant to the topic(s) of the guideline. If you are unable or unwilling to disclose the details of an interest that may pose a real or perceived conflict, you must disclose that a conflict of interest may exist and the Guidelines Committee may decide that you be totally recused from the meeting or work concerned, after consulting with you.

Contact information of the guideline development group member

Please complete the table below and indicate if you wish to be considered for a leadership role.

Name: ……

Institution, Address: ……

……

E-mail address: ……

Area(s) of expertise relevant to this work ……

☐ **I would like to be considered for a co-chair position**

Please disclose your direct financial interests (past 36 months), direct non-financial interests (past 12 months), and any anticipated interests during the guideline development period (~24 months).

Direct financial interests

(personally held by the individual and directly related to the guideline topic)

☐ I have no direct financial interests from the past three years to report, nor any anticipated during the development of this guideline.

☐ I have the following direct financial interests from the past three years or anticipated during the development of this guideline, to report (please complete the table below):

Type of interest	Description of the interest			Time period (from MM/YYYY-to MM-YYYY)
	From which company, organisation, or institution	Range of the amount of income or value of interest (≤€1000; €1001-€5000; €5001-€10000; >€10000) If not disclosed, assumed to be significant	Relevance to the guideline topic (please specify)	
Commercial business (proprietorships, partnerships, joint ventures, board memberships, etc.)				
Ownership (Stocks, bonds, stock options, etc.)				
Salary or position funding directly linked to activities related to the guideline topic				
Patents, royalties, or intellectual property				
Directorship or consultancy fee(s) (advisory boards, legal testimony, etc.)				
Honoraria, speaker fees, teaching payments, paid event attendance.				
Expert testimony or legal advice				
Personal research funding or grants including institutional grants on which you are explicitly named as recipient or investigator				
Travel, accommodation, or other -in-kind support				
Other financial benefits related to the topic of this guideline and not captured by the checklist.				

Note: Institutional grants on which the individual is not explicitly named as a recipient, investigator, or co-investigator are considered indirect financial interests and should be declared in the indirect interest's section.

Direct non-financial interests

(non-financial, professional, academic, or personal interests personally held and directly related to the guideline topic)

☐ I have no direct non-financial interests from the past 12 months to report, nor any anticipated during the development of this guideline.

☐ I have the following direct non-financial interests from the past 12 months or anticipated during the development of this guideline, to report (please complete the table below):

Type of interest	**Description of the interest**		**Time period (from MM/YYYY-to MM/YYYY)**
	From which company, organisation, or institution	**Nature of the relationship/ Relevance to the guideline topic**	
Advocacy or membership in lobbying or pressure group			
Unpaid board membership or leadership roles			
Office or position of authority in a professional organisation			
Involvement in an ongoing or scheduled trial/ research related to the topic of the guideline			
Published topic-related opinions			
Authorship or co-authorship of evidence relevant to one of the topics of the guideline			
Leadership or participation in related guideline development elsewhere			
Other non-financial benefits related to the topic of this guideline and not captured by the checklist.			
Note: Institutional affiliations are considered indirect non-financial interests unless the individual holds a leadership or topic-specific role directly related to the guideline topic.			

Indirect interests

(financial and non-financial interests not personally held and/ or not directly related to the guideline topic, including institutional, employer, and family-related interests)

- Do any third-party closely associated with you (e.g., close relatives, employer, department, research unit, or close collaborator) have any financial or non-financial interests related to this guideline and which may be perceived as unduly influencing your judgment?

☐ No.

☐ Yes (please provide details including the nature of the interest, the parties involved, and its relevance to the guideline topic).

……
……

- Do you hold any institutional roles, affiliations, or responsibilities, in an organisation whose activities is related to the topic of this guideline (without receiving a personal financial benefit)?

☐ No.

☐ Yes (Please describe the role, the organisation, and its relevance to the guideline topic).

……
……

☐ **Declaration: I hereby declare on my honour that the disclosed information is true and complete to the best of my knowledge**

Should there be any change to the above information during the development of the guideline and before the publication of the final document, I will promptly notify the responsible ESHRE staff and complete a new declaration of interest form that describes the change.

Date Signature

C. CONFIDENTIALITY FORM

As a writer of an ESHRE guideline you have been or may be exposed to certain confidential and/or proprietary information, materials or data. It is important to the integrity of the writing process and final work that this information should be kept strictly confidential and not disclosed at any time under any circumstance.

Guideline/ GPR title	

Contact information of the guideline development group member

Name: ……

Institution, Address: ……

……

E-mail address: ……

Statement of confidentiality

I will not disclose any confidential and/or proprietary information, materials or data related to Guideline Development Group's work to any third party but keep this information strictly confidential.

I will keep any confidential and/or proprietary information, materials or data in my possession in a safe and secure place to protect against inadvertent disclosure.

I will not use any confidential information and/or proprietary information, materials or data for any purpose other than participating in an ESHRE guideline development procedure.

