Appendix C. Data extraction of peer-reviewed articles

RQ1

Clarity of presentation

Table C1 The Guideline Implementability Decision Excellence Model (GUIDE-M): a mixed methods approach to create an international resource to advance the practice guideline field

Publication identification	
Authors (year)	Brouwers et al. 2015
Country	Canada
DOI	https://doi.org/10.1186/s13012-015-0225-1
Publication description	
Design	Mixed methods
Objective	▪ To create a comprehensive and evidence-informed model of guideline implementability o Provide common (English-language) nomenclature (labels, definitions) appropriate across geographic jurisdictions and disciplines to facilitate communication and local progression of scientific inquiry o Seek support for the model from the international practice guideline community through a formalised external consultation. ▪ To identify priority areas for further investigation: o Examine the extent to which major international practice guideline development, reporting and evaluation tools address the components of the final model o Identify those components in the model that are perceived to be more and less well studied.
Summary/Overview	This paper reports on how the Guideline Implementability Decision Excellence Model was developed. The development included a realist review, survey of international stakeholders (guideline developers and users) to gather feedback and to refine the model, content analysis comparing the model to existing practice guideline tools, and strategy to prioritise areas of the model for further research by members of the research team. It can be used by practice guideline developers and other knowledge users to optimise the implementability of practice guidelines and by researchers to advance the science and application of practice guidelines.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	The final GUIDE-M comprises six levels: (i) 3 core tactics, (ii) 7 domains, (iii) 19 subdomains, (iv) 44 attributes, and (v) 40 subattributes and elements. At its highest level, the three core tactics serve as the foundation of the GUIDE-M: (i) Developers of Content (practice guideline group), (ii) Creating Content (creating the practice guideline, and (iii) Communicating Content (disseminating the practice guideline). **Tactic: Developers of content** The Developers of Content tactic is comprised of three domains: comprehensive, knowledgeable and credible, and competing interests. This tactic advises on the types and characteristics of participants who ought to be recruited to create a comprehensive multidisciplinary practice guideline development group, the expected skills of the group members, and issues related to competing interests of the group members. Optimised stakeholder involvement and participation strategies will increase credibility and acceptability of resulting recommendations. ▪ Domain: Comprehensive o Clinical experts

	▪ Multidisciplinary and multijurisdictional ▪ Researchers and users. ○ Target population ▪ Individual patients ▪ Family members ▪ Groups representing patients. ○ Decision makers ▪ Multidisciplinary and multijurisdictional ▪ Researchers and users. ○ Methodologists ▪ Practice guideline experts ▪ Knowledge synthesis experts ▪ Health economics experts ▪ Ethicists ▪ Implementation experts. ▪ Domain: Knowledgeable and credible ▪ Domain: Competing interests ○ Financial ○ Professional and or academic ○ Advocacy. **Tactic: Creating Content** The Creating Content tactic comprises two core domains: evidence synthesis and deliberations and contextualisation. The evidence synthesis domain outlines how to create the evidence base, how to report it, and how to ensure its currency. The deliberations and contextualisation domain refers to the process of moving from the evidence to recommendations through the careful consideration of the clinical applicability, values of practice guideline stakeholders (patients, providers, policy makers, society, and developers), and issues of feasibility in applying the recommendations. ▪ Domain: Evidence synthesis ○ How: execution of methods to develop evidence base ▪ Systematic and reproducible ▪ Valid and reliable. ○ What: completeness of reporting evidence base ▪ Question ▪ Eligibility criteria ▪ Literature search strategy ▪ Critical appraisal ▪ Data extraction ▪ Data synthesis ▪ Reporting. ○ When: currency of evidence base. ▪ Domain: Deliberations and contextualisation

	○ Clinical applicability ▪ Clinical relevance ▪ Patient relevance ▪ Implementation relevance. ○ Values – Acceptability and or preferences of: ▪ Patient/client o Acceptability o Preferences. ▪ Provider o Acceptability o Preferences o Clinical flexibility o Clinical judgement. ▪ Guideline developer o Acceptability o Preferences. ▪ Population and or societal o Acceptability o Preferences o Diversity o Equity. ▪ Policy o Acceptability o Preferences. ○ Feasibility ▪ Local applicability o Local adaption o Application tools and strategies. ▪ Resources o Availability of resources o Economic evaluation. ▪ Novelty o Compatibility o Knowledge and skills. **Tactic: Communicating content** The Communicating Content tactic includes two domains, language and format. This tactic focuses on specific strategies to communicate practice guideline information to optimise its implementability. This includes how to create clear, simple, and persuasive messages and how to format messages into key components while also considering presentation styles and the design of multiple versions to address the needs of different users. ▪ Domain: Language

	○ Simple ▪ Succinct ▪ Uncomplicated. ○ Clear ▪ Actionable ○ Specific ○ Unambiguous. ▪ Effective writing. ○ Persuasive ▪ Framing ▪ Relative advantage. ▪ Domain: Format ○ Version ▪ Tailored ▪ Modalities ○ Electronic (dynamic, static) ○ Non-electronic. ▪ Document types. ○ Components ○ Presentation ▪ Document layout ○ Visual elements ○ Length. ▪ Structure ○ Match system to the real world ○ Grouping and or ordering. ▪ Information visualisation ○ Display (tables, algorithms, pictures, graphical display) ○ Context (framing, vividness, depth of field, evaluability).
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A

What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	Additional core component: clarity of presentation. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: N/A – not applicable; N/R – not reported.

Table C2 Guideline uptake is influenced by six implementability domains for creating and communicating guidelines: a realist review

Publication identification	
Authors (year)	Kastner et al. (2015)
Country	Canada
DOI	https://doi.org/10.1016/j.jclinepi.2014.12.013
Publication description	
Design	Realist review
Objective	Primary Objective: To identify factors associated with the implementability of clinical practice guidelines and recommendations through a comprehensive and multidisciplinary perspective. Secondary Objective: To determine what characteristics are posited to improve uptake by whom and under what circumstances.
Summary/Overview	This study describes the synthesis of 278 articles from a multidisciplinary body of literature, with the authors identifying six domains of guideline implementability that affect uptake of recommendations within two broad categories (the creation and communication of guideline content). The authors also created narratives of key concepts, which can be used to develop tools to determine their impact on building better guidelines aimed at increasing their uptake and promoting better care.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	**Summary of hierarchical narratives to describe guideline implementability:** ▪ Stakeholder involvement o The guideline development group should have appropriate composition; its members should have relevant, unbiased expertise and suitable credibility, and potential conflicts of interest should be disclosed. The target population of end users (patients, the public) should be clearly defined, and their views and preferences considered. Ensuring stakeholder involvement during guideline development facilitates uptake. ▪ Evidence synthesis o To enhance guideline validity and reproducibility, the necessary evidence must be specified, the method of synthesis clearly defined (ideally evidence-based, valid, reliable, and transparent), and the timing of sequential syntheses appropriate (balancing timeliness of the guideline with stability over time). ▪ Considered judgement o The guideline development group must supplement evidentiary factors (quality, quantity, and consistency) with considered judgment, making complex trade-offs between the competing benefits and harms, side effects, and risks of various options for managing the disease or condition. They must also consider clinical applicability (whether the guideline responds to variability among patients) and the values and preferences of patients, developers, and care providers (i.e., the relative worth or importance of a health state and consequences such as benefits, harms, and costs of a decision). ▪ Implementation feasibility o Feasibility reflects local applicability (i.e., strategies for adapting recommendations to local conditions), consideration of resource constraints (availability of resources and other economic implications), and the influence of novelty of (or familiarity with) the guidelines, where novelty refers to the degree to which the recommendations propose behaviours considered unconventional by clinicians or patients). Feasible

	implementation of guidelines allows for flexibility in individual clinical decisions, are in agreement with users' opinions and skills, and are suitable for routine use in intended settings. ▪ Message ○ Guideline messages should use simple, clear, and persuasive language. Simplicity can be achieved by limiting the number of elements, the number of steps within each recommendation, or the number of conditional factors influencing performance (to prevent the quantity of information from exceeding available cognitive capacity). Clarity is enhanced by using specific, unambiguous language and by applying a direct writing style, with active voice, suitable punctuation, short sentences, and bullet lists to convey series of points, and without awkward breaks, abbreviations, hyphenation, redundancy, or unnecessary jargon. Finally, the guideline messages should be clinically convincing and should be framed in terms of potential gains, to convey their relative advantage over any previous approach. ▪ Format ○ Guidelines may be formatted in multiple versions (e.g., research-based, information-gathering, analytical tool; briefer guide for clinical education; short version for point-of-care clinical use; lay-language version for patients). Formatting involves determining which components to include in various versions (e.g., scope and purpose, target audience, guideline development panel, update plan, and implementation considerations), their presentation (e.g., proper placement of visual elements and document length), and structure (i.e., matching the system to the real world and bundling); and the overall visualisation of information.
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? OR What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component: clarity of presentation (including formatting for multiple versions).

	Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: N/A – not applicable; N/R – not reported.

Table C3 Focus groups and interviews with the public led to the development of a template for a GRADE Plain Language Recommendation (PLR)

Publication identification	
Authors (year)	Santesso et al. (2022)
Country	Canada/Poland
DOI	https://doi.org/10.1016/j.jclinepi.2021.09.018
Publication description	
Design	Qualitative interviews/focus groups
Objective	To test a patient-focused template for communicating recommendations and explore public perceptions of, and attitudes towards, guidelines and recommendations.
Summary/Overview	This paper describes the development of a template that provides direction to guideline developers when creating a Plain Language Recommendations that facilitates communication of complex information and ultimately enhances use by the public.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	**Plain language recommendations template:** ▪ Provide the recommendation, its strength (with a symbol), and an explanation to go with it. o E.g., Strong, conditional recommendation. ▪ List the people, populations, or individuals to whom this recommendation applies. o E.g., 'Who is this recommendation for?' ▪ Give a rationale for the strength of the recommendation. o Why strong – This recommendation is strong because [describe the balance of benefits and harms and other factors; use words such as 'clearly outweighs' or 'is' or 'are'] o Why conditional – This recommendation is conditional because [describe the balance of benefits and harms and other factors; use words such as 'probably outweighs' or 'may be']. ▪ Discuss additional considerations for using the recommendation. o Additional information – Some people may [describe other factors that may play a role when following the intervention, e.g., values, preferences, costs, feasibility]. ▪ Describe the possible benefits and harms, and provide a link to a summary of findings. o Benefits and harms [describe the change of benefits and harms using informative statements using the GRADE approach, which include the size of effect and certainty of the evidence; provide absolute numbers if helpful]. ▪ Explain implications of the guideline and topics or questions that one should discuss with their doctor. o E.g., 'What does this mean for you?' o What can you do? – It might be important to [include information about how to apply the recommendation in daily life; who to speak to; links to other resources] o Speak with your healthcare professional – To decide whether [include discussions about your own risk; what an outcome means; what may happen next].
RQ2: Description of quality measures/criteria for clinical practice guidance development	

What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component: clarity of presentation. Category of evidence: Grade C
Associated handbook(s)	N/R

Key: GRADE – Grading Recommendations Assessment and Development Evidence; N/A – not applicable; N/R – not reported**.**

Table C4 Strategies for disseminating recommendations or guidelines to patients: a systematic review

Publication identification	
Authors (year)	Schipper et al. (2015)
Country	The Netherlands, Switzerland
DOI	https://doi.org/10.1186/s13012-016-0447-x
Publication description	
Design	Systematic review
Objective	To assess what dissemination strategies are feasible to inform and educate patients about recommendations or guidelines.
Summary/Overview	This article gives an overview of tools and strategies to disseminate recommendations to patients. Key factors of success were a dissemination plan, written at the start of the recommendation development process, involvement of patients in this development process and the use of a combination of traditional and innovative dissemination tools. The lack of strong evidence calls for more research of the effectiveness of different dissemination strategies as well as the barriers for implementing a strategic approach to dissemination.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	**Dissemination to patients or patient organisations** **Dissemination plan** ▪ Ideally developed in parallel with the development of the recommendations ▪ Should be developed during the project and not at the end of the project ▪ The target audience needs to be clarified at the start of the project, which will subsequently determine the scope, objectives, format, style and wording of the recommendations as well as the tools for dissemination. **Lay version** ▪ A lay version enables patients to better understand the goals of treatment, the different treatment options and the benefits and risks of each option. ▪ Patients who have access to lay versions are better equipped to prepare themselves for the consultation with their healthcare provider and are expected to become an active partner in their own treatment. ▪ The following aspects should be taken into account when developing a lay version: o The message should be customised to the target audience o The information should be made relevant for the target audience (patients) o The information in the recommendations should be consistent, unambiguous and credible o The information in the lay version has to be readable for patients. ▪ Information can be simplified by using less medical and technical terms or by giving an explanation of the terms ▪ Information about where more in-depth information can be found should be included in the lay version ▪ Information in the lay version should use familiar words of one or two syllables, presented in active voice in the present tense and use short sentences of 15 words or less, and short paragraphs of ten lines or less ▪ A successful lay version provides clear, explicit and specific information and some key messages ▪ For international guidelines, lay versions of guidelines should ideally be translated into different languages through forward translation, back translation and patient testing.

