Journal of Clinical Epidemiology 150 (2022) 51−62

Journal of Clinical Epidemiology

REVIEW ARTICLE

Evidence to decision frameworks enabled structured and explicit development of healthcare recommendations

Jose F. Meneses-Echavez[a,b,*], Julia Bidonde[a,c], Juan Jose Yepes-Nuñez[d,e], Tina Poklepović Peričić[f], Livia Puljak[g], Malgorzata M. Bala[h], Dawid Storman[i], Mateusz J. Swierz[h], Joanna Zając[h], Camila Montesinos-Guevara[j], Yuan Zhang[k], Nathaly Chavez Guapo[l], Holger Schünemann[k], Signe Flottorp[m], Pablo Alonso-Coello[n]

[a]*Norwegian Institute of Public Health, Oslo, Norway*
[b]*Facultad de Cultura Física, Deporte y Recreación. Universidad Santo Tomás, Bogotá, Colombia*
[c]*School of Rehabilitation Science, College of Medicine, University of Saskatchewan, Canada*
[d]*Universidad de los Andes, School of Medicine, Bogotá, Colombia*
[e]*Pulmonology Service, Internal Medicine Section, Fundación Santa Fe de Bogotá University Hospital, Bogotá, Colombia*
[f]*Department of Research in Biomedicine and Health, University of Split School of Medicine, Split, Croatia*
[g]*Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Zagreb, Croatia*
[h]*Chair of Epidemiology and Preventive Medicine, Department of Hygiene and Dietetics, Systematic Reviews Unit, Jagiellonian University Medical College, Krakow, Poland*
[i]*Chair of Epidemiology and Preventive Medicine, Department of Hygiene and Dietetics, Systematic Reviews Unit, Department of Adult Psychiatry, University Hospital Jagiellonian University Medical College, Krakow, Poland*
[j]*Centro de Investigación en Salud Pública y Epidemiología Clínica (CISPEC), Facultad de Ciencias de la Salud Eugenio Espejo, Universidad UTE, Quito, Ecuador*
[k]*Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada*
[l]*Facultad de Enfermería y Rehabilitación, Programa de Fisioterapia, Universidad de la Sabana, Chía, Colombia*
[m]*Institute of Health and Society, University of Oslo, Oslo, Norway*
[n]*Iberoamerican Cochrane Center, Biomedical Research Institute Sant Pau (IIB Sant Pau-CIBERESP), Barcelona, Spain*

Accepted 7 June 2022; Published online 13 June 2022

Abstract

Objective: The aim of this study is to identify and describe the processes suggested for the formulation of healthcare recommendations in healthcare guidelines available in guidance documents.

Methods: We searched international databases in May 2020 to retrieve guidance documents published by organizations dedicated to guideline development. Pairs of researchers independently selected and extracted data about the characteristics of the guidance document, including explicit or implicit recommendation-related criteria and processes considered, as well as the use of evidence to decision (EtD) frameworks.

Results: We included 68 guidance documents. Most organizations reported a system for grading the strength of recommendations (88%), half of them being the GRADE (Grading of Recommendations Assessment, Development, and Evaluation) approach. Two out of three guidance documents (66%) proposed the use of a framework to guide the EtD process. The GRADE-EtD framework was the most often reported framework (19 organizations, 42%), whereas 20 organizations (44%) proposed their own multicriteria frameworks. Using any EtD framework was related with a more comprehensive set of recommendation-related criteria compared to no framework, especially for criteria like values, equity, and acceptability.

Conclusion: Although limited, the use of EtD frameworks was associated with the inclusion of relevant recommendation criteria. Among the EtD structured frameworks, the GRADE-EtD framework offers the most comprehensive perspective for evidence-informed

Conflict of interest statement: The authors declare that they have no competing interests. Jose Meneses-Echavez is a doctoral candidate in Public Health and Methodology of Biomedical Research, at the Department of Pediatrics, Obstetrics, Gynecology and Preventive Medicine at Universitat Autonoma de Barcelona, Spain. Pablo Alonso-Coello, Signe Flottorp, and Holger Schünemann were involved in the development of the GRADE-EtD framework.

Funding Source: Jose F. Meneses-Echavez received funding from Universidad Santo Tomás, Bogotá, Colombia (FODEIN project code: 2115005).

* Correspondence author: Researcher, Division for Health Services. Norwegian Institute of Public Health, Sandakerveien 24C, Building D11, 4th floor office 434, Oslo, Norway 0477. Tel.: +47 486 89 405.

E-mail address: jose.meneses@fhi.no (J.F. Meneses-Echavez).

https://doi.org/10.1016/j.jclinepi.2022.06.004

Keywords: Practice guidelines; Decision making; Evidence-based practice; Methods; Evidence-to-decision framework; Clinical practice guidelines; GRADE; Health recommendations; Certainty of evidence; Systematic reviews; Evidence synthesis

1. Introduction

Guidelines for clinical, public health, and health policy, from here on guidelines, enhance decision making by translating complex scientific research findings into recommendations for practice [1,2]. Guideline developers share the same aim, establishing standards of care backed by strong scientific evidence; however, they do not share the same methodological expertise and resources for guideline development [1,2]. According to the National Academy of Medicine, trustworthy guidelines should ''be developed by a knowledgeable, multidisciplinary panel of experts and representatives from key affected groups; be based on a systematic review of the existing evidence; consider important patient subgroups and patient preferences as appropriate; be based on an explicit and transparent process that minimizes distortions, biases, and conflicts of interest; provide a clear explanation of the logical relationships between alternative care options and health outcomes and provide ratings of both the quality of evidence and the strength of the recommendations; and be reconsidered and revised as appropriate when important new evidence warrants modifications'' [2].

The process of moving from evidence to recommendations represents a cornerstone of guideline development [2,3]. This process implies a systematic and transparent integration of the evidence supporting the criteria influencing the recommendation, as expressed by item 13 in the Guidelines 2.0 checklist [3]. Due to the broad variety of recommendation-related criteria suggested by international organizations [2−7], the McMaster checklist, a tool for guideline development, points out that a framework outlining the criteria to be considered to arrive at a decision (e.g., the magnitude of the difference between benefits and harms, and resource use) should guide the recommendation formulation process [3].