SIGNATURE (OR STATE YOUR NAME):

……

DATE:

……

D. SCOPING CHECKLIST

WHAT IS/ARE THE OVERALL PURPOSE(S) OF THE PROPOSED GUIDELINE?

Specify health intents (i.e., prevention, diagnosis, treatment, etc.) and expected benefits or outcomes. E.g. preventing thromboembolic complications of patients undergoing elective orthopaedic surgery.

……

WHAT IS THE PROPOSED TARGET PATIENT POPULATION?

Specify subjects to whom those recommendations apply (i.e. patients, society, etc.). E.g. adults undergoing elective orthopaedic surgery, all women 40 years of age or older, etc.

……

WHAT IS THE PROPOSED HEALTH CARE SETTING?

Specify the level of health care (i.e. primary, secondary, etc.) where these recommendations are supposed to be implemented.

……

WHICH INTERVENTIONS SHOULD BE INCLUDED IN THE GUIDELINE?

Specify which preventive, therapeutic and diagnostic interventions will be covered, and which will be not.

……

WHICH OUTCOME(S) SHOULD BE ADDRESSED BY THE PROPOSED GUIDELINE?

Specify which outcome(s) would be preferred, which are commonly reported, and which are preferred by patients.

……

WHO ARE THE TARGET USERS OF THE PROPOSED GUIDELINE, AND WHO ARE THE KEY STAKEHOLDERS?

Specify all relevant professional groups, institutions, patients, public, etc. who are target users or beneficiaries of these guidelines and/or whose views should be sought.

……

WHAT PREFERS THE PROPOSED PATIENT POPULATION?

Is this already included? Which methodology/methodologies will be used to include patients' preferences?

……

WHAT ARE KEY RESOURCES TO CONSIDER?

Specify resources needed for the implementation of guidelines (i.e. need for additional human resources, equipment, infrastructure, system changes, etc.) and potential barriers to implementation.

……

WHAT IS THE RELATION TO OTHER DOCUMENTS?

List all existing documents/guidelines on the same or similar topic that are likely to be currently used in practice (e.g. guidelines developed by other organisations).

E. PICO CHECKLIST

For each PICO question, provide the following information:

KEY QUESTION:	
PATIENTS/POPULATION:	
INTERVENTION:	
COMPARISON:	
OUTCOMES:	CRITICAL:
	IMPORTANT:
	OTHER:
Other databases to be searched for this question[6]	
Suggested key words/ search terms/synonyms:	

[6] in addition to Medline/PubMed and Cochrane Library

F. RELEVANCE AND QUALITY CHECK

REFERENCE	TYPE	RELEVANT FULL-TEXT?	EXPLANATION (EXCLUSION CRITERION)	QUALITY FULL-TEXT	EXPLANATION (EXCLUSION CRITERION)
XX	SR	◯ YES ◯ NO	_____	◯ HIGH ◯ MODERATE ◯ LOW	_____

If a paper is found to be not-relevant, or of low quality, the GDG member should list the exclusion criteria.

(Remarks can be added in case of relevant good quality papers as well).

Possible exclusion criteria (non-exhaustive list):

- ➔ Not relevant:
  - O Publication type: Case report - Expert opinion - Editorial
  - O Relevant patients are not included, or only as subgroup
  - O Relevant intervention is not included
  - O Relevant outcomes are not assessed or inappropriately assessed
  - O Study is included in meta-analysis / More recent meta-analysis available

- ➔ Low Quality
  - O Selection bias
  - O Performance bias
  - O Attrition bias
  - O Detection bias
  - O Other sources of bias:
    - — Study population: for instance too small
    - — Methodology
    - — Insufficient data

Instructions for the GDG members

1. **You will receive all full text papers and a relevance and quality check table with all references.** *For some full text papers we do not have access to the full text.*
2. **Check whether the paper is relevant. If not, list an explanation/exclusion criterion**
3. **If relevant, assess the quality of the paper (with the appropriate checklist for risk of bias assessment). List the most important criteria for your judgement on quality (especially in case of low quality)**
4. **Send in the completed relevance and quality check table.**

G. QUALITY ASSESSMENT CHECK

(AMSTAR - Assessing the Methodological Quality of Systematic Reviews)

1. Did the research questions and inclusion criteria for the review include the components of PICO?	○ Yes ○ No For Yes: ○ Population ○ Intervention ○ Comparator group ○ Outcome Optional (recommended) ○ Timeframe for follow up
2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	○ Yes ○ Partial Yes ○ No For Partial Yes: The authors state that they had a written protocol or guide that included ALL the following: ○ Review question(s) ○ A search strategy ○ Inclusion/exclusion criteria ○ A risk of bias assessment For Yes: As for partial yes, plus the protocol should be registered and should also have specified: ○ A meta-analysis/synthesis plan, if appropriate, and ○ A plan for investigating causes of heterogeneity
3. Did the review authors explain their selection of the study designs for inclusion in the review?	○ Yes ○ No For Yes, the review should satisfy ONE of the following: ○ Explanation for including only RCTs ○ OR explanation for including only NRSI ○ OR explanation for including both RCTs and NRSI
4. Did the review authors use a comprehensive literature search strategy?	○ Yes ○ Partial Yes