	Combining dissemination strategies ▪ The use of passive dissemination strategies, such as a leaflet or brochure, has proven to be insufficient to educate patients or change daily routine because such information does not endure in the long term ▪ Strategies to disseminate lay versions need to be accompanied by the development of other materials ▪ Dissemination requires a combination of different, mutually reinforcing strategies; for example, the repetition of key messages from different credible sources such as well-known professionals ▪ Potential combinations of strategies include: organising press conferences, providing lay versions through public libraries, developing books to reach children and developing posters with 'trigger' stickers, or making a website endorsed by a well-known expert in the topic of the guideline'. ▪ Patient organisations can organise an annual national forum on a disease at which people share their experiences and take part in training and education programmes. Patient organisations can also provide telephone and online counselling and literature and other resources for patients and caregivers ▪ Provide automatic updates of new information and resources for patients who have opted in to receiving information, by interactive internet-based lectures and by developing a variety of learning tools like posters, summaries, handouts, pocket cards and slide sets for patients ▪ Education to patients should be pro-active, whereby face-to-face and contact by telephone can be used, in the same way as' education events for patients ▪ Guidelines should be easily searchable and accessible immediately ▪ Development of community 'champions' and knowledge brokers ▪ Combination of passive and active dissemination strategies is recommended, while taking into account social and cultural sensitivities and differences. **Patient involvement in developing recommendations** ▪ It is important to involve patients or patient organisations in the design and development of recommendations to enhance the dissemination of recommendations in healthcare and to local patient organisations. ▪ Patients should be involved from an ethical point of view; involvement is needed to give patients influence on the recommendations by incorporating their experiential knowledge and perspectives. ▪ The involvement of patients, if done properly, increases the comprehensiveness of the recommendations and facilitates the adaption of the recommendations to the target population by incorporating the knowledge and perspectives of patients. **Patient involvement in developing lay versions** ▪ The involvement of patients with different native languages enhances the likelihood that lay versions of international guidelines can be easily understood by many patients and that an English lay version can easily be translated into various languages. ▪ The G-I-N toolkit suggests that the translation of the English lay version in different languages should be done by patient organisations, using a heterogeneous group of patients with different disease status and educational levels. ▪ A participatory action research methodology could be used to involve patients in the development of guidelines and or lay versions. **Patient involvement in disseminating recommendations**

	▪ Successful change of clinical practice in accordance with the recommendations requires an effective implementation strategy in which the key stakeholders delivering and receiving care (including patients) are involved in the dissemination ▪ National patient organisations should, according to the G-I-N toolkit, disseminate the recommendations in their own countries. This can be done by using their own website, newsletters, brochures, other publications, phone calls, support groups, workshops, events, seminars, annual conferences, local or regional events, events for professionals and/or patients, press releases, print-ready ads, flyers or by including the recommendations in information packages provided to their members ▪ Personal stories of patients in media can also help to raise awareness of new recommendations. **Who should be involved?** ▪ G-I-N toolkit suggests that patients who have participated in the development of the recommendations can also actively contribute to the dissemination process. ▪ Establishment of permanent groups, networks or 'virtual panels' of patients. The network members are alerted when new recommendations or patient versions are published. They can raise awareness by distributing lay versions to health professionals, patients, patient organisations and members of the public. The network should in this case include members with different backgrounds. **Conditions for involvement** ▪ G-I-N toolkit suggests: o Informing patients about their role before participating o Clarifying expectations about the specific role of the patients and the time commitment required o Providing training in advance to prepare patients for their assigned role ▪ The training could be in technical areas such as how to understand the terminology or how to take part in the group effectively (e.g. assertiveness). o Supporting patients during the process ▪ Patients can be supported by providing networking opportunities for individuals or by providing a buddy. **Suggestions to make patient involvement more successful** ▪ Use of selection criteria in choosing patient representatives o A criterion may be the ability to consider the evidence objectively and to make recommendations that depart from preconceived views or self-interests. ▪ Involve patients in less traditional ways (e.g., as committee member) and choose more innovative ways such as the use of new media that better fit the patients' role, expectations and capabilities. The guideline development group may choose not to include a consumer representative but instead, may invite patients to review draft documents or attend a group meeting or internet forum to share their perspectives ▪ Provide patients with sufficient information and knowledge before and during the project. This empowers them to become effective partners in the dissemination and implementation process.
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A

What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component: clarity of presentation. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: G-I-N – Guidelines International Network; N/A – not applicable; N/R – not reported.

Gender equity

Table C5 Gender balance in WHO panels for guidelines published from 2008 to 2018

Publication identification	
Authors (year)	Bohren et al. (2019)
Country	Australia, Switzerland
DOI	https://doi.org/10.2471/BLT.18.226894
Publication description	
Design	Systematic review
Objective	To assess the gender composition of guideline contributors for all WHO guidelines published from 2008-2018.
Summary/Overview	This study assessed the gender composition of guideline panels for all WHO guidelines published over an 11-year period. Overall, women were under-represented across most roles. While gender balance appeared to be improving in recent years, only half of the guidelines had a guideline development group composed of 40.1–60.0% female members, and more than one-third of the guidelines had a guideline development group composed of less than 40.0% female members.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	**Gender balance in the guideline development and external review groups** (see Table C5a below)

Table C5a Individuals contributing to WHO guideline development, by gender and role in 219 guidelines published between 2008 and 2018.

Role	**No. (%) of individuals**			
	Female	Male	Unknown gender	Total
Guideline development group member (including chair)	2241 (45.6)	2606 (53.1)	65 (1.3)	4912
Guideline development group chair	96 (39.5)	145 (59.7)	2(0.8)	243
Guideline methodologist	179 (63.5)	102 (36.2)	1 (0.4)	282
WHO steering group member	1637 (43.9)	2044 (54.9)	40 (1.1)	3721
WHO headquarters staff	1366 (45.0)	1640 (54.0)	29 (1.0)	3035
WHO regional or country office staff	271 (39.5)	404 (58.9)	11 (1.6)	686
External review group staff	1045 (44.4)	1279 (54.3)	31 (1.3)	2355
Other	896 (43.5)	909 (44.1)	254 (12.3)	2059

- Under-representation of women across most roles in the guideline panels
- Only half of the guidelines had a guideline development group composed of 40.1–60.0% female members, and more than one-third of the guidelines had a guideline development group composed of less than 40.0% female members.

RQ2: Description of quality measures/criteria for clinical practice guidance development	

What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component: gender equity. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: N/A – not applicable; N/R – not reported; WHO – World Health Organization.

Table C6 The representation of women on Australian clinical practice guideline panels, (2010-2020)

Publication identification	
Authors (year)	Shalit et al. (2023)
Country	Australia
DOI	https://doi.org/10.5694/mja2.51831
Publication description	
Design	Survey study
Objective	▪ To assess the gender composition of Australian clinical practice guideline development panels. ▪ To explore guideline development-related factors that influence the composition of panels.
Summary/Overview	This study describes an analysis of the composition of Australian clinical practice guideline panels by gender. The overall proportion of women as contributors to guideline development (all roles) was 44.8%, as guideline panel members 41.1%, and as guideline panel chairs 42.1%. The proportion of female guideline panel members was less than 40% for 179 of 335 guidelines (53% of guidelines): Women comprised fewer than 40% of guideline panel members for 17 of 59 NHMRC-approved guidelines (29%). The proportion of guideline panel members who were women was below 40% for 22 of 50 cancer guidelines (44%), 31 of 39 cardiology guidelines (80%), 18 of 27 nephrology guidelines (67%), four of 13 paediatric medicine guidelines (31%), and two of 17 women's health guidelines (12%).
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	**Representation of women in clinical practice guidelines** Overall guideline contributors ▪ Of 7,472 contributors named in 335 guidelines (median per guideline, 12 people; interquartile range, 6–22 people), 3,514 were men (47.0%) and 3,345 women (44.8%). Gender could not be determined for 612 people (8.2%), and one person used they/them as pronouns in their online professional profile and did not specify a gender identity. Guideline panel chairs ▪ Of the 511 guideline panel chairs (6.8% of named guideline contributors), 280 were men (54.8%) and 215 were women (42.1%); gender could not be determined for 16 people (3.1%). ▪ The overall proportion of male chairs was larger than that of female chairs in all years apart from 2017 (38 female chairs of 61 chairs, 62% women). Guideline panel members ▪ Of 5,039 guideline panel members (67.4% of named guideline contributors), 2,566 were men (50.9%) and 2,071 women (41.1%); gender could not be determined for 402 people (8.0%). ▪ The proportion of women among guideline panel members was <40% for 179 guidelines (53% of guidelines considered; including 36 guideline panels with fewer than 10% women [11%]), 40–60% for 85 guidelines (25%), and >60% for 71 guidelines (21%).

	▪ Overall, a greater proportion of panel members were male in all years apart from 2011 (288 female panel members of 572 panel members, 50% women) and 2017 (297 female panel members of 536 panel members, 55% women) ▪ The median proportion of women among guideline panel members was below 50% each year, with the exceptions of 2017 and 2018. ▪ Women comprised fewer than 40% of guideline panel members for 17 of 59 NHMRC-approved guidelines (29%) and 162 of 276 guidelines without NHMRC approval (59%; $p<0.001$). Women comprised fewer than 40% of guideline panel members for 32 of 78 guidelines that used GRADE methodology (41%) and 147 of 257 guidelines that did not (57%; $p=0.043$).
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component: gender equity. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: GRADE – Grading of Recommendations, Assessment, Development, and Evaluations; N/A – not applicable; N/R – not reported; NHMRC – National Health and Medical Research Council.

Health equity

Table C7 GRADE equity guidelines 2: considering health equity in GRADE guideline development: equity extension of the guideline development checklist

Publication identification	
Authors (year)	Akl et al. (2017a)
Country	International
DOI	https://doi.org/10.1016/j.jclinepi.2017.01.017
Publication description	
Design	Mixed methods: review of guideline methodology articles, conceptual articles, and consensus.
Objective	To provide guidance for guideline developers on how to consider equity at key stages of the guideline development process.
Summary/Overview	This paper provides guidance for the key stages at which guideline developers could consider equity. These include setting priorities, guideline group membership, identifying the target audience(s), generating the guideline questions, considering the importance of outcomes and interventions, deciding what evidence to include and searching for evidence, summarising the evidence and considering additional information, wording of recommendations, and evaluation and use.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	Guidance is based on the stages of the guideline development process described in the comprehensive 18-item checklist suggested by Schünemann et al.[(1)] Although equity could potentially be taken into consideration at each of these 18 stages, specific suggestions are provided on how to consider equity for the following most relevant stages (the number in parenthesis indicates the position of each step within the sequence suggested by Schünemann et al., with order not reflecting relative importance): ▪ Setting priorities (step 2) ○ Consider dedicating part of or a whole guideline (as opposed to no part) to the care of disadvantaged populations. ▪ Guideline group membership (step 3) ○ Include representatives of the disadvantaged populations in the different guideline groups, particularly the voting panel ○ Ensure the method for recruitment of group members considers representatives of all relevant disadvantaged populations ○ Recruit a methodologist who is familiar with and mindful of equity issues ○ Ensure the chair of the voting panel is familiar with equity issues. ▪ Identifying the target audience(s) (step 5) ○ Specify relevant disadvantaged populations when identifying the target audience(s) ○ Involve representatives of disadvantaged populations when identifying the target audience(s). ▪ Generating the guideline questions (step 8) ○ Consider equity when specifying elements of the PICO questions ○ Consider “good-practice statements” that could help address equity issues. ▪ Considering the importance of outcomes and interventions (step 9)

	○ Involve representatives of disadvantaged populations in rating the importance of interventions and outcomes ○ Search selected databases (e.g., UK DUETs, COMET) for outcomes rated as important by disadvantaged populations ○ Consider separate recommendations for disadvantaged populations if their values and preferences are thought to differ substantively to the point of affecting the strength and/or direction of recommendation. ▪ Deciding what evidence to include and searching for evidence (step 10) ○ Seek evidence specific to disadvantaged populations, for example, baseline risks specific to those groups ○ Consider including evidence derived from fields other than health (e.g., social science) that address disadvantaged populations ○ Search literature published in the language relevant to the disadvantaged population. ▪ Summarising the evidence and considering additional information (step 11) ○ Consider the PROGRESS-plus elements when synthesising the evidence (i.e., Place of residence, Race/ethnicity/culture/language, Occupation, Gender/sex, Religion, Education, Socioeconomic status, or Social capital. The "Plus" in "PROGRESS-Plus" indicates that other characteristics, such as age, disability, sexual orientation, time-dependent situations, and relationships, may also be at the heart of health inequities as they pertain to disadvantaged populations) to refer to the factors associated with social disadvantage. ○ Follow the PRISMA-equity statement when reporting systematic reviews ○ Consider information on resource use, cost, effect on equity, feasibility, and acceptability from the perspective of disadvantaged populations. ▪ Wording of recommendations (step 14) ○ Be as specific as possible in defining the population to maximise the understanding that it applies to a disadvantaged population (when applicable) ○ Include the necessary remarks following the recommendation to ensure its appropriate implementation in disadvantaged populations ○ Ensure that language is used carefully so that the recommendation does not stigmatise already disadvantaged populations. ▪ Evaluation and use (step 17) ○ Produce tools to facilitate implementation and use among disadvantaged populations ○ Monitor and audit implementation and use among disadvantaged populations.
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A

RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component: health equity. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: COMET – Core Outcome Measures in Effectiveness Trials; DUETs – Database of Uncertainties about the Effects of Treatments; N/A – not applicable; N/R – not reported; PICO – Population/patient, Intervention, Comparison, Outcomes; PRISMA – Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PROGRESS-Plus – place of residence, race/ethnicity/culture/language, occupation, gender/sex, religion, education, socioeconomic status, social capital also including personal characteristics associated with discrimination, features of relationships, and time-dependent relationships

Table C8 Methodological guidance for incorporating equity when informing rapid-policy and guideline development

Publication identification	
Authors (year)	Dewidar et al. (2022)
Country	International
DOI	https://doi.org/10.1016/j.jclinepi.2022.07.007
Publication description	
Design	Mixed methods (consensus, guidance review)
Objective	To provide guidance for the consideration of equity in a rapid review through examples of published COVID-19 rapid reviews.
Summary/Overview	This study identified areas where researchers could consider equity in rapid review development, such as stakeholder involvement in the review process, incorporating equity in team values and composition (including training), developing research questions to assess health inequities, conducting literature searches in appropriate databases to capture health equity evidence, collecting, analysing and evaluating health equity evidence and adoption of reporting guidelines that capture health equity information. This guidance could be used by groups and agencies responsible for rapid decision-making during emergencies to ensure that populations experiencing inequities are considered when informing policy and developing guideline recommendations.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	Incorporation of equity in rapid reviews conducted for rapid policy and guideline develop. Equity can be incorporated in the following areas of rapid review: 1. Engaging relevant stakeholders in conducting, designing and interpreting the review ▪ Team members should ensure diversity in team expertise and lived experience ▪ Modification of best practices to suit the timelines for reviews in pandemics, such as engaging individuals with experience or developing alternative approaches for training ▪ Stakeholders could also critique the study question to ensure it is relevant to policy or clinical practice ▪ Stakeholders could also identify interdisciplinary libraries and grey literature sources, provide insights on participant characteristics, study design features or identify outcomes that may be relevant to addressing equity, provide their perspectives on the relevance of key findings, and participate in appropriately disseminating the evidence (e.g., plain language summaries). 2. Reflecting on equity in team values and composition ▪ Equity values should be formulated as part of the team values and culture ▪ To ensure that a supportive environment is provided within the research team, research team members should discuss participating in at least one of the potential equity, diversity and inclusion training activities ▪ Team members should consider completing training that improves team capacity building and effective stakeholder engagement ▪ Including people with lived experience relevant to the review topic as part of the team strengthens the review process by incorporating context-specific understanding, based on experience and tacit understanding of an issue. 3. Developing research questions to assess health inequities