A number of frameworks have been proposed for addressing this process, such as the GRADE-EtD (Grading of Recommendations Assessment, Development, and Evaluation evidence to decision) framework [4], the ''decision-making triangle'' [8], the Evidence and Value: Impact on DEcision Making framework [6], or the Guidance for Priority Setting in Health Care (GPS-Health) [9]. Some research groups and organizations have launched other EtD frameworks and processes, mostly based on the criteria and subcriteria contained in the GRADE-EtD framework [10,11]. Some of EtD frameworks emphasize specific criteria, such as equity in the GPS-Health [9], or ethics as the ''decision-making triangle'' [8]. Despite the variation in terms of the criteria being proposed, all the EtD frameworks aim to offer a comprehensive list of criteria needed to be considered by both decision makers and guideline developers.

One of the most popular frameworks is the GRADE-EtD. There exists a growing body of research about the role of the GRADE-EtD framework in guideline development [12]. The GRADE Working Group developed the EtD framework for different types of decisions, including recommendations [4]. Details about the development process of the GRADE-EtD framework are available elsewhere [13]. The GRADE-EtD framework aims to help panel members use evidence in a structured and transparent way to inform healthcare decisions. Besides, the framework helps guideline development groups (GDG) to considering the most relevant criteria that influence a decision by structuring discussions, identifying reasons for disagreements, and building the basis for a transparent decision [4,13] (Fig. 1). The GRADE-EtD framework has been used by international organizations including the World Health Organization (WHO), the European Commission [7], and major scientific societies [14−16]. Some panelists have acknowledged the relevance of the framework in facilitating both structured panel meetings and the consideration of all relevant criteria pertaining to decision making [17].

The controversy around the best fit for purpose framework remains in the literature as new proposals continue to emerge even in recent years [10]. In fact, despite the progress achieved in studying the importance of various recommendation-related criteria, to our knowledge, little is known on how GDGs work, the EtD processes they follow, and the criteria that organizations consider when formulating recommendations. This gap constrains the transparency of reporting of the methods through which healthcare recommendations were developed and affects the extent to which end users can understand the evidence grounding the recommendations and whether these might be implemented in their own setting.

Therefore, our study addresses the following question: What are the processes and frameworks suggested for the formulation of recommendations in healthcare guidelines available in guidance documents?

2. Methods

2.1. Design

Methodological study. We followed the methods for the conduct of a systematic review (e.g., double data

What is new?

Key finding

- One in three organizations do not report a structured Evidence to Decision (EtD) process in their guidance documents.
- Using an EtD framework is associated with a more comprehensive process, including additional criteria, like values, equity, and acceptability.

What this adds to what was known?

- The EtD process in guidance documents from guideline development organizations is often unstructured.
- Most organizations consider desirable effects, undesirable effects, and certainty of the evidence of effects. Other criteria, such as patients' values and preferences, cost-effectiveness, equity, acceptability, or feasibility are less frequently considered.

What is the implication and what should change now?

- Guidance documents for guideline development require improvements in the reporting of the EtD process.
- Guideline development organizations should follow a more systematic and explicit methodology for formulating recommendations.
- A more complete and detailed reporting of the EtD process should be included in guideline development guidance documents.
- Organizations may use our findings to either prepare their own guidance documents or to reassess their existing ones by identifying gaps in EtD criteria or to follow suggestions for improvement.

extraction) [18]. We published the protocol in Open Science Framework [19].

2.2. Inclusion criteria

We included the latest version of methodological guidance documents for guideline development, published from 2003 as organization reports, journal articles, or briefs, developed by institutions responsible for guideline development in any field of health care. We discarded guidance documents dealing with processes of updating or adapting guidelines as well as those endorsing included documents.

2.3. Literature searches

We systematically searched the Guidelines International Network library, MEDLINE, Google Scholar, and The Cochrane Methodology Register using key terms such as "handbook*," "guideline methodology," "clinical practice guideline*," and "manual." Then, we inspected the references listed in the included documents and websites of relevant organizations identified in previous research projects [1,20–22]. These procedures (search and inspection of references) were first conducted during October 2018, and further updated in May 2020. The searches had no language restrictions except for Google Scholar, which was restricted to English. We contacted experts in the field. Additional file 1 describes the search strategies.

2.4. Study selection and data extraction

We designed and piloted an ad hoc data abstraction form based on the WHO handbook for guideline development [7] because of its comprehensiveness, relevance, and formative influence on other guideline documents. The form covered information on the main characteristics of the guidance documents, as well as characteristics of the methodology [e.g., panel composition and management of conflicts of interest (COI)], and information on the EtD process, including the use of frameworks [4]. We defined a structured process as any systematic series of steps taken to fulfill a goal, and a framework as any structure of concepts underlying a structured process, in this case the process of formulating recommendations (EtD process). We explored the full-text of each guidance document for information about the suggested EtD process, including the use of frameworks, and listed all recommendation-related criteria considered. Both selection and data extraction processes were undertaken independently by pairs of reviewers, with discrepancies solved by consensus or by involving a third reviewer, if needed.

2.5. Data management and analysis

Pairs of reviewers read the guidance documents and extracted the information pertaining to the recommendation formulation process (e.g., direct quotes). We held rounds of virtual meetings to ensure accuracy and completeness of the data. All data extraction forms were then compiled by one reviewer (J.F.M.E.), who run descriptive analysis, with frequencies and percentages (e.g., characteristics of the organizations, panel composition, and COIs management). Based on the EtD process and the recommendation-related criteria identified in the guidance documents, we created the following categories:

1. Use of any framework for the EtD process
    - 1.1. Use of GRADE-EtD framework
    - 1.2. Use of other EtD framework
2. No framework

The last step of the analysis involved the study of the probability of addressing each criterion across the different categories of frameworks presented above (i.e., any framework vs. no framework; GRADE-EtD vs. other EtD framework; GRADE-EtD vs. no framework; other EtD framework vs. no framework), for which we estimated odds ratios (OR) with their corresponding 95% confidence intervals (CI). We used R-project software for the analysis [23]. J.F.M.E. carried out the statistical analysis, and a second reviewer (P.A.C.) audited it for accuracy.