	○ No For Partial Yes (all the following): ○ Searched at least two databases (relevant to research question) ○ Provided key word and/or search strategy ○ Justified publication restrictions (e.g. language) For Yes, should also have (all the following): ○ Searched the reference lists/ bibliographies of included studies ○ Searched trail/study registries ○ Included/consulted content experts in the field ○ Where relevant, searched for grey literature ○ Conducted search within 24 months of completion of the review
5. Did the review authors perform study selection in duplicate?	○ Yes ○ No For Yes, either ONE of the following: ○ At least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include ○ OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer
6. Did the review authors perform data extraction in duplicate?	○ Yes ○ No For Yes, either ONE of the following: ○ At least two reviewers achieved consensus on which data to extract from included studies ○ OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer.
7. Did the review authors provide a list of excluded studies and justify the exclusions?	○ Yes ○ Partial Yes ○ No For Partial Yes: ○ Provided a list of all potentially relevant studies that were read in full-text form but excluded from the review For Yes, must also have:

	○ Justified the exclusion from the review of each potentially relevant study
8. Did the review authors describe the included studies in adequate detail?	○ Yes ○ Partial Yes ○ No For Partial Yes (ALL the following): ○ Described populations ○ Described interventions ○ Described comparators ○ Described outcomes ○ Described research designs For Yes, should also have ALL the following: ○ Described population in detail ○ Described intervention in detail (including doses where relevant) ○ Described comparator in detail (including doses where relevant) ○ Described study's setting ○ Timeframe for follow-up
9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?	○ Yes ○ Partial Yes ○ No ○ Includes only NRSI **RCTs** For Partial Yes, must have assessed RoB from ○ Unconcealed allocation ○ Lack of blinding of patients and assessors when assessing outcomes (unnecessary for objective outcomes such as all-cause mortality) For Yes, must also have assessed RoB from: ○ Allocation sequence that was not truly random, and ○ Selection of the reported result from among multiple measurements or analyses of a specified outcome ○ Yes ○ Partial Yes ○ No ○ Includes only RCTs
	NRSI For Partial Yes, must have assessed RoB: ○ From confounding, and ○ From selection bias

10. Did the review authors report on the sources of funding for the studies included in the review?	For Yes, must also have assessed RoB: ◯ Methods used to ascertain exposures and outcomes, and ◯ Selection of the reported result from among multiple measurements or analyses of a specified outcome ◯ Yes ◯ No For Yes: ◯ Must have reported on the sources of funding for individual studies included in the review. Note: reporting that the reviewers looked for this information, but it was not reported by study authors also qualifies.
11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?	◯ Yes ◯ No ◯ No meta-analysis conducted **RCTs** For Yes: ◯ The authors justified combining the data in a meta-analysis ◯ AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present. ◯ AND investigated the causes of any heterogeneity ◯ Yes ◯ No ◯ No meta-analysis conducted
12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	**NRSI** For Yes: ◯ The authors justified combining the data in a meta-analysis ◯ AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present ◯ AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available ◯ AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review ◯ Yes ◯ No ◯ No meta-analysis conducted

	For Yes: ◯ Included only low risk of bias RCTs ◯ OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect.
13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?	◯ Yes ◯ No For Yes: ◯ Included only low risk of bias RCTs ◯ OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results
14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review??	◯ Yes ◯ No For Yes: ◯ There was no significant heterogeneity in the results ◯ OR if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review
15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	◯ Yes ◯ No ◯ No meta-analysis conducted For Yes: ◯ Performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias
16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	◯ Yes ◯ No For Yes: ◯ The authors reported no competing interests OR ◯ The authors described their funding sources and how they managed potential conflicts of interest

Checklist: Randomized controlled trials

SELECTION BIAS (SYSTEMATIC DIFFERENCES BETWEEN THE COMPARISON GROUPS)	
An appropriate method of randomization was used to allocate participants to treatment groups (which would have balanced any confounding factors equally across groups).	○ Yes ○ No ○ Can't answer ○ Not applicable
There was adequate concealment of allocation (such that investigators/participants cannot influence enrolment or treatment allocation)[8]	○ Yes ○ No ○ Can't answer ○ Not applicable
The groups were comparable at baseline, including all major confounding and prognostic factors.	○ Yes ○ No ○ Can't answer ○ Not applicable
Based on your answers to the above, in your opinion was selection bias present? If so, consider the likely direction of its effect?	○ Yes ○ No ○ Can't answer ○ Not applicable

[8] Lack of allocation concealment: Those enrolling patients are aware of the group (or period in a crossover trial) to which the next enrolled patient will be allocated (a major problem in "pseudo" or "quasi" randomized trials with allocation by day of week, birth date, chart number, etc.).