	▪ Review authors could focus on a population experiencing inequities (the PROGRESS-Plus framework can aid in the identification process) or consider such populations as subgroups of interest ▪ The review authors should supplement these decisions with an *a priori* definition of how the intervention is expected to influence health equity for the identified populations. 4. Conducting searches in relevant inter-disciplinary databases ▪ Reviewers may need to consider searches in social databases or other inter-disciplinary databases from low- and middle-income countries to identify relevant evidence for socio-economic impacts on different populations ▪ Local databases and governmental and non-governmental websites could be investigated as potential grey literature sources ▪ Review authors should also ensure that search terms capturing equity-related content have been included within the search string ▪ Authors should aim to adopt validated filters relevant to their topic when searching for studies that are equity relevant. If there are no validated filters, authors should be mindful that unvalidated equity filters could limit their searches and risk missing relevant evidence. 5. Collecting data for equity ▪ The review authors should capture elements of study design to evaluate the nature of participant inclusion or exclusion as it may influence the applicability of the results for populations experiencing inequities ▪ Review authors should also assess if the chosen methodology and theories by the primary authors articulate possible pathways to addressing inequities ▪ Reviewers should collect data on sample characteristics such as context and population demographics that interact with other contextual elements and influence health inequities ▪ Capturing information on retention and attrition across populations experiencing inequities is also essential, as this may affect the generalisability of the review findings ▪ When possible, outcome data should be collected in both relative and absolute differences between groups. 6. Analysing evidence on equity ▪ The review authors should consider checking for baseline imbalance across PROGRESS-Plus factors for quantitative evidence. ▪ When appraising qualitative evidence, the review authors should consider if the primary research authors designed the question to assess outcomes related to health equity (i.e., impact of intervention, acceptability) by evaluating if and how they included populations experiencing inequities. The sources of the quotations within the primary qualitative research and how they were analysed should also be considered. ▪ Additional synthesis methods may be needed to address questions related to equity. Subgroup analyses are usually conducted. Other methods such as moderator analysis, meta-regression and sensitivity analysis may be more relevant, depending on the question and how the review authors decide to consider equity at the question conceptualisation stage. All these analyses should be pre-planned,

	accompanied with a rationale linked to an analytical framework (i.e., logic model) and adhere to reporting standards to ensure their credibility. ▪ The analyses should also be pre-planned and accompanied by theory-based rationales. 7. Evaluating the applicability of the findings to populations experiencing inequities or other settings ▪ Influence on health equity should be interpreted from the findings of the review. The principles of interpretation include: (1) evaluating who was included in the studies and judging if they are representative of people with the condition; (2) if there were any differences in recruitment, retention, effects found, what are the potential impacts on policy and practice. 8. Adhering to reporting guidelines for communicating review findings ▪ Adopting reporting guidelines such as the PRISMA-Equity, SAGER guidelines and International Committee of Medical Journal Editors when constructing the review encourages the completeness of reporting of equity-relevant information.
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? OR What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component of health equity (and rapid policy). Category of evidence: Grade D.
Associated handbook(s)	N/R

Key: N/A – not applicable; N/R – not reported; PRISMA-Equity – Preferred Reporting Items for Systematic Reviews and Meta-Analyses Health Equity Statement; PROGRESS-Plus – place of residence, race/ethnicity/culture/language, occupation, gender/sex, religion, education, socioeconomic status, social capital also including personal characteristics associated with discrimination, features of relationships, and time-dependent relationships; SAGER – sex and gender equity in research.

Table C9 GRADE equity guidelines 4: considering health equity in GRADE guideline development: evidence to decision process

Publication identification	
Authors (year)	Pottie et al. (2017)
Country	International
DOI	https://doi.org/10.1016/j.jclinepi.2017.08.001
Publication description	
Design	Mixed methods: review of guideline methodology articles, conceptual articles, and consensus.
Objective	To provide detailed guidance on how to incorporate health equity within the GRADE evidence to decision process.
Summary/Overview	This paper describes guidance on how to consider health equity in the EtD process of guideline development. Authors suggest two approaches: (1) assessing the potential impact of interventions on equity and (2) incorporating equity considerations when judging or weighing each of the evidence to decision criteria.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	Considering the impact on health equity may be required, both in general guidelines and guidelines that focus on disadvantaged populations Two approaches are suggested to incorporate equity considerations: 1. Assessing the potential impact of interventions on equity 2. Incorporating equity considerations when judging or weighing each of the evidence to decision criteria. <u>Formulating the question</u> ▪ Health equity can be considered within panel formulation and development of equity-sensitive questions. ▪ Analysis focusing on resource limited settings, specific populations (e.g., disadvantaged) can lead to different recommendations. ▪ Equity-sensitive questions may highlight health equity in different ways, such as by focusing on a disadvantaged population or by assessing effects across subgroups that may be disadvantaged. <u>Is the problem a priority?</u> ▪ Many diseases disproportionately affect low-income populations (e.g., cardiovascular disease, infectious disease, motor vehicle accidents). ▪ In addition, disadvantaged populations may face limited access to appropriate treatment and care (e.g., cataracts in low-income countries or tuberculosis in indigenous populations). ▪ In these cases, guideline panels might consider these conditions as high priority because they place value on redressing unfair burden carried by disadvantaged populations. ▪ Similarly, undernutrition and its correlation to poor educational consequences in low-income neighbourhoods may mean school feeding programmes may reduce health inequities and will be seen as a priority. <u>How substantial are the undesirable and desirable anticipated health effects?</u> ▪ Assessing both desirable and undesirable effects for specific disadvantaged populations may lead to different recommendations. ▪ When assessing the size and importance of effects, it is important to consider both relative and absolute effects. ▪ Even though relative effects are similar across populations, absolute effects may differ if baseline risks are substantially different for disadvantaged populations.

	What is the overall certainty of the evidence of effects? ▪ Indirectness is an important criterion when considering disadvantaged populations; for example, the available evidence may be limited for disadvantaged populations as these groups are often underrepresented in clinical and population research. ▪ If there is a strong rationale to expect the relative treatment effect to substantially vary between the general population and one or more disadvantaged groups, the panel could rate down the certainty of effect for indirectness. ▪ Panels should be cautious in rating down certainty of effects for indirectness unless there is a strong rationale, such as evidence of subgroup effects or effect modifiers, for why the evidence would not apply to the specific disadvantaged group. ▪ Similarly, it is possible (but generally less likely) that a greater body of evidence may exist for a disadvantaged population (e.g., as a consequence of more research having been conducted in response to a high burden in this population). Under such a situation, it is possible that greater certainty in the evidence may result in a different recommendation that is targeted at disadvantaged groups. ▪ However, potential for stigma or other adverse consequences that may arise from targeted recommendations should be considered. Is there important uncertainty about or variability in how much people value the main outcomes? ▪ Certain disadvantaged populations may value the main outcomes differently than the general population. Explicitly considering the views of disadvantaged groups concerning the importance of outcomes helps panels to decide whether adapting recommendations for subgroups is warranted. Do the desirable effects outweigh the undesirable health effects? ▪ Deciding whether the desirable effects outweigh the undesirable health effects depends on the assessments for the three previous criteria relating to the size and importance of effects, the certainty of the evidence for those effects, and certainty and variability in valuation of outcomes for different populations. ▪ When relevant population subgroups need to be considered in guidelines, this may require making this assessment separately for each population subgroup of interest. Resource use ▪ When considering health equity in assessing resource use (costs), an intervention with lower resource requirements and lower effectiveness, compared to the next best alternative, may sometimes be preferred in some settings to promote health equity. ▪ Guideline development panels might identify economic evaluations that have explicitly considered health equity in one of four ways: 1. Background reviews 2. Health equity impact assessment 3. Analysing opportunity costs for equity considerations 4. Equity weighting of health outcomes. ▪ Equity weighting methods include adding an equity weight or social welfare function whereby society is prepared to sacrifice overall health benefits to promote a more equitable distribution of these benefits; statistical programming that quantifies the opportunity cost of equity; and multicriteria decision analysis based

	on trade-offs between a range of criteria of which cost may be one, that is, through the process of discrete choice analysis. ▪ Economic evaluations that have explicitly considered health equity may be helpful when rating equity-sensitive questions. <u>What would be the impact on health equity?</u> 1. Are there groups or settings that might be disadvantaged in relation to the problem or intervention (option) of interest? 2. Are there plausible reasons for anticipating differences in the relative effectiveness of the intervention (option) for disadvantaged groups or settings? 3. Are there different baseline conditions across groups or settings that affect the absolute effectiveness of the intervention (option) or the importance of the problem for disadvantaged groups or settings? 4. Are there important considerations that people implementing the intervention (option) should consider in order to ensure that inequities are reduced, if possible, and that they are not increased? ▪ The criterion of potential impact on health equity focuses on both relative and absolute effectiveness for the intervention, the importance of the problem, and identifying considerations to ensure health inequities are reduced or not increased. ▪ However, data on specific disadvantaged groups and settings may not be available. In such cases, considering health equity across the previous criteria may help in answering the above questions. ▪ By explicitly examining the potential impact on health equity, a panel may discover differential effects on disadvantaged populations (e.g., health equity in relation to specific characteristics: economic status, employment or occupation, education, place of residence, gender or ethnicity). ▪ The four questions to assess potential impact on health equity may also be considered at other points during the evidence to decision process. ▪ An organisation or panel might decide in advance to make explicit detailed judgements for one or more criteria. Alternatively, they may only make explicit detailed judgements when these judgements help resolve disagreements. <u>Is the intervention acceptable to key stakeholders, including patients?</u> ▪ The less acceptable an intervention is to key stakeholders (including disadvantaged people), the less likely it is that it should be recommended, or if it is recommended, the more likely it is that an implementation strategy will be needed to address concerns about acceptability. ▪ Acceptability of an intervention may differ for different populations due to four main reasons: 1. Different distribution of benefits, harms, and costs 2. Timing of outcomes (e.g., now or in the future) 3. Different values about the relative importance of desirable and undesirable health effects 4. Ethical considerations, such as patient autonomy and justice. ▪ It is important to collect input from key stakeholders from disadvantaged populations in considering acceptability, since assumptions by panel members may be biased by their personal experience. ▪ Differences in acceptability between groups, if strong enough, could impact the strength of a recommendation for specific groups and should be reported and addressed during implementation. ▪ Key stakeholders of disadvantaged populations may include communities, health workers, and interest groups.

	▪ Certain stakeholders may disapprove of the intervention on cultural or traditional, religious or moral grounds. ▪ Variation in acceptability may demand additional implementation resources for community-based testing to prevent inequities. <u>Is the intervention feasible to implement?</u> ▪ Panels may identify unique health equity concerns relating to access, barriers to implementation in certain settings, and programme feasibility. ▪ Detailed judgements may include consideration of barriers to the sustainability of the option(s). These barriers include guideline factors, individual health professional factors, patient factors, professional interactions, incentives and resources, capacity for organisational change, and social, legal, and political factors. <u>Drawing conclusions</u> ▪ The importance of each EtD criterion for a recommendation can vary depending on the population and limitations in resources. To formulate a recommendation, a panel must consider the trade-offs and health equity concerns of each judgement. ▪ Assessing health equity across EtD criteria will help to decide among three possible types of recommendations: 1. A general recommendation that can be applied across different populations and settings. Assessment of health equity across the criteria may increase the confidence of the panel that a general recommendation is warranted and that the intervention is applicable for disadvantaged populations and settings. 2. A general recommendation that can be accompanied with subgroup and implementation considerations to promote health equity or mitigate worsening health inequities. 3. A separate recommendation for a specific disadvantaged population when evidence of meaningfully different effects for a specific setting or subgroup is identified. <u>Methodologic challenges</u> ▪ One challenge in assessing potential impact on health equity across EtD criteria is that within a specific disadvantaged population, such as people living on low income, there is heterogeneity in the experience of health inequity due to variation across other social determinants such as gender, employment, and age. Thus, the assessment for one population may vary depending on the intersection of these different characteristics. Guideline panels may need to prioritise the most relevant disadvantaged populations for the problem or condition of interest and recognise that there is no "one-size-fits-all" approach to mitigating health inequities. ▪ EtD process requires a stepwise approach. It may not be possible or desirable to assess equity for every criterion. For example, if an intervention is simply not acceptable for a disadvantaged population, then it is redundant to assess feasibility and resource use considerations for this population. The acuity to assess and address the needs of disadvantaged groups may require additional research training and systematic processes. Assessing health equity may not be warranted for all criteria or scenarios.
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A

RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component: health equity. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: EtD – evidence to decision; GRADE – Grading Recommendations Assessment and Development Evidence; N/A – not applicable; N/R – not reported.