3. Results

The search resulted in a total of 8,838 records. After we removed 268 duplicates, 8,570 records remained. We excluded 8,478 at full-text and assessed the 92 full-text documents for eligibility (Additional file 2). Of these, we excluded 24 documents, either because they were not a guidance document, were a previous version of an included document, were impossible to retrieve, or were publications related to documents already included. We included a total of 68 guidance documents from 14 countries (Fig. 2). Additional file 3 presents the names of the organizations and their corresponding guidance documents included in this study. The median publication year of the documents was 2015 (range 2003–2020). Scientific societies published most of the documents (58%), followed by governmental (20%) and supranational organizations (13%). Nearly half of the documents were from North America (45%), followed by Europe (37%), Asia and Oceania (6%), and South America (3%). See Figure 3 and Additional file 4 for further details on the characteristics of the included guidance documents. Figure 2 does not show six guidance documents published by European organizations, WHO, and World Confederation for Physical Therapy.

Most of the guidance documents provided some degree of information about the composition of guideline panels. A multidisciplinary panel, including a steering group, methodologists with expertise in evidence synthesis, and health professionals specialized in the topic of interest was the most common composition. More than half of the documents (64%) suggested the involvement of patients or consumer representatives (Additional file 4). Half of the guidance documents presented a definition of COIs and presented a process to report the COIs. Although 87% of the guidance documents specified who should report COIs, only half of them stated that any declaration of COIs should be reviewed before making appointments to the GDG and specified the types of COIs to be declared (i.e., financial and nonfinancial). Half of the documents (59%) described a process for the management of COIs (see Additional file 4).

3.1. Panel composition and conflicts of interest management

3.1.1. Rating the quality of evidence and grading the strength of recommendations.

Almost all guidance documents (93%) suggested a structured approach or system for rating the quality of the evidence; most of them (85%) contained a specific section or chapter. GRADE was the most common approach suggested (53%), followed by the approaches proposed by the Australian National Health and Medical Research Council and the Scottish Intercollegiate Guidelines Network (SIGN), reported in three documents each (4%). Of note, seven documents (10%) suggested rating systems that were based on GRADE or other approaches. Finally, nine documents (13%) proposed their own systems, mainly based on the previous approaches. See Additional file 4 for further details.

Most of the guidance documents (88%) proposed a system for grading the strength of recommendations; half of them

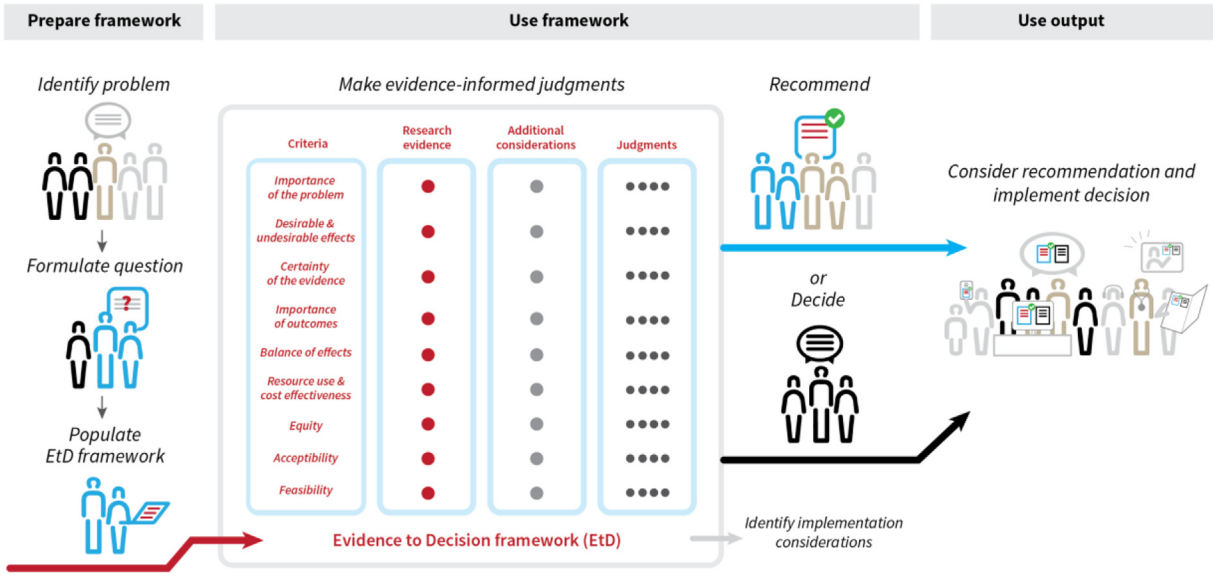

Fig. 1. GRADE-EtD frameworks workflow. Illustration: Sarah Rosenbaum, Norwegian Institute of Public Health. GRADE-EtD, Grading of Recommendations Assessment, Development, and Evaluation evidence to decision

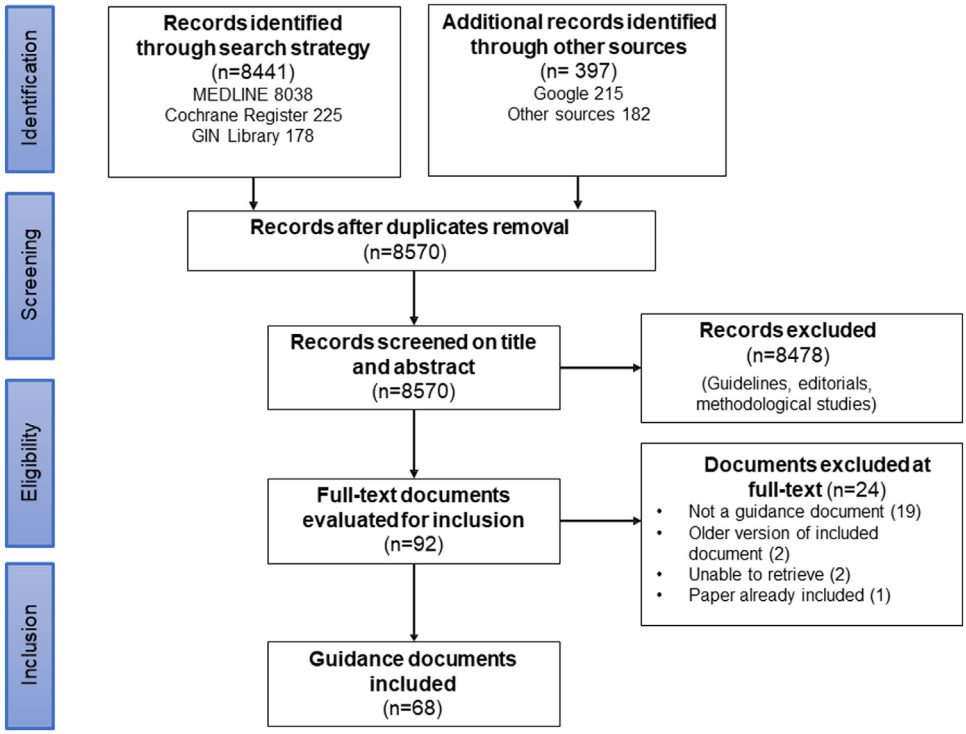

Fig. 2. Selection process.

suggested the GRADE approach. Around a fifth of the documents (22%) suggested their own approach, whereas 14% of the documents did not report any. Other approaches were SIGN (3%), Oxford Center for Evidence-Based Medicine (1 document), NICE (National Institute for Health Care and Excellence; 1 document), and US Preventive Services Task Force (USPSTF; 1 document). Five documents (8%) proposed other approaches, mostly adapted from GRADE (Table 1).