Checklist: Randomized controlled trials – updated RoB2 tool Cochrane (20190822_RoB_2.0_template_parallel_trial.docx - Google Docs)

Responses underlined in green are potential markers for low risk of bias, and responses in red are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Domain 1: Risk of bias arising from the randomization process

Signalling questions	Comments	Response options
1.1 Was the allocation sequence random?		Y / PY / PN / N / NI
1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?		Y / PY / PN / N / NI
1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?		Y / PY / PN / N / NI
Risk-of-bias judgement		Low / High / Some concerns
Optional: What is the predicted direction of bias arising from the randomization process?		NA / Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Domain 2: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)

Signalling questions	Comments	Response options
2.1. Were participants aware of their assigned intervention during the trial?		Y / PY / PN / N / NI
2.2. Were carers and people delivering the interventions aware of participants assigned intervention during the trial?		Y / PY / PN / N / NI

2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the trial context?		NA / Y / PY / PN / N / NI
2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome?		NA / Y / PY / PN / N / NI
2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups?		NA / Y / PY / PN / N / NI
2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?		Y / PY / PN / N / NI
2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?		NA / Y / PY / PN / N / NI
Risk-of-bias judgement		Low / High / Some concerns
Optional: What is the predicted direction of bias arising from the randomization process?		NA / Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Domain 2: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)

Signalling questions	Comments	Response options
2.1. Were participants aware of their assigned intervention during the trial?		Y / PY / PN / N / NI
2.2. Were carers and people delivering the interventions aware of participants assigned intervention during the trial?		Y / PY / PN / N / NI
2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important non-protocol interventions balanced across intervention groups?		NA / Y / PY / PN / N / NI

2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome?		NA / Y / PY / PN / N / NI
2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?		NA / Y / PY / PN / N / NI
2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention?		NA / Y / PY / PN / N / NI
Risk-of-bias judgement		Low / High / Some concerns
Optional: What is the predicted direction of bias arising from the randomization process?		NA / Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Domain 3: Missing outcome data

Signalling questions	Comments	Response options
3.1 Were data for this outcome available for all, or nearly all, participants randomized?		Y / PY / PN / N / NI
3.2 If N/PN/NI to 3.1: Is there evidence that the result was not biased by missing outcome data?		NA / Y / PY / PN / N
3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?		NA / Y / PY / PN / N / NI
3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?		NA / Y / PY / PN / N / NI
Risk-of-bias judgement		Low / High / Some concerns

Optional: What is the predicted direction of bias arising from the randomization process?	NA / Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Domain 4: Risk of bias in measurement of the outcome

Signalling questions	Comments	Response options
4.1 Was the method of measuring the outcome inappropriate?		Y / PY / PN / N / NI
4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?		Y / PY / PN / N / NI
4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants?		NA / Y / PY / PN / N / NI
4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?		NA / Y / PY / PN / N / NI
4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?		NA / Y / PY / PN / N / NI
Risk-of-bias judgement		Low / High / Some concerns
Optional: What is the predicted direction of bias arising from the randomization process?		NA / Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Domain 5: Risk of bias in selection of the reported result

Signalling questions	Comments	Response options
5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized		Y / PY / PN / N / NI

before unblinded outcome data were available for analysis?		
Is the numerical result being assessed likely to have been selected, on the basis of the results, from...		
5.2. ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?		Y / PY / PN / N / NI
5.3 ... multiple eligible analyses of the data?		Y / PY / PN / N / NI
Risk-of-bias judgement		Low / High / Some concerns
Optional: What is the predicted direction of bias arising from the randomization process?		NA / Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Overall risk of bias

Signalling questions	Comments	Response options
Risk-of-bias judgement		Low / High / Some concerns
Optional: What is the predicted direction of bias arising from the randomization process?		NA / Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable

Checklist: Case control studies

SELECTION BIAS (SYSTEMATIC DIFFERENCES BETWEEN THE COMPARISON GROUPS)	
The cases and controls are taken from comparable populations	○ Yes ○ No ○ Can't answer ○ Not applicable
The same exclusion criteria are used for cases and controls	○ Yes ○ No ○ Can't answer ○ Not applicable
The participation rate was similar between cases and controls, and participants and non- participants are compared to establish their similarities and differences	○ Yes ○ No ○ Can't answer ○ Not applicable
Cases are clearly defined and differentiated from controls	○ Yes ○ No ○ Can't answer ○ Not applicable
PERFORMANCE BIAS (SYSTEMATIC DIFFERENCES BETWEEN GROUPS IN THE CARE PROVIDED, APART FROM THE INTERVENTION UNDER INVESTIGATION)	
Measures were taken to prevent knowledge of primary exposure from influencing case ascer- tainment	○ Yes ○ No ○ Can't answer ○ Not applicable
Exposure status is measured in a standard, valid and reliable way	○ Yes ○ No ○ Can't answer ○ Not applicable
The main potential confounders are identified and taken into account in design and analysis	○ Yes ○ No ○ Can't answer ○ Not applicable