Table C10 Methodology for senior-proof guidelines: A practice example from the Netherlands

Publication identification	
Authors (year)	Van Munster et al. (2017)
Country	The Netherlands
DOI	https://doi.org/10.1111/jep.12738
Publication description	
Design	Consensus methodology – nominal group technique
Objective	To develop a methodology to increase the focus on older people in the development of guidelines. This methodology is intended both for general guidelines and for guidelines that are specifically aimed at older people as a target group, and it may be used by a guideline working group without specialists in aged care.
Summary/Overview	This study describes a methodology, developed by a core group of internal medicine and geriatrics specialists and guideline methodologists, to increase the focus on older people in guideline development during the guideline preparation and development stages. The framework for crafting clinical practice guidelines that are relevant to the care and management of people with multimorbidity was used as an example for the relevant topics.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	As treatment may differ depending on the health state of the older person, four categories of older adults aged ≥65 years were considered (category 4 may overlap with categories 2 and 3): 1. Relatively healthy older people 2. Older people with comorbidity (that is, one additional specific (interfering) comorbid condition or disease). Interfering comorbidity refers to coexisting conditions that impact upon the disease or disorder that is the subject of the guideline. For example, osteoporosis in older people with COPD, as corticosteroid medications used in connection with COPD negatively impacts bone density and increases the risk of fractures. 3. Older people with multimorbidity (several coexisting chronic conditions), where the guideline does not concern a single, specific comorbid condition 4. Frail older people. <u>Preparation stage</u> ▪ The extent to which a specific focus on older people is required or desirable within a guideline may be determined by applying the below criteria: o the prevalence of the disease or condition o the level of suffering o the social relevance o the expectation that a guideline might improve the quality of care. ▪ The decision to focus on a specific group of older people as defined above depends on whether or not prevalent interfering comorbidity or multimorbidity occurs in the target population of the guideline ▪ Examination of evidence specific to older people in the form of systematic reviews ▪ Availability of adjacent guidelines or expertise from subject matter experts, such as specialists in geriatric medicine ▪ State the decision and rationale to focus (or not) specifically on older people in the introduction to the guideline, alongside other considerations for the demarcation of the subject of the guideline

	▪ The target group (i.e., older people, older people with 1 specific [interfering] comorbid condition, older people with multimorbidity or frail older people) and the starting questions that take this into account should be explicit. ▪ Considerations for participation of a specialist in older persons' care and organisations for older people in guideline working groups o When most key questions on older people pertain to relatively healthy older people or older people with an interfering comorbid condition that is (medically) treated, deployment may be limited to the provision of feedback within the framework of peer review o In case of an interfering comorbid condition, it is also possible to consider a medical specialist in the field of this specific comorbid condition o When most key questions on older people pertain to those with multimorbidity, a choice between participation during the development stage and providing feedback during the development stage depends on the nature of the multimorbidity o When most key questions on older people pertain to frail older people, then participation or the provision of feedback during the development stage by a specialist in geriatric medicine is an obvious choice. <u>Development stage</u> ▪ Inventory and analysis of problem areas o The beliefs, values, or preferences of older patients are of importance for the preparation of the appropriate recommendations, in particular, during the weighing of the pros and cons of treatment options o Involvement of patient organisations for older people and organisations of informal carers within problem areas is therefore desirable. o Additionally, a literature research is recommended, preferably in MEDLINE and the Cochrane Library. "Patient preferences," "patient satisfaction," "patient experiences," "patient participation," "physician-patient relations," and "shared decision making" may be used as (controlled) keywords, in combination with the subject, and possibly limited to systematic reviews and/or 'older[tiab]' or 'AGED[Mesh]'. ▪ Starting questions and outcome measures o Next to outcome measures such as mortality and morbidity, it is desirable to consider outcome measures such as quality of life, hospitalisation, cognitive functioning, functional status, or treatment burden in relation to older people o The required follow-up duration regarding the outcome measures should be considered because this may be of importance in connection with the prognosis o It is recommended to determine the relative importance of the outcome measures for this target group. The guideline development group should explicitly state which outcome measures are crucial and important. ▪ Systematic search for evidence o Search strategies should be tailored for each of the four categories of older adults as appropriate.

	○ The search results may be narrowed down in a later stage to articles with a high average age, or studies in which meta-analyses (of individual patient data), subgroup analyses, or results of meta-regression are reported in relation to the treatment effect and interactions between the treatment effect and comorbidity and/or age categories of older people. ▪ Summarising study characteristics ○ For clinical studies that focus on the treatment of diseases that frequently occur in older people, it should be considered whether these studies produce the required evidence for older people by verifying the following: ▪ Whether older people are well represented in the studies, ▪ Whether it concerns a representative population of older people, and ▪ Whether the results (effectiveness and toxicity) of interventions in older people are specified separately. ▪ Determining the quality of evidence for every outcome measure as well as the overall quality of evidence ○ The overall quality of evidence may vary for the different categories of older people due to: ▪ Use of various outcome measures ▪ The relative importance of an outcome measure may differ ▪ The quality of evidence may differ for each outcome measure, among others, due to indirect evidence. ○ In the event of a lack of subgroup analyses or results of meta-regressions that give information on the effects of (diagnostic or therapeutic) interventions in older people, it may be checked whether effects for a specific outcome measure in studies that present overall results vary according to the average age of the study population or the percentage of patients with, for example, the relevant comorbidity in each study. ▪ Formulating and implementing recommendations ○ When examining the balance of desired and undesired effects for (the various categories of) older people, consider the following: ▪ Absolute benefits and absolute risks of interventions; ▪ Medicine interactions; ▪ Drug-disease interactions; ▪ Treatment burden; ▪ Physical, mental, and emotional capacity of a patient; ▪ Prognosis (remaining life expectancy, functional status, years spent with limitations, quality of life); and ▪ The values and preferences of patients. ○ When implementing recommendations for older people, consider what specific impeding factors (at the level of individual care providers, at the level of the organisation, and at the level of the system) exist for their application, and in particular, how they may be addressed.
RQ2: Description of quality measures/criteria for clinical practice guidance development	

What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component: health equity. Category of evidence: Grade D.
Associated handbook(s)	N/R

Key: COPD – Chronic Obstructive Pulmonary Disease; N/A – not applicable; N/R – not reported.

Table C11 GRADE equity guidelines 1: considering health equity in GRADE guideline development: introduction and rationale

Publication identification	
Authors (year)	Welch et al. (2017a)
Country	International
DOI	https://doi.org/10.1016/j.jclinepi.2017.01.014
Publication description	
Design	Mixed methods: review of guideline methodology articles, conceptual articles, and consensus.
Objective	To provide the rationale for three subsequent articles considering health equity explicitly in GRADE guidelines throughout the process (Akl et al.), rating certainty of evidence (Welch et al.), and in the evidence to decision framework (Pottie et al.).
Summary/Overview	This paper describes the methodology and rationale for the consideration of health equity throughout the guideline development process.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	When to think about health equity in guideline development: ▪ Question formulation and priorities, scope definition and group membership o Consideration of health equity ▪ What are the priorities of disadvantaged groups or populations, and how does this affect the key questions? ▪ Evidence assessment (i.e., in systematic review of the evidence) o Consideration of health equity ▪ Analysis of differences of effect (baseline risk and effectiveness) ▪ Targeted interventions ▪ Quality assessment of directness. ▪ Evidence to recommendation o Consideration of health equity ▪ Balance of likely impact on health equity with other factors.
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR**	N/A

What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component: health equity. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: GRADE – Grading of Recommendations, Assessment, Development and Evaluations; N/A – not applicable; N/R – not reported.

Table C12 GRADE equity guidelines 3: considering health equity in GRADE guideline development: rating the certainty of synthesized evidence

Publication identification	
Authors (year)	Welch et al. (2017b)
Country	International
DOI	https://doi.org/10.1016/j.jclinepi.2017.01.015
Publication description	
Design	Mixed methods: review of guideline methodology articles, conceptual articles, and consensus.
Objective	To describe a conceptual framework for how to consider health equity in the GRADE guideline development process.
Summary/Overview	This paper describes how health equity should be considered in the guideline development process and recommends that research on health inequity and guidelines prioritise the identification of examples where health equity has been considered explicitly in guidelines.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	Five methods can be used to assess health equity with the GRADE approach: 1. Include health equity as an outcome 2. Consider patient-important outcomes relevant to health equity 3. Assess differences in the magnitude of effect in relative terms between disadvantaged and more advantaged individuals or populations 4. Assess differences in baseline risk and hence the differing impacts on absolute effects for disadvantaged individuals or populations 5. Assess indirectness of evidence to disadvantaged populations and or settings.
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? OR What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	

Reviewer notes	RQ1: Additional core component: health equity. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: GRADE – Grading Recommendations Assessment and Development Evidence; N/A – not applicable; N/R – not reported.

Health outcome descriptors

Table C13 Development and use of health outcome descriptors: a guideline development case study

Publication identification	
Authors (year)	Baldeh et al. (2020)
Country	International
DOI	https://doi.org/10.1186/s12955-020-01338-8
Publication description	
Design	Mixed methods
Objective	To determine which aspects of the development, content and use of health outcome descriptors are valuable to guideline developers
Summary/Overview	This study describes the experiences of health outcome descriptor development for a healthcare guideline and provides guidance for future efforts in this area. The resulting standardised health outcome descriptor format may be useful for facilitating a common understanding of the outcomes chosen for the healthcare questions covered in a guideline, and thus improving the transparency of the guideline methods used.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	▪ Health outcome descriptors as developed by researchers at McMaster GRADE Centre. ▪ The rating and selection of important health outcomes occurs before the search for evidence because it helps narrow the search. ▪ Clear definitions and agreement by a guideline panel on what constitutes an outcome is required to search for evidence, balance benefits and harms, communicate with the public, and conduct research. ▪ To promote transparency of guideline development methods, guideline end-users require clear explanations of what constitutes each important outcome. **Template for health outcome descriptors** <u>Name of health outcome- importance rating</u> ▪ Symptoms o List most common symptoms. ▪ Time Horizon o Describe how long symptoms will persist for and how they might change over time o Describe approximate timing of relevant healthcare. ▪ Testing and treatment o Describe relevant healthcare or interventions. ▪ Consequences o Describe relevant consequences resulting from the health outcome or relevant healthcare. The format is purposefully designed to be concise; written at a Grade 8 reading level (as indicated by the Flesch–Kincaid readability tests) from the perspective of the healthcare recipient, who is the primary beneficiary of any healthcare guideline.
RQ2: Description of quality measures/criteria for clinical practice guidance development	

What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component: health outcome descriptors. Category of evidence: Grade C
Associated handbook(s)	N/R

Key: GRADE – Grading of Recommendations Assessment, Development and Evaluation; N/A – not applicable; N/R – not reported.

Table C14 Development and application of health outcome descriptors facilitated decision-making in the production of practice guidelines

Publication identification	
Authors (year)	Wiercioch et al. (2021)
Country	International
DOI	https://doi.org/10.1016/j.jclinepi.2021.04.016
Publication description	
Design	Consensus methodology – nominal group technique
Objective	▪ To describe methods for developing health outcome descriptors (HODs) in the context of real-world experience when developing a clinical practice guideline. ▪ To describe how the approach facilitated prioritising and rating health outcomes and how it was incorporated in decision making by guideline panels.
Summary/Overview	This study describes an approach for HOD development and provides instruction on how guideline developers may implement it. The approach may be useful to provide a reference point for a common understanding of outcomes considered in the development of a guideline. HODs explicitly define health outcomes based on experiences of affected individuals or patients.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	**Template for health outcome descriptors** <u>Health Outcome Descriptor Title, including lay health outcome descriptor title</u> ▪ Symptoms o Common symptoms due to the health state. o Note that grade of severity can be labelled mild, moderate or severe, and will be used as a descriptor of the HOD, not as part of the symptom. ▪ Time Horizon o Within which timeframe does the health state occur? ▪ Testing and treatment o Which tests and treatments are commonly applied for this health state? ▪ Consequences o Including prognosis and side effects. When a health outcome is found to present significant variability in its characteristics (e.g., acute and chronic pain, or mild and severe allergic reaction), then separate HODs were created to reflect this variation. There are two HOD templates, one for importance rating and one for utility rating. Both templates had the same structure and domains, but the template for utility rating was intended to expand on the outcome importance rating with additional details (e.g. 2 – 4 bullet points per domain rather than one or two points). This was intended to allow for more specificity in eliciting a utility value on a 0 – 100 visual analogue scale, where a value of 0 indicates the state of being dead and 100 indicates the state of full health. ▪ **Structure of health outcome descriptor** The structure and details of the HODs allow the definition of outcomes across a range of diseases, including describing different severities of an outcome that would result in different consequences for a person experiencing that outcome.

	What this means for guideline developers is that outcomes considered in decision-making that may typically be defined with a simple label (e.g. allergic reaction), but can have a broad range of consequences, can be more specifically defined with a HOD. HODs accomplish this by covering the common symptoms experienced due to a health outcome, how long the health outcome lasts, the tests and treatments that a person experiencing the outcome is expected to undergo, and consequences such as long-term effects. ▪ **Use of HODs in the evidence-to-decision process** Use of the HODs provided a reference point throughout the guideline development process when considering health outcomes. HODs provided explicit definitions of outcomes and could be contextualised into Summary of Findings tables presenting research evidence on effects of interventions and in EtD frameworks to facilitate the panels' judgements and decision making. While making judgements about EtD criteria to arrive at a recommendation, panels were able to view the specific HODs. This helped ensure that panellists were considering the same outcome (i.e. with the same consequences and severity) during their deliberations. The HODs served to inform the panels' decisions and judgements regarding magnitude of desirable effects (or benefits) and undesirable effects (or harms). When it came to deciding about the balance between desirable and undesirable effects, HODs were intended to ensure a common understanding between panellists to allow for appropriate weighing of that balance (e.g., in favour or against an intervention).
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component: health outcome descriptors. Category of evidence: Grade D.
Associated handbook(s)	N/R

Key: EtD – evidence to decision; HOD – health outcome descriptor; N/A – not applicable; N/R – not reported.