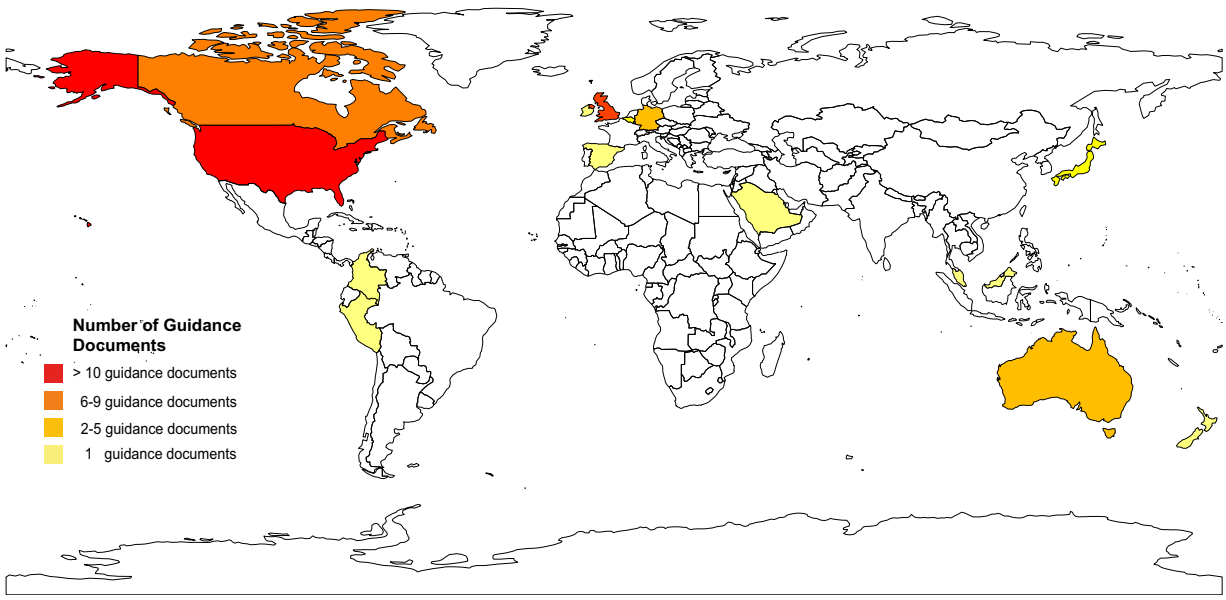

Fig. 3. Geographical distribution of the included guidance documents.

3.2. Recommendation formulation

Two out of three guidance documents (66%) suggested a structured process for formulating recommendations, with the entire panel being involved in more than half (56%) of the documents. In contrast, 28% of the guidance documents failed to report who was involved in the process of formulating recommendations. The technical team needed to share preliminary material (e.g., evidence summaries and evidence-to-recommendation tables) with the guideline panel ahead of GDG meetings in 40% of the guidance documents, and 48% of the documents reported that the technical team should make preliminary judgments about the quality of the evidence. A lower number of documents (37%) declared that preliminary judgments on the certainty of the evidence or any other factor (e.g., values and preferences, equity, resources required) should be made before the panel meetings. A smaller number of documents (22%) suggested the same for deciding about the direction and strength of recommendations (10%) (Table 1).

3.3. Use of evidence to decision frameworks

The GRADE-EtD framework was the most often reported (42%), followed by NICE's framework (8%), SIGN (4%), and USPSTF (4%). Twenty guidance documents (76%) reported their own multicriteria framework. Nearly half of the documents (57%) provided guidance for formulating recommendations when there is insufficient evidence or no evidence available, or low/very low-quality evidence. Most of the documents (82%) suggested a method to reach agreement among guideline panel members [e.g., consensus, voting (majority rule), or nominal group techniques] (Table 1).

3.4. Setting, perspective, and subgroup considerations

We found that 42% and 24% of the guidance documents that included an EtD framework specified the setting and perspective, respectively. These rates were considerably lower among documents that did not suggest any framework (17% setting and 9% perspective). Guidance documents suggesting the GRADE-EtD framework were the most common ones in reporting the inclusion of both setting and perspective (42%), followed by those that suggested another EtD framework (42% and 11%, respectively). The inclusion of subgroups considerations showed a similar pattern among the guidance documents (i.e., 31% in the use of any EtD framework and 4% in the use of no framework).

3.5. Recommendation-related criteria

A total of 14 recommendation-related criteria were identified across the guidance documents. Overall, guidance documents that suggested an EtD framework considered a more comprehensive set of criteria than those who did not suggest any framework. All guidance documents had a major focus in their EtD frameworks on desirable effects, undesirable effects, and certainty of the evidence of effects; guidance documents that suggested any framework for guiding the EtD process had higher reporting rates (>80%) than those who suggested no framework (52%). Other criteria such as equity, acceptability, and feasibility were reported in less than half of the documents that guided the EtD process through a systematic framework (any EtD framework), whereas considerably lower rates of use were observed among the documents that did not suggest a framework to guide the EtD process (e.g., from 0% to 22%). Of note, one document reported on the consideration of legal consequences (Domus Medica) and on bioethical considerations [Italian National Center for Clinical Excellence, Quality, and Security (CNEC)], respectively.