Checklist: Case control studies

DETECTION BIAS (BIAS IN HOW OUTCOMES ARE ASCERTAINED, DIAGNOSED OR VERIFIED)	
Have confidence intervals been provided	◯ Yes ◯ No ◯ Can't answer ◯ Not applicable
OVERALL QUALITY OF CASE-CONTROL STUDIES	
Is the methodology used for the cohort study sufficiently robust to permit a valid conclusion?	◯ Yes ◯ No ◯ Can't answer ◯ Not applicable
What is your overall assessment of the methodological quality of this study	◯ HIGHT QUALITY ◯ MODERATE QUALITY ◯ LOW QUALITY ◯ VERY LOW QUALITY

Note: the outcome of the quality assessment should be documented in the relevance and quality check document.

Checklist: Studies of diagnostic accuracy

Checklist based on QUADAS tool (Whiting et al., 2004)

Was the spectrum of patients representative of the patients who will receive the test in practice?	○ Yes ○ No ○ Can't answer ○ Not applicable
Were selection criteria clearly described?	○ Yes ○ No ○ Can't answer ○ Not applicable
Is the reference standard likely to classify the target condition correctly?	○ Yes ○ No ○ Can't answer ○ Not applicable
Is the period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	○ Yes ○ No ○ Can't answer ○ Not applicable
Did the whole sample or a random selection of the sample receive verification using a reference standard of diagnosis?	○ Yes ○ No ○ Can't answer ○ Not applicable
Did patients receive the same reference standard regardless of the index test result?	○ Yes ○ No ○ Can't answer ○ Not applicable
Was the reference standard independent of the index test (i.e. the index test was not part of the reference standard)?	○ Yes ○ No ○ Can't answer ○ Not applicable

Checklist: Studies of diagnostic accuracy

Was the execution of the index test described in sufficient detail to permit replication of the test?	○ Yes ○ No ○ Can't answer ○ Not applicable
Was the execution of the reference standard described in sufficient detail to permit its replication?	○ Yes ○ No ○ Can't answer ○ Not applicable
Were the index test results interpreted without knowledge of the results of the reference standard?	○ Yes ○ No ○ Can't answer ○ Not applicable
Were the reference standard results interpreted without knowledge of the results of the index test?	○ Yes ○ No ○ Can't answer ○ Not applicable
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	○ Yes ○ No ○ Can't answer ○ Not applicable
Were uninterpretable/intermediate test results reported?	○ Yes ○ No ○ Can't answer ○ Not applicable
Were withdrawals from the study explained?	○ Yes ○ No ○ Can't answer ○ Not applicable

Checklist: Studies of diagnostic accuracy

OVERALL QUALITY OF THE DIAGNOSTIC STUDIES	
Is the methodology used for the diagnostic cohort study sufficiently robust to states the conclusion is reliable?	○ Yes ○ No ○ Can't answer ○ Not applicable
What is your overall assessment of the methodological quality of this study	○ HIGHT QUALITY ○ MODERATE QUALITY ○ LOW QUALITY ○ VERY LOW QUALITY

Note: the outcome of the quality assessment should be documented in the relevance and quality check document.

H. EVIDENCE TABLES

Intervention studies

Evidence table

REFERENCE	STUDY TYPE	PATIENTS NO. OF PATIENTS PATIENT CHAR-ACTERISTICS + GROUP COM-PARABILITY	INTERVENTIONS (+COMPARISON) INCLUDE: STUDY DURATION / FOLLOW-UP	OUTCOME MEASURES INCLUDE: HARMS / ADVERSE EVENTS	EFFECT SIZE	AUTHORS CONCLUSIONS	COMMENTS

Details

REFERENCE (+PMID)	Author, journal, year, volume and pages + PMID
STUDY TYPE	Meta-analysis/ systematic review (SR), randomized controlled trial (RCT), non-randomized cohort study (CS), case report (CASE), opinion documents (OPINION), other (specify)...
PATIENTS NO. PATIENTS CHARACTERISTICS + GROUP COMPARABILITY	Total number of patients, and the number of patients involved in each group Describe relevant baseline characteristics (age, disease status, inclusion / exclusion criteria,..), if stated per group and add comment for comparability of groups.
INTERVENTIONS (+COMPARISON) INCLUDE STUDY DURATION / FOLLOW-UP	Specify the interventions per group (treatment/procedure (dose, regimen, length,..) / placebo/ alternative treatment / expectant management) duration patients participate the study from inclusion to a specified endpoint (e.g. implantation) or the end of data collection, mention per outcome if relevant
OUTCOME MEASURES (PRIMARY/SECONDARY) INCLUDE HARMS / ADVERSE EVENTS	All outcome measures (positive and negative): e.g. OHSS occurrence rate, implantation rate, pregnancy, quality of life, satisfaction,.. Divide between primary (1) and secondary (2) outcomes when this is specified by the author
EFFECT SIZE (INCLUDE HARMS)	Absolute risk reduction, relative risks, numbers needed to treat/harm, or odds ratios with confidence intervals; State clearly which outcome measure is used. Add p-values, if available. Define and describe observed harms per group as reported in the paper. Precise mean, percentages and p-values, if available
AUTHORS CONCLUSION	Report the Authors conclusion
COMMENTS	Additional characteristics/interpretations or flaws of the study, additional calculations made by the reviewer (NNT, RR, OR, CI,..) If relevant, mention the source of funding, or any competing interests. If the paper is to be excluded, mention the exclusion criterion here.