Quality indicators

Table C15 Approaches of integrating the development of guidelines and quality indicators: a systematic review

Publication identification	
Authors (year)	Langendam et al. (2020)
Country	International
DOI	https://doi.org/10.1186/s12913-020-05665-w
Publication description	
Design	Systematic review
Objective	▪ To identify and describe approaches that are utilised to develop guideline recommendations and quality indicators, i.e., in an integrated framework. ▪ To evaluate the effects of an integrated guideline and quality indicator development approach on individual health outcomes as well as process and structure outcomes (e.g., time required to develop recommendations and quality indicators, feasibility, acceptability by key stakeholders and development costs).
Summary/Overview	The authors conducted an extension and update of a previous systematic review to identify approaches to the integrated development of guidelines and related quality indicators. They identified 30 articles describing these approaches; however, in general, these were not based on well-defined conceptual frameworks and lacked full integration of the two areas. Key findings indicate a lack of coherence between the two fields and heterogeneity in methods. For example, the quality of the guidelines was not assessed in the majority of the articles. This suggests that although quality indicator development is often done on the basis of recommendations by reputable organisations, the suitability and quality of the recommendations may not coincide with the goals of quality indicators.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	▪ Quality indicators (QIs) are used to monitor guideline adherence as they measure structures, processes and health outcomes of care. ▪ Ideally, development of the quality indicators should be integrated in the guideline development process to establish a direct link with the recommendations. **Methods of guideline-based quality indicator development matched to the steps in the guideline development process** The GIN-McMaster Guideline checklist was used to match the domains in the guideline process with the accompanying domain in QI development. 1. Organisations, budget, planning and training ▪ Funding: 14 of the 30 approaches included in the systematic review were publicly funded, 3 were privately funded, 2 were funded both publicly and privately, 1 did not receive funding and funding was not reported for 10. 2. Priority setting ▪ See item 5. 3. Guideline group membership ▪ Criteria for selection of Guideline Development Group members were reported in six articles. Four articles reported selection of a multidisciplinary panel, including methodological competence, experience in quality improvement, policy decision making and knowledge translation. All 6 articles mentioned clinical

	expertise for the specific healthcare topic as competence. Criteria for selection of QI development panel members were mentioned in 15 articles. Clinical expertise was a criterion in all 15 articles, methodological experience was reported in six of the 15 articles. Patients/lay persons were part of three panels. Six reports did not use a formal panel and in nine articles the criteria were unclear. 4. Establishing guideline group processes: Group processes were not described in any article. 5. Identifying target audience and topic selection ▪ Fifteen articles reported criteria for selecting the QI topics and the target audience. The criteria, and phrasing of the criteria, varied from article to article. Criteria for topic selection included relevance for the specific care domain (e.g., primary care), quality of care gap, sound evidence base, feasibility, availability, measurability, reliability, validity, regulatory requirements, unknown quality adherence, expected impact on quality of life, costs, work load, disease severity, potential to reduce health inequities and covering all aspects of the care process. 6. Consumer and stakeholder involvement ▪ Patients were included in the QI selection process in nine of the 30 articles. 7. Conflict of interest: Conflicts of interest considerations for the QI development process were not mentioned in any of the papers. 8. PICO question generation ▪ See item 5 (QI topic selection). 9. Considering importance of outcomes and interventions, values, preferences and utilities ▪ Seventeen articles reported criteria for QI selection. In nine of these articles patient outcomes, health gain or importance or clinical effectiveness were part of the criteria. 10. Deciding what evidence to include and searching for evidence ▪ See item 11. 11. Summarising evidence and considering additional information ▪ All articles used evidence-based guideline recommendations as the starting point for QI development (this was an inclusion criterion). Thirteen articles reported additional sources, e.g., literature searches for existing QI sets or available data. In seven articles QI development was based on multiple guidelines, and in nine articles QI were developed based on one guideline. In one article this was not specified. 12. Judging quality, strength or certainty of a body of evidence ▪ Eight of the 13 articles that reported criteria for selecting recommendations as a basis for QI development used level of evidence as a criterion; three of the eight approaches used GRADE and suggested that only strong recommendations should be considered for translation into QI. 13. Developing recommendations and determining their strength ▪ See item 12. 14. Wording of recommendations and of considerations of implementation, feasibility and equity ▪ Feasibility was mentioned as a criterion for selecting QI (10 articles). Equity was mentioned once, as a criterion for selecting the topic for which the QIs were developed. 15. Reporting and peer review: Reporting and peer review of QI were not mentioned. 16. Dissemination and implementation.

- Implementation: 12 articles reported a QI implementation plan as part of their approach, mostly consisting of development of tools and software, and audits.

17. Updating: Updating of QI was not explicitly mentioned in any of the papers.

Guideline-based QI development reporting standard items and reporting of these criteria in the method papers

The GIN Reporting standards for guideline-based performance measures includes nine items of the quality indicator development process. These items were used to provide a detailed overview of the 30 included articles

- Guideline selection: criteria
  - Selection of guidelines was based on topic and
    - evidence-based development (n = 18)
    - methodological quality of the guideline (n = 2)
    - use of GRADE (n = 1)
    - structured format (n = 1)
    - no other criterion (n = 1)
    - unclear (n = 7).
- Guideline selection: appraisal of guidelines
  - AGREE (n = 8)
  - criteria not fully specified (n = 4)
  - not reported (n = 18).
- Selection of guideline recommendations
  - based on topic (n = 2)
  - impact on patient outcome (n = 4) / burden of illness (n = 1) / clinical utility (n = 1) / available treatment (n = 1)
  - relevance (n = 4) / appropriateness (n = 1)
  - value for money (n = 1)
  - practice variability (n = 1)
  - scope for improvement (n = 1) / gap in quality of care (n = 1)
  - priority / feasibility for implementation (n = 3)
  - validity (n = 2) / reliable (n = 1)
  - (high) level of evidence (n = 8) / adequate scientific proof (n = 1)
  - direct link to aim of guideline (n = 1)
  - common to more than one guideline (n = 1)
  - unclear (n = 3)
  - no selection (n = 6).
- Selection of performance measures from recommendations
  - formal panel method (n = 11)
  - other or informal consensus method (n = 13)
  - not reported (n = 2)
  - unclear (n = 4).
- Core attributes of performance measures (criteria for selecting QI)
  - relevance (n = 4)

	○ potential for improvement (n = 9) / likely to change current practice (n = 2) / gap in quality of care (n = 2) / importance for health care (n = 4) ○ burden of illness (n = 2) / improving patient outcomes (n = 9) ○ cost-effectiveness (n = 4) ○ influenced by service provider (n = 3) ○ appropriateness (n = 1) ○ evidence base/scientific soundness (n = 7) ○ (strength of) association with patient important outcome (n = 2) ○ feasibility (n = 7) ○ no risk for unintended consequences (n = 3) ○ unambiguous definition (n = 2) / clear (n = 1) ○ data routinely collected (n = 1) ○ measurable (n = 4) / interpretable (n = 1) / actionable (n = 2) ○ applicable (n = 3) / acceptable (n = 1) / adherence (n = 1) ○ reliable (n = 6) / face validity (n = 2) / construct validity (n = 1) / content validity (n = 2) ○ precision (n = 1) ○ minimum bias (n = 1) ○ not reported (n = 4) ○ unclear (multiple criteria per methodological framework) (n = 3). ▪ Specification of performance measures ○ denominator: population eligible to receive the clinical interventions, numerator: desired intervention and subset of population that should receive it (n = 6). ○ based on algorithm (n = 1) ○ formulation of numerator and denominator in line with formulation of recommendation (n = 1) ○ numerator and denominator including risk adjustment factors (n = 3) ○ clinical researcher drafted an expanded text for each recommendation, using logical operators (e.g., 'AND' and 'OR') to link descriptive statements to produce numerators and denominators (n = 1) ○ method not specified in detail (n = 15) ○ not reported (n = 3). ▪ Intended use of performance measure ○ quality improvement (n = 10) ○ quality of care delivered (n = 2) ○ monitoring compliance with guideline (n = 4) ○ implementation of care (n = 1) ○ clinical audit (n = 1) ○ pay for performance program (n = 1) ○ not specified (n = 8) ○ unclear (n = 1) ○ not reported (n=2). ▪ Practice test of performance measures

	○ planned (n = 18) ○ retrospective (n = 2) ○ implicit (n = 1) / ad hoc (n = 1) ○ not reported (n = 8). ▪ Review and evaluation of performance measure ○ plan for evaluation and updating (n = 3) ○ evaluation including criteria for retiring (n = 1) ○ mentioned, but not explained in detail (n = 2) ○ evaluation not reported, often because QIs were developed but not yet implemented (n = 24). ▪ Composition of the panel ○ monodisciplinary (n = 2) ○ multidisciplinary (n = 23) ○ panel composition not reported (n = 5). ▪ Composition of the panel: patient involvement ○ yes (n = 10) ○ no (n = 17) ○ depends on guideline (n = 1) ○ not reported (n = 2).
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? OR What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core components: quality indicators.

	Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: AGREE – Appraisal of Guidelines for REsearch & Evaluation; GDG – guideline development group; GIN – Guidelines International Network; GRADE – Grading of Recommendations, Assessment, Development, and Evaluations; N/A - not applicable; N/R – not reported; QI – quality indicator.

Table C16 Reporting standards for guideline-based performance measures

Publication identification	
Authors (year)	Nothacker et al. (2016)
Country	International
DOI	https://doi.org/10.1186/s13012-015-0369-z
Publication description	
Design	Mixed methods
Objective	To develop and agree on a set of core methodological standards for guideline-based performance measures with an associated rationale.
Summary/Overview	The authors developed a reporting standard for guideline-based performance measures with nine criteria, using formal written consensus methods (two Delphi rounds) and systematic literature search to identify core criteria.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	**Reporting standards for guideline-based performance measures** ▪ Guideline selection o State the currency of the guideline(s) used for guideline-based performance measure development and state if it/they meet the criteria set out by the G-I-N. Describe the guideline quality using a validated guideline appraisal tool, such as AGREE II. o Indicate additional sources, if used and the rationale for their use. ▪ Selection of guideline recommendations o State the strength of evidence and/or the grade of recommendation qualifying the guideline recommendations to be used for guideline-based performance measures. ▪ Selection process of performance measures from guideline recommendations o Describe clearly and in detail the methods used to develop the performance measures from the supporting clinical guideline recommendations. ▪ Core attributes of performance measures o State if the following attributes within the development process of guideline-based performance measures were considered: ▪ Relevance (as a minimum: potential for improvement/clinical relevance) ▪ Scientific soundness (as a minimum: the evidence supporting the measure) ▪ Feasibility (as a minimum: clarity of definition and measurability). ▪ Specification of performance measures o State that numerator and denominator of the guideline-based performance measure are specified unambiguously and in detail. ▪ Intended use of performance measures o State if there is a clear description of the intended use of the performance measure (quality improvement, quality assurance with or without accountability purposes, pay for performance) and at what level in the health system it is used (local, regional, national). ▪ Practice test of performance measures

	○ If a practice test (piloting) is carried out prior to using the guideline-based performance measure, provide a full description of the process. If no practice test is done, provide the rationale for this. Provide information about any other validation process in use. ▪ Review and re-evaluation of performance measures ○ Report the currency of the performance measures in use. State if there are criteria for deciding to change or stop using performance measures. ▪ Composition of the panel deciding on guideline-based performance measures ○ Describe clearly the composition of the panel deciding on guideline-based performance measures with information on participation of multidisciplinary experts, stakeholders in the field, experts in quality measurement, and patient representatives.
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? OR What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: AGREE ll – Appraisal of Guidelines for REsearch & Evaluation II; G-I-N – Guidelines International Network; N/A – not applicable; N/R – not reported.

Table C17 International experiences in the development and implementation of guideline-based quality indicators: a qualitative study

Publication identification	
Authors (year)	Nothacker et al. (2021)
Country	Germany
DOI	https://doi.org/10.1136/bmjopen-2020-039770
Publication description	
Design	Qualitative study
Objective	▪ To explore the implicit and explicit processes in the development of guideline-based QI. ▪ To explore the international experiences in the development of guideline-based QI. ▪ To explore the factors that hinder or facilitate the development of guideline-based QI from clinical guidelines internationally.
Summary/Overview	In this study, international guideline experts were interviewed to reveal insights into all development steps of guideline-based QI. Authors identified common patterns, feasible approaches and factors facilitating or hindering the development of guideline-based QI. Four main topics were identified: (i) organisation/context of guideline and QI development process, (ii) panel composition and decision making, (iii) QI selection criteria/attributes and (iv) intended use and implementation of QI.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	Guideline-based QI can be defined as 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'. The following four topics were found to be the most important for QI development methods from guidelines: **1. Organisation/context of guideline and QI development process** ○ Various ways of organising the development of guideline-based QI ▪ Develop guideline and QI in one process ▪ Develop the guideline first and the QI in a later process with a different team, although individual experts may be involved in both processes ▪ Guideline organisation acting at the local level could adopt existing QI developed by third party groups related to the guideline topic ▪ Develop guidelines and QI simultaneously, giving consideration to QI development early on from the beginning of the guideline process rather than towards the end. ▪ Develop informal audit measures with the guideline and develop national QI in a separate process led by a national group ▪ To drive QI development forward in a guideline organisation, a dedicated person with QI experience was deemed crucial by several clinicians and methodologists. ○ QI development and especially implementation usually requires external cooperation and resource support depending on the institutional position ○ Suggested approach: ▪ Have a person or a team in the guideline organisation or a collaborating organisation that is responsible for the process of development of guideline-based QI

	▪ Seek cooperation with partners in quality improvement. Adapt the QI process according to resources available for development and implementation, for example: o Consider alignment with existing QI o Limit development process to QI selection. **2. Panel composition and decision making** o Selection of a QI panel that represents different professional groups and settings, is balanced in its representation of interest groups and has a high level of knowledge about QI o Involvement of professionals/education ▪ Involvement of different professionals is favourable, including members with methodological knowledge as well as collaboration with future implementers. The group needs education and a shared understanding of the process. ▪ Patient participation: There are different views on patient participation in QI development and direct patient relevance of QI. Half the interviewees stated that patient involvement was crucial. On the other hand, it was argued that most patients on the panels were not 'vocal about measurement' and some careful selection or prior education on measurement was necessary. ▪ Decision making: There is no shared concept of a particular decision-making process. Difficulties of reaching consensus were mostly connected with feasibility aspects including expected decisions of payers. ▪ Collaborating with future implementers: A collaborative approach that involves feedback loops with clinical stakeholders in the field was emphasised by several interviewees as a facilitator for QI development and implementation. Pilot testing was seen by clinicians and methodologists as an essential opportunity for adapting QI based on feedback from implementers, although it was realised in only about half of the organisations due to time and resource constraints. o Suggested approach: ▪ Recruit a panel that is representative of the relevant health professionals, include members with methodological knowledge as well as future implementers. Train the QI developing team on QI methodology, possibilities and limitations. ▪ Discuss patient perspectives, patient participation and relevance to patients of guideline-based QI at the beginning of the QI process. ▪ If the GRADE approach is used in the guideline, QI can be linked to prioritised patient relevant outcomes. Instruct patients participating in the QI process and methodology as part of the panel. **3. QI selection criteria/attributes** o The evidence base of QI is most important and should be transparent. o To assess the relevance of QI, knowledge of regional quality gaps/ variability is needed. o Measurability as a key attribute remains a challenge. o Difficulties are reported concerning measuring patient reported outcomes, shared decision making and individualised care for QI based on weak recommendations. o To assess measurability, piloting in cooperation with future implementers is preferable. In terms of feasibility, existing QI should be considered. QI should also be acceptable to clinicians. o Suggested approach:

	▪ Use explicit evidence-based guidelines for QI development with transparent evidence base for each recommendation ▪ Ensure familiarity with regional/national quality gaps for assessing the need for QI preferably using healthcare data, or if not available, using expert consensus ▪ Pilot QI with those who will/must implement them. ▪ Consider alignment with existing QI ▪ Consider 'resource use/expense' also for clinicians as one criterion when assessing feasibility. **4. Intended use and implementation of QI** ○ Guideline groups (and institutions) aim for their important recommendations to be implemented to ensure clinically meaningful quality improvement. ○ If a QI is suitable for a certain purpose, it can only be appraised after a pilot test. ○ Suggested approach: ▪ Make sure that guideline-based QI recommendations are made available to decision makers in charge of QI ▪ Conduct piloting to make sure the QI is suitable for the intended use.
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? OR What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component: quality indicators. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: GRADE – Grading of Recommendations, Assessment, Development, and Evaluations; N/A – not applicable; N/R – not reported; QI – quality indicator.