Among the documents that suggested a framework for guiding the EtD process, those that suggest the GRADE-EtD framework addressed a larger number of criteria compared to those that suggested another framework for the EtD process (Table 3). Rates of use differed between the two categories. For instance, 95% of the guidance documents that suggested the GRADE-EtD framework reported the consideration of patients' values relative to 46% of those that used another EtD framework. A larger difference was observed for equity considerations, 42% in the GRADE-EtD framework and 11% in another EtD framework, respectively. Table 2 presents all the recommendation-related criteria reported in the guidance documents. Results from the bivariate analysis are presented in the section below.

The documents failed to provide specific definitions of the criteria. Although the wording of some recommendation-related criteria may vary across the documents, which might be explained by the organization preferences, this refers to the same set of criteria described in Table 2. Criteria wording variation are expected for organizations that adapt a framework and process for arriving at recommendations. Thus, some organizations may have addressed the magnitude of the problem before assessing the evidence pertaining to the EtD process, and therefore might not report this as an independent criterion in their EtD framework.

To illustrate this better, the GRADE-EtD framework presents a set of additional considerations, the so-called detailed judgments, that assist the panel when considering the evidence that underlie the main criteria [4]. For example, legal consequences, reported by Domus Medica as a specific criterion, would be covered by the GRADE-EtD framework as one of the detailed judgments included under feasibility (i.e., Are there important legal or bureaucratic constraints that make it difficult or impossible to

Table 1. Recommendation formulation information reported in the included guidance documents

Recommendation formulation information	*n* (%)
Contains specific section in the document	55 (81%)
Details of people involved	Entire panel/GDG, 8 (56%)
	Panel + other, 5 (7%)
	Subgroup of the panel, 3 (4%)
	Not specified, 19 (28%)
Technical team shares materials (e.g., evidence summaries) with the guideline panel ahead of meeting	27 (40%)
Technical team or someone else makes preliminary judgments on the different criteria (e.g., certainty of evidence)	25 (37%)
Technical team or someone else makes preliminary judgment on the strength of recommendations (e.g., strong, conditional)	15 (22%)
Technical team or someone else makes preliminary judgments about the direction of recommendations (e.g., in favor, against)	7 (10%)
Approach to grading the strength of recommendations	60 (88%)
Approach suggested for grading the strength of recommendations	GRADE, 35 (51%)
	NHMRC, 2 (3%)
	SIGN, 2 (3%)
	CEBM, 1 (1.5%)
	NICE, 1 (1.5%)
	USPSTF, 1 (1.5%)
	Adapted systems (5, 7%)
	GRADE + NICE, 1 (1.5%)
	NICE + SIGN, 2 (3%)
	GRADE + SIGN, 1 (1.5%)
	GRADE + AHRQ + USPSTF, 1 (1.5%)
	Other or not specified, 21 (36%)
Use of a framework for the EtD process	45 (66%)
Frameworks suggested for the EtD process	GRADE-EtD, 19 (42%)
	Other approaches, 26 (58%)
	Own approach, 20 (76%)
	NICE, 2 (8%)
	SIGN, 1 (4%)
	USPSTF, 1 (4%)
	SIGN + NICE, 1 (4%)
	GRADE + SIGN + AHRQ, 1 (4%)
Explicit method to reach agreement among panel members (e.g., consensus, nominal group techniques)	56 (82%)
Documentation of judgements made	27 (38%)

AHRQ, The Agency for Healthcare Research and Quality (USA); CEBM, The Center for Evidence-Based Medicine, based in the Nuffield Department of Primary Care Health Sciences at the University of Oxford; EtD, evidence to decision; GDG, guideline development group; GRADE, The Grading of Recommendations Assessment, Development, and Evaluation; NHMRC, National Health and Medical Research Council (Australia); NICE, The National Institute for Health Care and Excellence (UK); SIGN, Scottish Intercollegiate Guidelines Network; USPSTF, The US Preventive Services Task Force.

cover the intervention?). The same principle applies to ethical considerations, which is suggested by the CNEC, and is addressed as a detailed judgment under acceptability by the GRADE-EtD framework [i.e., Are there key stakeholders who would disapprove of the intervention morally, for reasons other than its effects on people's autonomy (such as regarding ethical principles such as no maleficence, beneficence, or justice)?] [4].

Table 2. Recommendation-related criteria in the EtD process

Criteria	All guidance documents 68 (100%)	Any framework 45/68 (66%)	No framework 23/68 (34%)	GRADE-EtD[a] 19/45(42%)	Other framework 26/45(58%)
Problem priority	40 (59%)	30 (67%)	10 (43%)	12 (63%)	18 (69%)
Desirable effects	49 (72%)	37 (82%)	12 (52%)	17 (89%)	20 (77%)
Undesirable effects	50 (73%)	38 (84%)	12 (52%)	17 (89%)	21 (81%)
Certainty of the evidence of effects	50 (73%)	38 (84%)	12 (52%)	18 (95%)	20 (77%)
Values (outcome importance)	35 (51%)	30 (67%)	5 (22%)	18 (95%)	12 (46%)
Balance of effects	39 (57%)	32 (71%)	7 (30%)	17 (89%)	15 (58%)
Resources required	37 (54%)	31 (69%)	6 (26%)	18 (95%)	13 (50%)
Certainty of evidence of required resources	17 (25%)	16 (36%)	1 (4%)	9 (47%)	7 (27%)
Cost-effectiveness	36 (53%)	30 (67%)	6 (26%)	16 (84%)	14 (54%)
Equity	11 (16%)	11 (24%)	0	8 (42%)	3 (11%)
Acceptability	19 (28%)	17 (38%)	2 (9%)	9 (47%)	8 (31%)
Feasibility	23 (34%)	18 (40%)	5 (22%)	9 (47%)	9 (35%)

GRADE, The Grading of Recommendations Assessment, Development, and Evaluation; EtD, evidence to decision.
[a] Guidance documents that suggested the GRADE-EtD frameworks for the process of formulating recommendations.

3.6. Drawing conclusions as part of the evidence to decision process

Overall, 41% of the guidance documents reported a process to summarize the judgments made about the different recommendation-related criteria. This step was more common in the documents that suggested the use of a framework (49%) compared to documents that did not (26%). We did not observe major differences in this step between those suggesting the use of the GRADE-EtD framework and another framework (47% vs. 50%, respectively). Similarly, the justification of the recommendation's strength and direction was more common in documents that used a framework (53% vs. 17%, respectively). This trend was also observed for other steps, such as considerations for relevant subgroups, implementation, monitoring, and evaluation, as well as about the formulation priorities for further research (Table 3).