When no element can be added, include:

- ➔ *"Not applicable (NA)" when an item is not to be informed, or*
- ➔ *"Not described (ND)" when an item must be informed but no information is given in the publication)*

Diagnostic studies

Evidence table

REFERENCE	STUDY TYPE	PATIENTS NO. OF PATIENTS PATIENT CHARACTERIS-TICS + GROUP COMPARABILITY	DIAGNOSTIC TEST EVALUATED REF-ERENCE STANDARD TEST INCLUDE: TIME INTERVAL AND TREATMENT	OUTCOME MEAS-URES: PREVA-LENCE ACCURACY REPRODUCIBILITY	AUTHORS CONCLU-SIONS	COMMENTS

Details

REFERENCE (+PMID)	Author, journal, year, volume and pages + PMID
STUDY TYPE	Meta-analysis/ systematic review (SR), randomized controlled trial (RCT), non-randomized cohort study (CS), case report (CASE), opinion documents (OPINION), other (specify)...
PATIENTS NO. PATIENTS CHARACTERISTICS + GROUP COMPARABILITY	Total number of patients, and the number of patients involved in each group Describe relevant baseline characteristics (age, disease status, inclusion / exclusion criteria,..), if stated per group and add comment for comparability of groups. Setting: Multicentre, countries, healthcare setting,.
DIAGNOSTIC TEST EVALUATED REFERENCE STANDARD TEST INCLUDE: TIME INTERVAL AND TREATMENT	Describe the evaluated test (what, by whom and how, when,..), cut offs, blinding to clinical information and/or index test results Describe the evaluated test (what, by whom and how, when,..), cut offs, blinding to clinical information and/or index test results Specify the time interval and treatments administered between the tests (if any)
OUTCOME MEASURES : ➔ **PREVALENCE** ➔ **ACCURACY** ➔ **REPRODUCIBILITY**	Disease prevalence Accuracy: Give all available figures with confidence intervals (if available): Sensitivity (Se), Specificity (Sp), positive predictive value (PPV), negative predictive value (NPV), likelihood ratios (LR, LR-), area under the ROC curve Reproducibility: Give all available figures with confidence intervals (if available): ➔ Quantitative test (number of repetitions, extent of values, agreement, correlation) ➔ Qualitative test (reliability, correlation coefficient,..)
AUTHORS CONCLUSION	Report the Authors conclusion
COMMENTS	Additional characteristics/interpretations or flaws of the study, additional calculations made by the reviewer If relevant, mention the source of funding, or any competing interests. If the paper is to be excluded, mention the exclusion criterion here.

When no element can be added, include:

- ➔ *"Not applicable (NA)" when an item is not to be informed, or*
- ➔ *"Not described (ND)" when an item must be informed but no information is given in the publication)*

Instructions for the GDG members

1. **You will receive all full text papers and a relevance and quality check table with all references.**
   - ➔ *For same full text papers, we do not have access to the full text.*

2. **For some papers, you do not have to complete the evidence table. In the last column remarks have been added.**
   - ➔ "GRADE evidence profile": this meta-analysis or RCT will be added to a GRADE profile. You can add your comments to the meta-analysis or RCT, but you do not need to complete all details.
   - ➔ "EXCLUDE": Usually studies that are relevant but have been included in a meta-analysis. You can leave this blank. If at some point the meta-analysis is found to be of low quality, we may go back and complete the information in the evidence table.
   - ➔ Any other comment on the paper, you can decide whether to complete the evidence table.

3. **Read the paper thoroughly**
   - ➔ Complete the evidence table: fill in as much information as possible that is relevant for answering the PICO question. (see detailed instructions below for intervention and diagnostic studies).
   - ➔ Even though you have already judged the paper as "good quality" and "relevant", you can exclude papers at this stage. If you would like to exclude a paper, you do not have to complete the evidence table, just add a remark in the last column: "excluded due to…".
   - ➔ In the last column of the evidence table, formulate any concerns, comments, or questions you have regarding the content of the paper.