Table C18 Bringing two worlds closer together: a critical analysis of an integrated approach to guideline development and quality assurance schemes

Publication identification	
Authors (year)	Piggott et al. (2021)
Country	International
DOI	https://doi.org/10.1186/s12913-020-05819-w
Publication description	
Design	Mixed methods
Objective	To identify key issues in the integration of guidelines and quality assurance (QA).
Summary/Overview	This study describes the findings of a mixed methods approach centring on the development of a methodological framework for integrating guideline and QA schemes. The authors present seven key themes resulting from an iterative process and conclude that the integration of guidelines and QA is feasible. The integration of guidelines and QA presents clear benefits and that the challenges identified are surmountable. Of note, extensive methodological work to more effectively integrate and evaluate guidelines and QA schemes is required. In particular, attention to the use of evidence and transparency of integrated processes is critical.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	**Key considerations for integrating guideline and QA schemes** A QA scheme is a collection of requirements to support healthcare services in improving the quality of care provided. Development of guidelines and QA schemes in health traditionally operate in two different worlds despite the fact that they are both critical, interdependent health improvement processes, designed to ensure that the best possible health recommendations are developed and that the recommended interventions ultimately meet the specified quality standards. ▪ Evidence-based integrated guideline and QA frameworks ○ Integrated guideline and QA schemes should be based on the best available evidence (for QA schemes grey literature may be more relevant), usually synthesised and assessed in a systematic review. ○ Evidence reviews should include not only the benefits but potential harms, and other considerations important for decision-making (e.g., GRADE EtD framework criteria). ○ The study participants recommended that an integrated framework should begin with a model on the health topic (e.g., logic model/analytical pathway/disease model/analytical PICO framework) that addresses the issue comprehensively from prevention to diagnosis and treatment. Within the analytical pathway, quality gaps should be identified for which quality indicators are deemed important to improve healthcare processes and outcomes. ▪ Transparency ○ The steps involved in linking evidence to guideline and quality assurance recommendations by an integrated framework should be clearly documented in a transparent manner. ○ Building from the GRADE EtD framework, the incorporation of QA scheme development would enable clear and transparent linkage of quality indicators to the evidence and guideline recommendations. ▪ Declaration of interests and management of conflicts ○ Both financial and intellectual conflicts of interest for participants in an integrated guideline and QA scheme should be clearly declared and appropriately managed to limit interference in the process.

	▪ Selection of quality indicators (QIs) ○ Follow reporting standards on the selection of quality indicators from guideline recommendations. ○ Prioritise patient-important QIs that are measurable, feasible, cannot be easily manipulated and are sensitive to change. ○ First select a small but sufficient number of candidate QIs for review. If QIs are not derived from guideline recommendations, clearly document their source and rationale. ▪ Retirement of QIs ○ A QI should be retired if, for example, it no longer addresses a quality gap, it becomes associated with unintended consequences, or harm emerges. ▪ Risks of integrated guideline and QA group ○ The authors identified potential risks for a joint guideline and QA group, including challenges with group processes, focusing on patient-important outcomes, unintended consequences, piloting of quality indicators, and achieving multi-stakeholder engagement. ○ It was concluded that these risks would be manageable and that the benefits of an integrated scheme outweighed the risks. ▪ Extension of guideline checklist to incorporate QA considerations ○ The authors added steps to the *Guidelines International Network-McMaster Guideline Development Checklist* to incorporate unique QA considerations, such as searching for QIs, setting QA priorities, and whether expert subgroups within an integrated process are required to address QA.
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? OR What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component: quality indicators.

	Category of evidence: Grade B.
Associated handbook(s)	N/R

Key: EtD – evidence to decision; GRADE – Grading of Recommendations, Assessment, Development, and Evaluations; N/A – not applicable; N/R – not reported; PICO – Population/patient, Intervention, Comparison, Outcomes; QA – quality assurance; QI – quality indicators.

Table C19 The GIN-McMaster guideline tool extension for the integration of quality improvement and quality assurance in guidelines: a description of the methods for its development

Publication identification	
Authors (year)	Piggott et al. (2023)
Country	International
DOI	https://doi.org/10.1016/j.jclinepi.2022.04.002
Publication description	
Design	Mixed methods
Objective	To develop an extension of the widely used GIN-McMaster Guideline Development Checklist and Tool for the integration of Quality Assurance and Improvement (QAI) schemes with guideline development.
Summary/Overview	This paper describes the process for the development of a 40-item checklist extension to the GIN-McMaster guideline development checklist to be considered for integrated QAI and guideline development. This checklist presents steps for the integration of QAI into guideline development across the existing 18 topics and a newly created topic specific to QAI.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	**GIN-McMaster Checklist Extension for Quality Assurance and Quality Improvement** Includes guideline development topics from the GIN-McMaster Checklist with additional items for extension to QAI where relevant: Organisation, budget, planning, and training 1. Define the care pathway related to the health topic, and identify the data and input parameters from different parts of it that implementers need to report on people-/patient-important outcomes and quality indicators. 2. Explore whether outsourcing of specific tasks [e.g., systematic review conduct or QA work] is required to conduct the work. 3. Determine the scope of the QAI scheme; in particular, clarify whether the group will be expected to specify performance indicators (thresholds, that is, measures of performance that are based on standards determined through evidence-based academic literature or through the consensus of experts when evidence is unavailable) linked to quality indicators (that is, measurable elements 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) and performance measures (that is, measurement of the degree to which healthcare organisations meet key goals). This may require specific expertise. Priority setting 4. Modification to original item 4: Identify the perspective that is taken for the QAI scheme (clinical individual, clinical health system, coverage decisions, public health, or population/societal). This may be the same as or different from the perspective taken for the guideline. 5. Search for quality indicators and performance measures already existing for the topic. 6. Consider where quality indicators should be assessed in relation to the evidence pertinent to the decision-making process: parallel groups assessing the guideline recommendations and QAI scheme, integrated with the recommendations, or sequentially (if sequentially, the QAI scheme would usually follow completion of the guideline). 7. Identify current use and gaps in accreditation/certification schemes on the QAI topic and any existing evidence on the use of these schemes to improve outcomes. Consider which quality indicators may be feasible for use in

	certification and/or accreditation schemes (e.g., those that are easy to measure and collect, already available, or ready to benchmark). This may affect acceptability and feasibility of any quality indicators emerging from the guideline/QA scheme. Guideline group membership 8. Determine whether project subgroups are required for focused work on specific topics, including the QA scheme development. Determine which individuals will be needed in subgroups and how they will interact with the larger panel. Establishing guideline group processes 9. Predefine the process that will be used by the group to select final quality indicators; performance measures; and, if applicable, performance indicators from a candidate list of guideline recommendations and, if appropriate, from existing quality indicators and performance measures (e.g., consensus, nominal group process). Identifying target audience and topic selection 10. Decide whether the evaluation of proposed existing quality indicators currently in use might be an intervention question with evidence reviewed by the guideline or conducted following publication of the guideline. Consumer and stakeholder involvement 11. Modification to original item 1: Identify the appropriate stakeholders to involve and/or consult in the development of the guideline and QA scheme to incorporate views of all who might be affected by them (e.g., professional groups, health managers, policymakers, industry representatives). 12. Modification to original item 2: Identify the appropriate consumers to involve in the group and/or consult in the development of the guideline and QA scheme to incorporate views of all who might be affected by them (e.g., individual people/patients, people who provide non-reimbursed care and support [such as family/caregivers] to patients, members of the public as potential patients and as funders of health care through taxation, QA experts and implementers, community organisations that represent the interests of patients, and advocates representing the interests of patients and people who care for them). Conflict of interest considerations 13. Apply the same declaration of interest and conflict of interest management rules to all participants and their institutions involved in the guideline and QA. 14. Consider potential conflicts of interest of the institution of participants in the process (e.g., if an institution is doing very well on an indicator and advocates for it to be included). PICO question generation 15. Visually describe where PICO questions being addressed and possible quality indicators sit using a logic model/analytical pathway/disease model/analytical PICO framework, including steps from prevention to diagnosis to treatment to outcomes. 16. Determine whether outcomes identified in the PICO question are appropriate as quality indicators for QA purposes. If they are not appropriate, consider whether surrogate outcomes are necessary for PICO question outcomes to be included as quality indicators (see item 27 below). 17. If a guideline question and the related recommendation are focused on whether to recommend the use of a quality indicator, the group should consider using an intervention EtD framework to assess the quality indicator.

	18. Identify all relevant input parameters from the different parts of the pathway to the overall people-/patient-important outcomes and quality indicators. Considering importance of outcomes and interventions, values, preferences, and utilities 19. Rate or select a small and well-defined but sufficient number of candidate quality indicators. The number should be linked to the number and breadth of recommendations in a guideline. The feasibility of implementation and monitoring should be considered. We propose no more than 10 quality indicators as reasonable for a moderate-sized guideline/QA scheme. 20. Consider the relation of candidate quality indicators to the people-/patient-important outcomes. 21. When considering candidate outcomes as quality indicators, consider the perspective of key stakeholders (people/patients, clinicians, decision makers) on the appropriateness of quality indicators (see item 27 below). Deciding what evidence to include and searching for evidence 22. Evaluate whether evidence supports that the use of a quality indicator improves people-/patient-important outcomes, particularly if the performance measure selected relates to a process outcome. 23. Consider which outcomes are measurable, feasible, scientifically sound, and relevant as quality indicators or performance measures. Summarising evidence and considering additional information None Judging quality, strength, or certainty of a body of evidence None Developing recommendations and determining their strength None Wording of recommendations and of considerations of implementation, research, monitoring, and evaluation (including considerations about QA) 24. Select quality indicators and performance measures based on the prioritisation of people-/patient-important outcomes. 25. Consider which outcomes, linked to recommendations, are appropriate as quality indicators (see item 27 below). 26. Identify, assess, and mitigate unintended consequences that a quality indicator may have on the target population (people/patients, clinicians, or decision makers). Reporting and peer review 27. Develop or adopt a standardised format for reporting the guideline and QA framework, with specific structure, headings, and content. 28. Report in the monitoring and evaluation section of the EtD framework how the quality indicator fits in relation to item 27 considerations of an appropriate quality indicator. Dissemination and implementation (including considerations about QA) 29. Refer quality indicators that were flagged for potential use for certification and/or accreditation schemes to the appropriate individuals/organisations for implementation. 30. Report in the monitoring and evaluation section of the EtD framework the appropriateness of the quality indicators and performance measures (see item 27 above).

	31. Propose mechanisms to document quality indicators or performance measures in a standardised (even anonymous) fashion to allow synthesis of data, collaboration, and shared learning across different healthcare systems and jurisdictions. This might provide a feedback mechanism for larger-scale improvement and updating of guidelines. Evaluation and use 32. Consider providing guidance on when to keep using, modify, or retire/cease using a performance measure. 33. Consider pilot testing the quality indicators and performance measures with the target end-users (e.g., with members of the target audience and stakeholders who participated in the development group). The type of pilot testing may be different for different groups depending on the timeline and feasibility of an integrated guideline and QA scheme; however, this is a critical step for ensuring feasibility and implementation. 34. Provide clarity on accountability (that is, criteria, support and tools) to making the changes in quality indicators, if applicable to the guideline/QA process. Updating 35. Periodically re-evaluate the quality indicators, performance measures, and performance indicators in a pre-specified time frame. The time frame may vary, but we suggest no longer than 2 to 3 years following release. Preparation for QA and selection of quality indicators (QAI-specific topic) 36. Consider the appropriateness of outcomes as quality indicators. Appropriate quality indicators should have high certainty, also known as scientific soundness (supporting evidence is at low risk of bias, precise, direct, relevant, consistent, and without publication bias); should be responsive or sensitive to change (may also be considered under risk of bias); and should be feasible to measure, implement and monitor. 37. Use a predefined process to select from the list of candidate quality indicators. 38. Consider potential performance measures linked to selected quality indicators, including their specification and intended use. 39. Use a predefined process to select performance measures linked to each quality indicator. 40. If applicable to the scope of the QA scheme, select and specify performance indicators (thresholds) for each performance measure.
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR**	N/A

What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core components: quality indicators; clarity of presentation. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: EtD – evidence to decision; GIN - Guidelines International Network; N/A - not applicable; N/R – not reported; PICO – Population/patient, Intervention, Comparison, Outcome; QA – quality assurance; QAI – quality assurance and improvement.