3.7. Bivariate analysis

As stated in the methods section, we express the results of the bivariate analysis as the probabilities of addressing each recommendation-related criterion in the different categories of the EtD process (i.e., any framework vs. no framework; GRADE-EtD vs. other EtD framework; GRADE-EtD vs. no framework; other EtD framework vs. no framework). The following is a summary of the main findings. We refer the reader to Additional file 5 for a full description of the data.

The use of an EtD framework for guiding the recommendation formulation compared to no framework resulted in higher probability of incorporating both perspectives (OR 2.8, 95% CI 0.6–13.8) and subgroup considerations (OR 7.2; 95% CI 0.9–57.9). Similarly, the documents that incorporated the GRADE-EtD framework were more likely to incorporate these criteria in the recommendation formulation process than those that suggested another EtD

Table 3. Drawing conclusions

Criteria	All guidance documents 68 (100%)	Any framework 45/68 (66%)	No framework 23/68 (34%)	GRADE-EtD[a] 19/45 (42%)	Other framework 26/45 (58%)
Summary of the judgments made about the different criteria considered	28 (41%)	22 (49%)	6 (26%)	9 (47%)	13 (50%)
Justification of the recommendation	28 (41%)	24 (53%)	4 (17%)	11 (58%)	13 (50%)
Subgroup considerations	13 (19%)	12 (27%)	1 (4%)	7 (37%)	5 (19%)
Implementation considerations	38 (56%)	31 (69%)	7 (30%)	14 (74%)	17 (65%)
Monitoring and evaluation considerations	30 (44%)	25 (56%)	5 (22%)	14 (74%)	11 (42%)
Research priorities	19 (28%)	18 (40%)	1 (4%)	9 (47%)	9 (35%)

GRADE, The Grading of Recommendations Assessment, Development, and Evaluation; EtD, evidence to decision.
[a] Guidance documents that suggested the GRADE-EtD frameworks for the process of formulating recommendations.

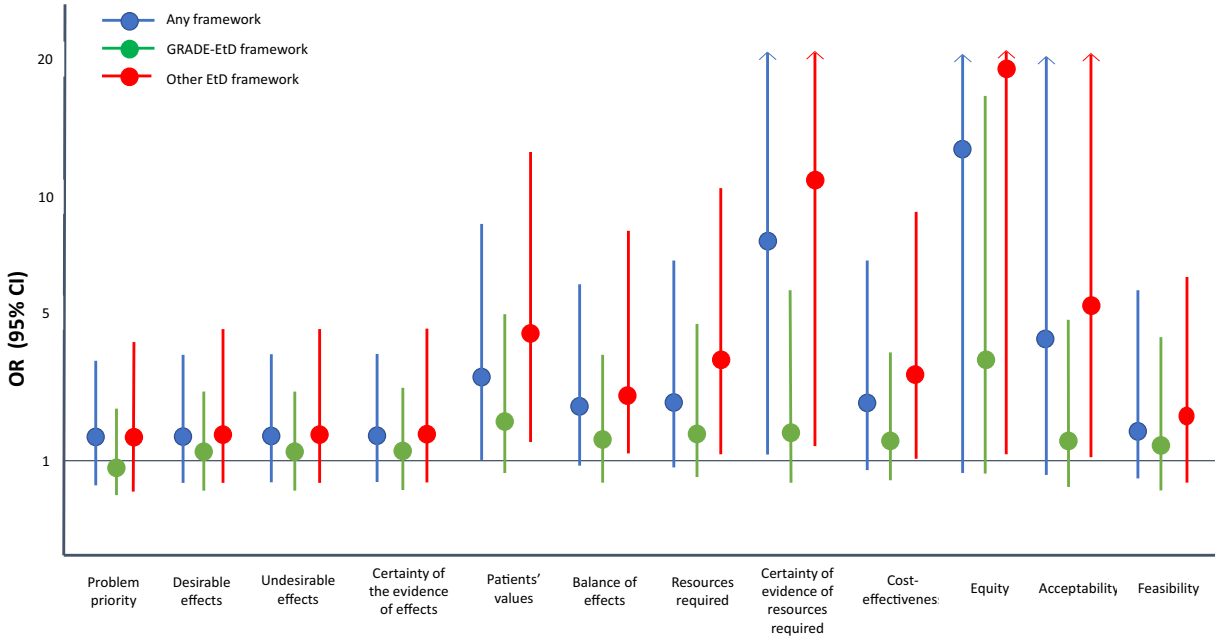

Fig. 4. Bivariate analysis for the recommendation-related criteria in the guidance documents. The odds ratios are based on the comparison of each category of EtD framework relative to the use of no framework. The vertical lines illustrate the 95% confidence intervals. GRADE, The Grading of Recommendations Assessment, Development and Evaluation; EtD, evidence to decision.

framework or no framework, with ORs ranging from 1.4 to 8.4, respectively.

The probability of using all the recommendation-related criteria identified in this study were higher in documents that suggested the use of any EtD framework relative to no framework, as well as for the documents that suggested the GRADE-EtD framework compared to those that suggested another framework or no framework (Fig. 4).

For instance, guidance documents that suggested the use of any EtD framework were more likely to consider patients' values in the recommendation formulation process compared to those that did not follow any framework (OR 3.1; 95% CI 1−8.9). The odds of including patients' values were twofold greater in documents that suggested the GRADE-EtD framework relative to those that suggested another framework (OR 2; 95% CI 0.8−5.4). The OR increased to 4 when the GRADE-EtD was compared to no framework (OR 4.4; 95% CI 1.4−13.9) (Fig. 4).

The guidance documents that suggested any EtD framework were more likely to present evidence on the balance between desirable and undesirable effects when formulating recommendations than documents that did not suggest any framework (OR 2.3; 95% CI 0.9−6.1). The odds were larger for the comparison of the documents that suggested the GRADE-EtD framework with those that used no framework (OR 2.9; 95% CI 1.1−8.6). Similar odds were observed for the criterion related to resources required.

Suggesting the use of an EtD framework was associated with higher odds of including cost-effectiveness considerations when formulating recommendations relative to the use of no framework (OR 2.5; 95% CI 0.9−7.1). Alike other criteria, the documents that suggested the use of the GRADE-EtD framework were more likely to incorporate cost-effectiveness considerations in the recommendation formulation process than those that suggested another framework (OR 1.6; 95% CI 0.6−3.9) or no framework (OR 3.2; 95% CI 1.1−9.8) (Fig. 4).