4. **Send in the completed evidence table**
   - ➔ The evidence table and GRADE profiles are the basis to write a summary of evidence and recommendations.
   - ➔ You can continue writing the summary of evidence and draft recommendations.
   - ➔ We will check and if needed complete the evidence table. Please remember that each paper in the evidence table should be mentioned in the summary of evidence or excluded with an appropriate exclusion criterion.

I. FRAMEWORK FOR RECOMMENDATIONS

The justification should comprise the following considerations:

	CRITERIA	JUDGEMENTS	RESEARCH EVIDENCE	ADDITIONAL INFORMATION
BENEFITS/HARMS	Are the desirable effects large relative to the undesirable effects?	○ Benefits outweigh harms/burden ○ Benefits slightly outweigh harms/burden ○ Benefits and harms/burden are balanced ○ Harms/burden slightly outweigh benefits ○ Harms/burden outweigh benefits	*Per outcome*	
EVIDENCE	What is the overall quality of evidence?	○ High ○ Moderate ○ Low ○ Very low ○ No included studies	*Per outcome*	
VALUES	What are the patient values and what certainty do we have about them?	○ Little uncertainty and similar values ○ Some uncertainty or some variation ○ Significant uncertainty or large variation		
RESOURCES	Is the incremental cost (or resource use) small relative to the benefits?	○ Cost is very small relative to the benefits ○ Cost is small relative to the benefits ○ Cost is borderline relative to the benefits ○ Cost is high relative to the benefits ○ Cost is very high relative to the benefits		

HEALTH SYSTEM PERSPECTIVE	What would the impact be on health equity?	◯ Reduced ◯ Probably reduced ◯ Probably no impact ◯ Probably increased ◯ Increased ◯ Varies ◯ Don't know		
	Is the intervention acceptable to key stakeholders?	◯ No ◯ Probably No ◯ Probably yes ◯ Yes ◯ Varies ◯ Don't know		
	Is the intervention feasible to implement?	◯ No ◯ Probably No ◯ Probably yes ◯ Yes ◯ Varies ◯ Don't know		
OTHER	*Subgroup considerations*[15]			
	Implementation considerations[16]			
	Research priorities[17]			

[15] Are the recommendations applicable to subgroups, and is this mentioned in the recommendation?

[16] Are there any barriers that could restrict the implementation of this recommendation? If so, how can we handle this?

[17] Is there a need for further research on the topic, and is it a priority? If yes, please provide details on how new studies should ideally be performed (RCT, large multicentre) and what their focus should be (subgroup of patients, specific intervention). All recommendations for research will be added as an annex to the guideline.

J. REVIEWER DISCLOSURE FORM

All reviewers are expected to provide completed and signed disclosure statements about all financial, personal, or professional relationships with industry, individuals, or organisations from the last three years to avoid perception of COI.

Guideline title	

Contact information of the reviewer

Name: ……

E-mail address: ……

- ○ I HAVE NO POTENTIAL COI FROM THE LAST THREE YEARS TO REPORT.
- ○ I HAVE THE FOLLOWING POTENTIAL COI FROM THE LAST THREE YEARS TO REPORT:
  - ○ RESEARCH GRANT(S) FROM ONE OR MORE COMPANIES, FROM
    ……
  - ○ CONSULTING FEE(S) FOR E.G. SERVICES ON AN ADVISORY BOARD OR LEGAL TESTIMONY, FROM
    ……
  - ○ SPEAKER FEE(S) FOR INSTANCE AS COMPENSATION FOR LECTURING AND TRAVEL, FROM
    ……
  - ○ SALARY OR POSITION FUNDING, FROM
    ……
  - ○ OWNERSHIP INTEREST BY STOCK (OPTIONS) OR PARTNERSHIP OF A HEALTHCARE COMPANY, FROM
    ……
  - ○ OTHER (FINANCIAL) BENEFIT E.G. BY INSITUTIONAL COI IN THE TOPICS OR ISSUES ADDRESSED IN THE DOCUMENT
    ……

SIGNATURE (OR STATE YOUR NAME):

……

DATE:

……

K. REVIEWER COMMENTS FORM

Guideline title	
Review period	

Contact information of the reviewer

Name: ……

E-mail address: ……

I AM PARTICIPATING

- ○ AS AN INDIVIDUAL
- ○ ON BEHALF OF A (INTER)NATIONAL ORGANISATION, NAMELY

……

- ○ ON BEHALF OF A COMPANY, NAMELY

……

Statement of confidentiality

As a reviewer of this ESHRE document you have been or may be exposed to certain confidential and/or proprietary
information, materials or data. It is important to the integrity of the writing process and final work that this information should be kept strictly confidential and not disclosed at any time under any circumstance.

- ➔ I will not disclose any confidential and/or proprietary information, materials or data related to Working Group's work to any third party but keep this information strictly confidential.
- ➔ I will keep any confidential and/or proprietary information, materials or data in my possession in a safe and secure place to protect against inadvertent disclosure.
- ➔ I will not use any confidential information and/or proprietary information, materials or data for any purpose other than participating in the review procedure.