Table C20 Conceptual considerations on the integration of quality indicators into clinical pathways

Publication identification	
Authors (year)	Richter et al. (2016)
Country	Germany
DOI	https://doi.org/10.3233/978-1-61499-678-1-38
Publication description	
Design	Narrative review
Objective	To analyse the potential of the integration and utilisation of QIs in clinical pathways.
Summary/Overview	This study describes the potential of the integration and utilisation of QIs in clinical pathways. It comprises the first steps of a design-orientated research process, i.e., problem motivation and objective definition. A conceptual framework for pathway-integrated quality indicators is proposed. Potentials of the approach are outlined in three use case scenarios. The analysis highlights the potential for quality management on institutional and network level (that is, coordinated care across healthcare disciplines and institutions) and for bridging the gap between medical research and practice.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	**QIs for clinical pathway** ▪ Clinical pathways detail and structure the significant steps in the care process for patients with specific health conditions. They are multidisciplinary care plans that are adapted to local structures, resources and conditions and that focus on quality and efficiency of the care process with regard to the recommendations of clinical practice guidelines. However, quality is not always systematically monitored in clinical pathways. ▪ Enhancing integrated care pathways with QIs allows an easier monitoring and assessment of the care provided and the progress of a patient across care within/between healthcare institutions/settings. Therefore, it is necessary to identify network quality indicators and how they can be represented in clinical/ integrated care pathways. ▪ The translation of clinical practice guidelines into clinical pathways has been addressed by use of disease-specific indicators with methods for the vertical integration of clinical practice guidelines and clinical pathway models as well as with domain-specific modelling languages that integrate clinical practice guidelines into clinical pathways. However, these have not as yet detailed QIs. Adapting such approaches for the purpose of quality management could help to detect prevalent systematic deviations from quality goals defined in clinical practice guideline recommendations, which in turn could reveal potentials for improvements in the guideline development process or for revision of recommendations. Thus, guideline monitoring in practice could be facilitated, which is assumed to be a key element to improving guideline implementability and usage. Furthermore, results of evaluating QIs on the basis of process instances and routine data could be used to generate knowledge and hypotheses for future medical research activities. **Framework for sources, integration and utilisation of quality indicators in clinical pathways** ▪ QI sources/quality scheme o Definition and selection of QIs: The definition and selection of an adequate set of QIs for a healthcare institution depend on the intended usage (e.g., certification, internal objectives) and are the first necessary steps for pathway-based quality management.

	○ Integration of QIs into clinical pathways: If the sources suggest revised or new QIs, the implemented QI scheme should be adapted correspondingly. This makes it necessary to preserve the link between QI sources and the QIs implemented in clinical pathways. The specified QI scheme then needs to be integrated into the clinical pathways of a healthcare institution, either during or after development. ▪ Management and operational level ○ Utilisation: documenting, monitoring, controlling and evaluating QIs ▪ Data sources (to retrieve and develop QIs) ○ Process (pathway) instances (e.g., time, amount, resource input) ○ Routine data ○ Documented QIs ○ Continuous monitoring and formative evaluation. ○ Ex-post evaluation and further use ▪ Internal and external quality reports ▪ Budget planning ▪ Benchmarking ▪ Care network quality assessment ▪ Feedback to guideline development group ▪ Evidence for research.
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? OR What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	RQ1: Additional core component: quality indicators.

	Category of evidence: Grade D.
Associated handbook(s)	N/R

Key: N/A – not applicable; N/R – not reported; QI – quality indicator.

RQ2

Quality measures and or criteria to examine methodological robustness of clinical practice guidance development identified in peer-reviewed articles (not evaluated)

Table C21 A Reporting Tool for Practice Guidelines in Health Care: The RIGHT Statement

Publication identification	
Authors (year)	Chen et al. (2017)
Country	International
DOI	https://doi-org/10.7326/M16-1565
Publication description	
Design	Delphi study
Objective	To develop a tool—the RIGHT checklist—focusing on the essential items for reporting guidelines.
Summary/Overview	The 22-item RIGHT checklist may assist guideline developers in reporting guidelines, support journal editors and peer reviewers when considering guideline reports, and help healthcare practitioners understand and implement a guideline. The checklist may be useful for clinical practice guidelines and persons in public health and other healthcare fields. It provides users and evaluators with a clear, explicit description of the processes and procedures used to develop a guideline and access to the evidence used to formulate each recommendation.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	The RIGHT checklist.
What criteria does the tool use to assess quality?	RIGHT checklist items: ▪ Title/subtitle o Identify the report as a guideline, that is, with “guideline(s)” or “recommendation(s)” in the title o Describe the year of publication of the guideline o Describe the focus of the guideline, such as screening, diagnosis, treatment, management, prevention or other. ▪ Executive summary o Provide a summary of the recommendations contained in the guideline. ▪ Abbreviations and acronyms o Define new or key terms and provide a list of abbreviations and acronyms if required. ▪ Corresponding developer o Identify at least one corresponding developer or author who can be contacted about the guideline. ▪ Brief description of the health problem(s) o Describe the basic epidemiology of the problem, such as the prevalence/incidence, morbidity and mortality, and burden (including financial) resulting from the problem. ▪ Aim(s) of the guideline and specific objectives

	○ Describe the aims of the guideline and specific objectives, such as improvements in health indicators (e.g., mortality and disease prevalence), quality of life, or cost savings. ▪ Target population(s) ○ Describe the primary population(s) that is addressed by the recommendation(s) in the guideline ○ Describe any subgroups that are given special consideration in the guideline. ▪ End-users and settings ○ Describe the intended primary users of the guideline (such as primary care providers, clinical specialists, public health practitioners, programme managers, and policy-makers) or other potential users of the guideline ○ Describe the setting(s) in which the guideline is intended to be primarily used, such as general practice (primary care), low- and middle-income countries, inpatient facilities. ▪ Guideline development groups ○ Describe how all contributors to guideline development were selected and their roles and responsibilities (e.g., steering group, guideline panel, external reviewers, systematic review team, and methodologists) ○ List all individuals involved in developing the guideline, including their title, roles and institutional affiliation(s). ▪ Healthcare questions ○ State the key questions that were the basis for the systematic reviews in PICO or other formats as appropriate ○ Indicate how the outcomes were selected and sorted. ▪ Systematic reviews ○ Indicate whether the guideline is based on systematic reviews, and if so, whether these were conducted specifically for this guideline or whether existing systematic reviews were used or updated ○ If the guideline developers used existing systematic reviews, reference these and describe how those reviews were identified and assessed (provide the search strategies and the selection criteria, and describe how risk of bias was evaluated) and whether they were updated. ▪ Assessment of the certainty of the body of evidence ○ Describe the approach used to assess the certainty of the body of evidence. ▪ Recommendations ○ Provide clear, precise, actionable recommendations ○ Present separate recommendations for important subgroups if the evidence suggests that there are important differences in factors influencing recommendations, particularly the balance of health benefits and harms across subgroups ○ Indicate the strength of recommendations and the certainty of evidence. ▪ Rationale/explanation for recommendations ○ Describe whether values and preferences of the target population(s) were considered in the formulation of each recommendation. If yes, describe the approaches and methods used to elicit or identify these values and preferences. If values and preferences were not considered, provide an explanation

	○ Describe whether cost and resource implications were considered in the formulation of recommendations. If yes, describe the specific approaches and methods used (such as cost-effectiveness analysis) and summarise the results. If resource issues were not considered, provide an explanation ○ Describe other factors taken into consideration when formulating the recommendations, such as equity, feasibility and acceptability. ▪ Evidence to decision processes ○ Describe processes and approaches used by the guideline development group to make decisions, particularly the formulation of recommendations (such as how consensus was defined and achieved and whether voting was used). ▪ External review ○ Indicate whether the draft guideline underwent independent review, and if so, how this was executed and the comments considered and addressed. ▪ Quality assurance ○ Indicate whether the guideline was subject to a quality assurance process. If yes, describe the process. ▪ Funding source(s) and role(s) of the funder ○ Describe the specific sources of funding for all stages of guideline development ○ Describe the role of funder(s) in the different stages of the guideline development and in the dissemination and implementation of the recommendations. ▪ Declaration and management of interest ○ Describe what types of conflicts (financial and non-financial) were relevant to guideline development ○ Describe how conflicts of interest were evaluated and managed, and how users of the guideline can access the declarations. ▪ Access ○ Describe where the guideline, its appendices, and other related documents can be accessed. ▪ Suggestions for further research ○ Describe the gaps in the evidence and/or provide suggestions for future research. ▪ Limitations of the guideline. ○ Describe any limitations in the guideline development process (such as the development groups were not multidisciplinary or patients' values and preferences were not sought), and indicate how these limitations might have affected the validity of the recommendations.
What are the strengths and limitations of the tool?	N/R
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR**	N/A

What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	Category of evidence: Grade D.
Associated handbook(s)	N/R

Key: N/A – not applicable; N/R – not reported; RIGHT – Reporting Items for practice Guidelines in HealTHcare.

Table C22 The International Guideline Evaluation Screening Tool (IGEST): development and validation

Publication identification	
Authors (year)	D'Angelo et al. (2022)
Country	Italy
DOI	https://doi.org/10.1186/s12874-022-01618-5
Publication description	
Design	Mixed methods: Development and validation study
Objective	To describe the development and validation of a new tool to screen trustworthy clinical practice guidelines for their adoption/adaption: IGEST.
Summary/Overview	This study describes the development and validation of the IGEST screening tool for clinical practice guidelines. This tool is intended to be used by groups interested in adopting/adapting existing clinical practice guidelines to their own context to facilitate the process of quickly assessing their quality and deciding whether to use them as source clinical practice guidelines. It is a generic tool that can be applied to guidelines of any speciality, population or healthcare setting.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	The IGEST instrument
What criteria does the tool use to assess quality?	IGEST comprises four preliminary conditions and 12 criteria. The tool is intended to be used by groups in adopting/adapting existing clinical practice guidelines to their own context to facilitate the process of quick quality assessment and decision making on whether or not to use them as a source for clinical practice guidelines. The IGEST scoring system includes a yes/no answer option for the preliminary conditions and a four-point Likert scale for the 12 criteria. **Dimensions covered by IGEST** ▪ Preliminary condition a) The full disclosure of any financial COI for each decision voted by panellists is reported. b) The strategy for systematic review of the literature (i.e., search strategy and study selection) is clearly described. c) A full description of the affiliation and professional profile of panellists is reported. d) The external review carried out by independent experts is reported. *If all preliminary conditions are fulfilled, proceed; otherwise, reject.* ▪ Dimension 1: Conflict of interest 1. The guideline should describe how any identified conflicts were recorded and resolved. 2. Non-financial conflict of interest is managed. 3. COI of any guideline development group members are examined and managed by an oversight committee.

	4. Chair and co-chair are not allowed to have any relevant financial COI. *Rating scale:* Poor: no criteria met. Fair: only criterion 1 met. Good: criterion 1 + any of the remaining 2-4 met. Excellent: all the criteria met. ▪ Dimension 2. Quality and consistency 5. Quality of evidence is rated according to study type, and there is no explicit link between quality of evidence and strength of recommendations. 6. Quality of evidence is rated according to study type, and there is an explicit link between the quality of evidence and strength of recommendations. 7. Quality of evidence is rated according to both study type and risk of bias, and there is an explicit link between the quality of evidence and strength of recommendations. 8. Rating quality of evidence and grading strength of recommendations are based on GRADE or GRADE-like method. *Rating scale:* Poor: criterion 5 met. Fair: criterion 6 met. Good: criterion 7 met. Excellent: criterion 8 met. ▪ Dimension 3: Panel composition 9. Only one clinical specialty is involved. 10. More than one clinical specialty is involved. 11. Different relevant clinical specialities, general practitioners, and other professional groups are involved. 12. Different relevant clinical specialities, general practitioners, other professional groups, and at least one patient representative are involved. *Rating scale:* Poor: only criterion 9 met. Fair: criterion 10 met. Good: criterion 11 met. Excellent: criterion 12 met.
What are the strengths and limitations of the tool?	Strengths ▪ Intended to be used to support adoption or adaptation of international clinical practice guidelines to the local context ▪ Validity of the IGEST has been tested ▪ Can be completed relatively quickly. Limitations

	▪ No formal evaluation of the appraisal tool ▪ A new systematic review was not performed so important instruments might have been missed ▪ The tool is still in development and some aspects have yet to be tested, including psychometric properties and reliability, training of personnel to use the tool and the exact number of appraisers needed for the most reliable assessment ▪ The focus group and validation process was conducted with participants from only Italian agencies and organisations; hence there is a lack of involvement of wider international groups.
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: COI – conflict of interest; GRADE – Grading of Recommendations, Assessment, Development, and Evaluation; IGEST – International Guideline Evaluation Screening Tool; N/A – not applicable; N/R – not reported.