Of note, the odds of including recommendation-related criteria such as equity, acceptability, and feasibility followed the same pattern. That is, the use of any EtD framework and the GRADE-EtD framework resulted in higher probability of including those criteria when formulating recommendations relative to the use of no framework or another EtD framework.

Finally, and in line with the associations observed for the recommendation-related criteria, the documents that suggested both the use of any framework or the GRADE-EtD framework were more likely to provide a justification of the judgements made, implementation considerations, as well as monitoring and evaluation considerations than the documents that suggested no framework or another framework. Additional file 5 presents further details.

4. Discussion

4.1. Main findings

Our study presents a complete and systematic evaluation of the frameworks and processes suggested when moving from evidence to recommendations in guideline development. We documented the diversity of both frameworks and processes in this area. GRADE emerged as the most common approach for grading both the certainty of the

evidence and the strength of recommendations. Our analyses revealed that only slightly more than half of the organizations presented a structured process for formulating recommendations, in their guidance documents. However, the analyses also showed that there are gaps in important aspects like the sharing of preliminary material or judgements through the recommendation formulation process.

Fourteen recommendation-related criteria were identified in the guidance documents. The use of an EtD framework was associated with higher probabilities of addressing a comprehensive set of recommendation-related criteria compared to the use of no framework, especially for criteria like patients' values and preferences, equity, and acceptability. Similarly, the use of the GRADE-EtD framework was associated with higher probabilities of including those criteria when compared to other EtD frameworks or no framework. However, caution is advised when drawing inferences from the analysis due to the small number of guidance documents included in our analysis.

4.2. Our results in the context of previous research

Previous research suggests that the GRADE-EtD framework is a widely used tool for assisting comprehensive and transparent evidence-informed decision making [12,13,24,25]. Our findings confirm this notion, because the methods developed by the GRADE Working Group were the most used approach. For example, a case study conducted by members of the GRADE working group [12] showed that the GRADE-EtD coverage framework had been fully accepted by some stakeholders in the United States [24]. Similarly, a more recent real-time assessment of the use of the GRADE-EtD framework in guideline panels concluded that it was essential to structure panel meetings and ensure the consideration of all relevant criteria [17]. In that study, formal GRADE-EtD's criteria guided 94% of panel discussions, whereas other external criteria guided the remaining 6% of the discussions (e.g., clinical experience, political environment, and legal implications). Nevertheless, both the extent to which the criteria contained in the GRADE-EtD framework restraint discussions among panelists, and the need for additional criteria remain unexplored [17].

Despite the diversity of both, frameworks and processes in this area, there is still ongoing debate and research on the adequacy and validity of available frameworks for different types of decisions [26] and organizations [10]. Recently, different authors have suggested the need for including additional criteria in the GRADE-EtD framework. For example, burden of treatment [5] or human rights and sociocultural acceptability, equality and nondiscrimination, societal implications, and health system considerations by other authors [10]. The need for tailoring for different types of organizations or goals was considered in the GRADE-EtD frameworks in the original publications. For example, ethical considerations are considered as a detailed judgment under acceptability in the GRADE-EtD frameworks. Therefore, an organization might consider treating ethical considerations as a separate criterion, rather than as a detailed judgment under acceptability [4].

4.3. Limitations and strengths

Our study had some limitations. We do not rule out having excluded documents that organizations may have published in other languages. Nevertheless, we do not anticipate missing relevant information, as our searches covered a representative number of guidance documents derived from the most relevant organizations in the field.

Regarding the strengths, we followed methodological standards for the conduct and reporting of evidence synthesis studies, such as systematic search of the literature, double screening process, and a priori piloting of the data extraction procedures. Data were analyzed through iterative consensus between researchers. The project team was strategically positioned to carry out this study, as the researchers are experienced in working with guideline panels and are experts in evidence synthesis methods and the guideline development process.

4.4. Implications for practice and research

Our findings provide a systematic analysis of different criteria contained in the frameworks proposed for the recommendation formulation process among guidance documents for guideline development. Different stakeholders (e.g., patients and clinicians) may use our findings to better understand the processes the organizations followed to formulate recommendations (e.g., quality of the evidence assessment, COI management, voting dynamics, and EtD criteria considered). This will help them decide whether recommendations can be implemented in their own setting.

It is important to remark that the information extracted from the guidance documents was reliant on how complete the information about each criterion was presented in the documents, and so was our analysis. As an example of this, stating the organization used the GRADE-EtD framework did not equal the inclusion of all the recommendation-related criteria proposed by the framework; this was also true for other frameworks proposed by the organizations. Organizations may use our findings to either prepare their own guidance documents or to reassess their existing ones by identifying gaps in EtD criteria or to follow suggestions for improvement. In addition, guidance documents might provide clear definitions for each EtD criterion, so that GDG members understand what evidence to look for and how to put this in a decision-making context (e.g., considerations on equity or acceptability). All organizations should follow a more systematic and explicit methodology for formulating recommendations and ensure a complete reporting in their guidance documents. These might enhance transparency and credibility, enabling end users to determine how much confidence they

can have in the recommendations; facilitate later adaptation to contexts other than the ones where they were originally developed; and improve usability and communicability of the EtD frameworks.

We systematically documented the use of EtD frameworks in guidance documents; however, the extent to which the mention of the different criteria in the guidance documents translates into a real use in published guidelines is still unknown. Besides, because both guideline developers and panel members have expressed favorable experiences with the use of EtD frameworks [12,17,24], further qualitative research exploring the use of the frameworks among panelists and methodologist is also warranted. Our findings could also foster further research on the need of modifications of available frameworks.

5. Conclusion

The use of systematic and structured processes for moving from evidence to decisions is still limited among international organizations devoted to guideline development. The use of EtD frameworks facilitates the inclusion of relevant recommendation criteria. Among the structured frameworks, the GRADE-EtD framework is the most widely used approach, offering the most comprehensive perspective for evidence-informed decision-making processes. More complete and detailed reporting in the guidance documents is warranted.