SIGNATURE (OR STATE YOUR NAME):

……

DATE:

……

Comments on the document

PAGE	LINE	COMMENT

(Add more lines if you need to)

Please send completed forms (as word-document or pdf) to guidelines@eshre.eu before XX.

All comments will be revised by the guideline development group and assessed. If the comment is accepted by the guideline development group, it will result in a modification of the document. If not, the guideline group will formulate a reply to the reviewer. The details of the review procedure, the comments, modifications and replies will be summarized in a review report which will be available on the ESHRE website.

By submitting this form, you will be listed as an expert reviewer of the guideline. The list of reviewers will be published in the review report and in an annex of the main guideline document.

For more information on the review, you can contact guidelines@eshre.eu.

L. DOCUMENT ASSESSMENT TOOL FOR UPDATING GUIDELINES

Adapted from(Agbassi et al., 2014)

NUMBER AND TITLE OF THE GUIDELINE	
CURRENT REPORT DATE	
LAST LITERATURE SEARCH DATE	
DATE ASSESSED	
RESEARCH SPECIALIST	
OUTCOME (FOR INTERNAL USE)	

ASSESSMENT:
For each document, please respond YES or NO to all the questions listed below. Provide an explanation of each answer where necessary.

1. IS THE DOCUMENT STILL RELEVANT (CLINICALLY OR TO THE CARE SYSTEM AS A WHOLE IN SOME WAY)?

2. SHOULD FULL ASSESSMENT AND REVIEW OF THIS DOCUMENT BE DEFERRED UNTIL NEXT YEAR?

Consider YES if:

- ➔ The document is less than three years old, and there is no reason to doubt the recommendations
- ➔ The document is between three and five years old, and a justification can be provided as to why the recommendations can be considered trustworthy for another year

3. DO THE QUESTIONS AND SEARCH CRITERIA AS THEY ARE IN THE DOCUMENT ADDRESS CURRENT NEEDS, SUCH THAT AN UPDATED LITERATURE SEARCH WOULD BE USEFUL AND IDENTIFY RELEVANT EVIDENCE?

Consider NO if:

- ➔ The standard of care has shifted significantly since the last version of the document, such that the questions only address the topic in part
- ➔ There are new, significant options (for treatment, diagnosis, etc.) available that are not covered by the current questions, such that new questions would need to be added to the document
- ➔ In general, if you believe that for the document to still be useful it will have to substantially be rewritten
- ➔ The document has been repeatedly deferred, and is now older than five years

4. DOES THE DOCUMENT HAVE AN IMPACT ON ACCESS TO CARE (THAT IS, ARE DECISIONS ABOUT ACCESS OR PAYMENT FOR CARE MADE BY THE MINISTRY, CCO, OR OTHER ORGANIZATIONS BASED ON THE RECOMMENDATIONS IN THIS DOCUMENT)? *Consider YES if:* ➔ Ministry funding decisions have been, are, or will be made based on this document ➔ An indication for a chemotherapy regimen was funded, or rejected, based on the document ➔ Case by case review or out of country requests are known to be decided based on the document ➔ Funding for some screening, diagnostic, staging or treatment procedure was or is determined
5. IS THERE KNOWN EVIDENCE THAT HAS BEEN PUBLISHED SINCE THIS DOCUMENT'S LAST LITERATURE SEARCH (SEE ABOVE) THAT WOULD RESULT IN SIGNIFICANT CHANGES TO THE RECOMMENDATIONS?
6. SHOULD THIS DOCUMENT BE TAKEN OFF THE WEBSITE WHILE IT AWAITS FULL REVIEW, OR CAN IT BE LEFT THERE WITH AN "IN REVIEW" WATERMARK? *Consider YES if:* ➔ If followed, even in error, the recommendations have the potential to cause harm to patients.

PLEASE LIST ANY ADDITIONAL FACTORS THAT SHOULD BE CONSIDERED IN PRIORITIZING THIS DOCUMENT FOR REVIEW:

.....

OUTCOME:	○ ENDORSE	The recommendations are still current and relevant for decision making. This can happen when there is a very strong justification to conclude that without a search for new evidence, the recommendations are still valid[18].
	○ DEFER	The document remains current and credible enough to wait until the next assessment.
	○ REVIEW	The document will undergo a review for currency and relevance.
	○ ARCHIVE	The document cannot be endorsed or deferred, and the recommendations will no longer be maintained.[19]

[18] For example, in cases in which added evidence will not change the recommendations because the existing evidence is so definitive, of high quality, and adequate quantity or in cases in which no additional evidence in that topic will be forthcoming because it is no longer an area of inquiry.

[19] This may happen because the recommendations are no longer clinically relevant and applicable to current practice. Or it may be because the developing group has little or no interest in maintaining them; for example, the topic areas may have changed so much that developing a new document is a more practical option than updating the existing one. Archived documents may still be useful for academic or other information purposes.