CRediT authorship contribution statement

JFME: conceptualization, methodology, data extraction, formal analysis, original draft preparation, project administrator, rewriting, and editing. **PAC:** conceptualization, methodology, formal analysis, original draft preparation, and editing. **HS and SF:** methodology, supervision, original draft preparation, rewriting, and editing. **JB, JJYN, TPP, LP, MMB, DS, MS, JZ, CMG, YZ, and NCG:** data extraction, formal analysis, original draft preparation, rewriting, and editing. All authors read, provided feedback, and approved the final version of this manuscript.

Acknowledgments

The authors thank Andrea Juliana Sanabria and Anna Selva for their support and assistance with the literature searches. Besides, the authors thank the Norwegian Institute of Public Health for their support with the publication fee for this study.

Appendix A

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2022.06.004.

References

[1] Alonso-Coello P, Martinez Garcia L, Carrasco JM, Sola I, Qureshi S, Burgers JS. The updating of clinical practice guidelines: insights from an international survey. Implementation Sci 2011;6:1−8.
[2] IOM (Institute of Medicine). Clinical Practice Guidelines We Can Trust. Washington, DC: The National Academies Press; 2011.
[3] Schunemann HJ, Wiercioch W, Etxeandia I, Falavigna M, Santesso N, Mustafa R, et al. Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise. CMAJ 2014;186(3):E123−42.
[4] Alonso-Coello P, Schünemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ 2016;353:i2016.
[5] Dobler CC, Harb N, Maguire CA, Armour CL, Coleman C, Murad MH. Treatment burden should be included in clinical practice guidelines. BMJ 2018;363:k4065.
[6] Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D. Evidence and Value: Impact on DEcisionMaking–the EVIDEM framework and potential applications. BMC Health Serv Res 2008;8:1−16.
[7] World Health Organization. WHO handbook for guideline development. 2nd ed. Geneva: World Health Organization; 2014.
[8] Tannahill A. Beyond evidence–to ethics: a decision-making framework for health promotion, public health and health improvement. Health Promot Int 2008;23(4):380−90.
[9] Norheim OF, Baltussen R, Johri M, Chisholm D, Nord E, Brock D, et al. Guidance on priority setting in health care (GPS-Health): the inclusion of equity criteria not captured by cost-effectiveness analysis. Cost effectiveness and resource allocation. Cost Eff Resour Alloc 2014;12:1−8.
[10] Rehfuess EA, Stratil JM, Scheel IB, Portela A, Norris SL, Baltussen R. The WHO-INTEGRATE evidence to decision framework version 1.0: integrating WHO norms and values and a complexity perspective. BMJ Glob Health 2019;4(Suppl 1):e000844.
[11] Morgan RL, Thayer KA, Bero L, Bruce N, Falck-Ytter Y, Ghersi D, et al. GRADE: assessing the quality of evidence in environmental and occupational health. Environ Int 2016;92-93:611−6.
[12] Neumann I, Brignardello-Petersen R, Wiercioch W, Carrasco-Labra A, Cuello C, Akl E, et al. The GRADE evidence-to-decision framework: a report of its testing and application in 15 international guideline panels. Implementation Sci : IS 2016;11:1−8.
[13] Rosenbaum SE, Moberg J, Glenton C, Schünemann HJ, Lewin S, Akl E, et al. Developing evidence to decision frameworks and an interactive evidence to decision tool for making and using decisions and recommendations in health care. Glob Challenges 2018;2(9):1−9.
[14] American College of Chest Physicians (CHEST). Available at https://www.chestnet.org/Guidelines-and-Resources/About-CHEST-Guidelines/Guideline-Development. Accessed January 15, 2021.
[15] Canadian Task Force on Preventive Health Care. Canadian Task Force on preventive health care procedure manual. 2014. Available at https://canadiantaskforce.ca/methods/. Accessed January 15, 2021.
[16] Centro Nazionale per l'Eccellenza Clinica la Qualità e la Sicurezza delle Cure. Manuale metodologico per la produzione di linee guida di pratica clinica. 2019. Available at https://snlg.iss.it/wp-content/uploads/2019/04/MM_v1.3.2_apr_2019.pdf. Accessed January 15, 2021.
[17] Li SA, Alexander PE, Reljic T, Cuker A, Nieuwlaat R, Wiercioch W, et al. Evidence to Decision framework provides a structured "roadmap" for making GRADE guidelines recommendations. J Clin Epidemiol 2018;104:103−12.
[18] Higgins JPT TJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.2. Available at www.training.cochrane.org/handbook. Accessed January 20, 2021.
[19] Meneses-Echavez JF, Bidonde J, Yepes-Nuñez JJ, Storman D, Puljak L, Zając J, et al. Moving from evidence to decisions when

making recommendations in healthcare: a protocol of an analysis of clinical practice guidelines guidance documents. Charlottesville, VA: Open Science Framework; 2020.

[20] Vernooij RW, Sanabria AJ, Sola I, Alonso-Coello P, Martinez Garcia L. Guidance for updating clinical practice guidelines: a systematic review of methodological handbooks. Implementation Sci 2014;9:1−9.

[21] Selva A, Sanabria AJ, Pequeno S, Zhang Y, Sola I, Pardo-Hernandez H, et al. Incorporating patients' views in guideline development: a systematic review of guidance documents. J Clin Epidemiol 2017;88:102−12.

[22] Schunemann HJ, Oxman AD, Akl EA, Brozek JL, Montori VM, Heffner J, et al. Moving from evidence to developing recommendations in guidelines: article 11 in Integrating and coordinating efforts in COPD guideline development. An official ATS/ERS workshop report. Proc Am Thorac Soc 2012;9(5):282−92.

[23] R Core Team. R. In: A language and environment for statistical computing. Vienna, Austria: I: R Foundation for Statistical Computing; 2017.

[24] Dahm P, Oxman AD, Djulbegovic B, Guyatt GH, Murad MH, Amato L, et al. Stakeholders apply the GRADE evidence-to-decision framework to facilitate coverage decisions. J Clin Epidemiol 2017;86:129−39.

[25] Meneses-Echavez JF, Rosenbaum S, Rada G, Flottorp S, Moberg J, Alonso-Coello P. Users' experiences with an interactive Evidence to Decision (iEtD) framework: a qualitative analysis. BMC Med Inform Decis making 2021;21:1−11.

[26] Morgan RL, Kelley L, Guyatt GH, Johnson A, Lavis JN. Decision-making frameworks and considerations for informing coverage decisions for healthcare interventions: a critical interpretive synthesis. J Clin Epidemiol 2018;94:143−50.