Table C23 GIN-McMaster Guideline Development Checklist extension for rapid recommendations

Publication identification	
Authors (year)	Morgan et al. (2018)
Country	International
DOI	https://doi.org/10.1186/s12961-018-0330-0
Publication description	
Design	Mixed methods
Objective	To develop an extension of the GIN-McMaster Guideline Development Checklist for Rapid Guideline development.
Summary/Overview	This paper summarises the process for the expansion of the GIN-McMaster guideline development checklist (GDC) to the development of rapid guidelines, based on a systematic review of developed rapid guidelines and qualitative research.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	GIN-McMaster Guideline Development Checklist extension for rapid recommendations.
What criteria does the tool use to assess quality?	**GIN-McMaster Guideline Development Checklist extension for rapid recommendations:** Organisation, Budget, Planning and Training 1. Define the amount of time available for development of the RG and the elements from the GDC that should be followed. 2. Develop RG-related standard operating procedures; develop templates for RGs; identify peer reviewers early on; and plan panel meetings as early as possible. Priority-setting 3. Define the rationale motivating the RG (e.g. new evidence about efficacy/cost-effectiveness/safety, emergent/dangerous situations, etc.). 4. Address whether there is a need for temporary and/or emergency guidance. Guideline Group Membership 5. Involve relevant individuals in the guideline oversight committee. 6. Develop a database of topic-specific experts by area of expertise to consult when establishing the guideline oversight committee. Establishing Guideline Group Processes 7. When the timelines are short, greater emphasis should be placed on using virtual meetings (alone or along with face-to-face meetings). Identifying Target Audience and Topic Selection 8. Alert the target audience to the RG before release. Consumer and Stakeholder Involvement None Conflict of Interest Considerations

	9. RG guideline development panels may need a rapid process for implementing conflict of interest policies. PICO Question Generation 10. RGs should address a limited number of questions. Considering Importance of Outcomes and Interventions, Values, Preferences and Utilities 11. Outcome prioritisation process for each PICO should be brief. 12. Information on patients' values and preferences can be informed by multiple methods, such as qualitative literature or patient advocacy groups. Deciding what Evidence to Include and Searching for Evidence 13. Consider the resources (both time and financial) needed and available for when defining the process for conducting the systematic review. Scoping or rapid reviews may inform eligibility criteria and prioritisation. Summarising Evidence and Considering Additional Information 14. Relevant primary studies and evidence solicited from experts may be used to inform 'additional information' in the evidence to decision table. Judging Quality, Strength or Certainty of a Body of Evidence None Developing Recommendations and Determining their Strength 15. Use pre-meeting voting and virtual meetings to expedite the decision-making process. Wording of Recommendations and of Considerations of Implementation, Feasibility and Equity 16. Finalise the wording of the final recommendations during the panel meeting(s). Reporting and Peer Review 17. Define and transparently record the process used when evidence is determined to be limited. 18. Expedited options for internal and external review of the RG should be explored, and if deemed possible, the process should be outlined in the RG. Dissemination and Implementation 19. RG implementation strategy should reflect the scope of the PICO. 20. RGs should outline and address any potential obstacles to implementation. Evaluation and Use None Updating 21. When developing an interim guideline, the date for when the RG or full practice guideline will be conducted should be defined. If developing an RG, the date for when the full practice guideline will be conducted should be defined.
What are the strengths and limitations of the tool?	Strengths ▪ The tool is developed based on the findings from a systematic survey of guidelines and methods by prominent organisations that produce RG globally. ▪ The development was also supported by in-depth interviews with key RG developers at World Health Organization. Limitation ▪ The authors did not propose differential weighting for different elements of the checklist. Hence, it cannot be indicated which elements are more important.

RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: GDC – Guideline Development Checklist; GIN – Guideline International Network; N/A – not applicable; N/R – not reported; PICO – Population/patient, Intervention, Comparison, Outcomes; RG – rapid guideline.

Table C24 A reporting tool for adapted guidelines in health care: The RIGHT-Ad@pt Checklist

Publication identification	
Authors (year)	Song et al. (2022)
Country	International
DOI	https://doi.org/10.7326/M21-4352
Publication description	
Design	Mixed methods: Review of adapted guidelines, semi-structured interviews, Delphi consensus survey, external review, final assessment of adapted guidelines.
Objective	To develop an extension of the RIGHT statement for the reporting of adapted guidelines (including recommendations that have been adopted, adapted, or developed de novo): the RIGHT-Ad@pt checklist.
Summary/Overview	The authors developed an extension of the RIGHT statement for the reporting of adapted guidelines—the RIGHT-Ad@pt checklist (comprising 7 sections, 27 topics and 34 items)—through an exhaustive process that included a literature review as well as input and consensus from a full range of relevant stakeholders, including guideline adaptation experts. Validity testing, but no formal evaluation, was conducted.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	The RIGHT-Ad@pt checklist
What criteria does the tool use to assess quality?	The RIGHT-Ad@pt checklist comprises 7 sections, 27 topics and 34 items: ▪ Basic information ○ Title/subtitle 1. Identify the report as an adaptation of practice guideline(s), that is include “guideline adaptation”, “adapting”, “adapted guideline/recommendation(s)”, or similar terminology in the title/subtitle 2. Describe the topic/focus/scope of the adapted guideline. ○ Cover/first page 3. Report the respective dates of publication and the literature search of the adapted guideline 4. Describe the developer and country/region of the adapted guideline. ○ Executive summary/abstract 5. Provide a summary of the recommendations contained in the adapted guideline. ○ Abbreviations and acronyms 6. Define key terms and provide a list of abbreviations and acronyms (if applicable). ○ Contact information of the guideline adaptation group 7. Report the contact information of the developer of the adapted guideline. ▪ Scope

	○ Source guideline(s) 8. Report the name and year of publication of the source guideline(s), provide the citation(s) and whether source authors were contacted. ○ Brief description of health problem(s) 9. Provide the basic epidemiological information about the problem (including the associated burden), health systems relevant issues, and note any relevant differences compared to the source guidelines. ○ Aim(s) and specific objectives 10. Describe the aim(s) of the adapted guideline and specific objectives, and note any relevant differences compared to the source guideline(s). ○ Target population(s) 11. Describe the target population(s) and subgroup(s) (if applicable) to which the recommendation(s) is addressed in the adapted guideline, and note any relevant differences compared to the source guideline(s). ○ End-users and settings 12. Describe the intended target users of the adapted guideline, and note any relevant differences compared to the source guideline(s). 13. Describe the setting(s) for which the adapted guideline is intended, and note any relevant differences compared to the source guideline(s). ▪ Rigor of development ○ Guideline adaptation group 14. List all contributors to the guideline adaptation process and describe their selection process and responsibilities. ○ Adaptation framework/methodology 15. Report which framework or methodology was used in the guideline adaptation process. ○ Source guideline(s) 16. Describe how the specific source guideline(s) was (were) selected. ○ Key questions 17. State the key questions of the adapted guideline using a structured format, such as PICO, or another format as appropriate 18. Describe how the key questions were developed/modified, and/or prioritised. ○ Source recommendation(s) 19. Describe how the recommendation(s) from the source guideline(s) was (were) assessed with respect to the evidence considered for the different criteria, the judgements and considerations made by the original panel. ○ Evidence synthesis 20. Indicate whether the adapted recommendation(s) is/are based on existing evidence from the source guideline(s), and/or additional evidence 21. If new research evidence was used, describe how it was identified and assessed. ○ Assessment of the certainty of the body of evidence and strength of recommendation

	22. Describe the approach used to assess the certainty/quality of the body/ies of evidence and the strength of recommendations in the adapted guideline and note any differences (if applicable) compared to the source guideline(s). ○ Decision-making processes 23. Describe the processes used by the guideline adaptation group to make decisions, particularly the formulation of recommendations. ▪ Recommendations ○ Recommendations 24. Report recommendations and indicate whether they were adapted, adopted or de novo 25. Indicate the direction and strength of the recommendations and the certainty/quality of the supporting evidence and note any differences compared to the source recommendations (if applicable). 26. Present separate recommendations for important subgroups if the evidence suggests important differences in factors influencing recommendations and note any differences compared to the source recommendations (if applicable). ○ Rationale/explanation for recommendations 27. Describe the criteria/factors that were considered to formulate the recommendations or note any relevant differences compared to the source guideline(s) (if applicable). ▪ External review and quality assurance ○ External review 28. Indicate whether the adapted guideline underwent an independent external review. If yes, describe the process. ○ Organisational approval 29. Indicate whether the adapted guideline obtained organisational approval. If yes, describe the process. ▪ Funding, declaration and management of interest ○ Funding source(s) and funder role(s) 30. Report all sources of funding for the adapted guideline and source guideline(s), and the role of the funders. ○ Declaration and management of interests 31. Report all conflicts of interest of the adapted and the source guideline(s) panels, and how they were evaluated and managed. ▪ Other information ○ Implementation 32. Describe the potential barriers and strategies for implementing the recommendations (if applicable). ○ Update 33. Briefly describe the strategy for updating the adapted guideline (if applicable). ○ Limitations and suggestions for further research

	34. Describe the challenges of the adaptation process, the limitations of the evidence, and provide suggestions for future research.
What are the strengths and limitations of the tool?	Strengths ▪ It retains the strength of the RIGHT statement while contextualising the guideline adaptation process. ▪ The tool was tested for usability with external reviewers and two assessments with published adapted guidelines. Limitation ▪ Further validation for real-life use is required.
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? OR What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: N/A – not applicable; N/R – not reported; PICO – Population/patient, Intervention, Comparison and Outcome; RIGHT – Reporting Items for practice Guidelines in HealTHcare.

Table C25 An extension of the RIGHT statement for introductions and interpretations of clinical practice guidelines: RIGHT for INT

Publication identification	
Authors (year)	Zhou et al. (2022)
Country	China
DOI	https://doi.org/10.1111/jebm.12466
Publication description	
Design	Mixed methods: Systematic review of published articles relating to guideline interpretation; Delphi study (following the methods recommended by the EQUATOR network and the RIGHT Statement).
Objective	To describe the development process of the extension of the RIGHT Statement for INTroductions and INTerpretations of Clinical Practice Guidelines (RIGHT for INT) checklist.
Summary/Overview	The RIGHT for INT checklist was developed following the methods recommended by the EQUATOR Network and formed by scientific consensus. The checklist contains 27 items with examples and explanations, providing guidance for interpreting clinical practice guidelines.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	The RIGHT Statement for INTroductions and INTerpretations of Clinical Practice Guidelines (RIGHT for INT) checklist.
What criteria does the tool use to assess quality?	**The extension of the RIGHT Statement for INTroductions and INTerpretations of Clinical Practice Guidelines (RIGHT for INT) checklist:** <u>Title</u> ▪ Title ○ Identify the article as an "interpretation" and include the original guideline title. <u>Abstract</u> ▪ Basic information ○ Describe the publication source, developer, publication year, and version of the guideline. ▪ Content ○ Describe the scope of the interpretation. ▪ Significance ○ The significance of the guideline to clinical practice (international, national, or regional level) and development of guidelines and research. ▪ Limitations ○ Describe the limitations of the guideline to local clinical practice (e.g., development method, usage, etc.). <u>Background of guideline interpretation</u> ▪ Reasons for interpretation ○ Describe the reason for writing an interpretation of the guideline. ▪ Interpreter ○ Describe the clinical or methodological background of the interpreter.

	▪ Conflict of interest o Describe whether the interpreter has financial or nonfinancial conflicts of interest related to the guideline. <u>Background of guideline development</u> ▪ Version and update o State the version, publication year, and updated content of the guideline. ▪ Source o State the publication source or platform of the guideline. ▪ Developer o State the developer of the guideline. ▪ User and target population o Describe the target users and target population of the guideline. ▪ Setting o Describe the application setting of the guideline (such as general hospital, specialist hospital, or primary medical institution). ▪ Funding o Describe the sources and types of funding that the guideline received. ▪ Conflicts of interest o Describe the conflicts of interest of the guideline development group members and how they are managed. <u>Guideline development methodology</u> ▪ The composition of guideline group o Describe the details of the guideline working group (such as recruitment methods, and the distribution of fields of expertise and geographical region). ▪ Evidence o Describe the sources, types, and synthesis methods of the evidence. In particular, it is necessary to clarify whether the guideline is based on systematic reviews. ▪ Grading methods o Describe the approach for evaluating the quality of evidence and strength of recommendations used by the guideline and its interpretation. <u>Recommendations</u> ▪ Selection of recommendations o State whether all or part of the recommendations of the guideline are interpreted. If some recommendations are selected for interpretation, the reasons should be explained. ▪ Content of recommendations o Interpret each recommendation in its specific context considering e.g., the population, type of intervention, resource need, applicability, and feasibility to formulate an appropriate, clear and comprehensive explanation. ▪ Supplementary material

	○ Interpret the appendices (such as the search strategy, evidence summary table, or implementation tools) of the guideline, so that readers can understand and apply the guideline comprehensively. <u>Strengths and limitations</u> ▪ Strengths ○ Describe the strengths of the guideline in terms of methodology and other aspects. ▪ Limitations ○ Describe the limitations of the guideline in terms of methodology and other aspects. <u>Implications for local guidelines and clinical research</u> ▪ Implications for development or updating of guideline ○ Describe the significance of the guidelines for the development or updating of relevant guidelines in the local setting. ▪ Implications for clinical research ○ Describe the significance of the guideline for launching clinical research in the local setting. <u>Dissemination and implementation</u> ▪ Dissemination and implementation ○ Describe strategies for efficient dissemination and implementation of the guideline locally. <u>Reporting quality</u> ▪ Guideline reporting quality. ○ Use the RIGHT checklist to evaluate the guideline.
What are the strengths and limitations of the tool?	Strengths ▪ Developed using methods recommended by the EQUATOR Network and formed by scientific consensus ▪ First reporting standard for interpreting clinical practice guidelines and hence it fills a gap in the area of guideline interpretation. Limitations ▪ The reliability and validity have not been evaluated. ▪ As working group members were all from the same country, international perspectives were not represented.
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	

Reviewer notes	Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: EQUATOR – Enhancing the QUAlity and Transparency Of health Research; N/A – not applicable; N/R – not reported; RIGHT – Reporting Items for practice Guidelines in HealTHcare; RIGHT for INT – RIGHT Statement for INTroductions and INTerpretations of Clinical Practice Guidelines.

Quality measures and or criteria to examine methodological robustness of clinical practice guidance development identified in peer-reviewed articles (evaluated)

Table C26 Developing a Clinician Friendly Tool to Identify Useful Clinical Practice Guidelines: G-TRUST

Publication identification	
Authors (year)	Shaughnessy et al. (2017)
Country	Canada, USA
DOI	https://doi.org/10.1370/afm.2119
Publication description	
Design	Modified Delphi study
Objective	To develop the G-TRUST instrument for clinicians to easily identify useful clinical practice guidelines.
Summary/Overview	This study describes the modified Delphi process through which the authors developed an 8-item checklist designed to help clinicians quickly identify useful guidelines to follow in practice.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	The G-TRUST instrument.
What criteria does the tool use to assess quality?	**G-TRUST instrument items** ▪ Relevance and utility ○ The recommendations focus on improving patient orientated outcomes, explicitly comparing benefits versus harms to support clinical decision making. ○ The recommendations are clear and actionable. ○ The patient populations and conditions are relevant to the clinician's clinical setting. ▪ Trustworthiness ○ The guidelines are based on a systematic review of the research. ○ The recommendation statements important to the clinician are based on graded evidence and include a description of the quality (e.g., strong, weak) of the evidence. ○ The guideline development includes a research analyst, such as a statistician or epidemiologist. ▪ Interpretation ○ The Chair of the guideline development committee and a majority of the rest of the committee are free of declared financial conflicts of interest, and the guideline development group did not receive industry funding for developing the guideline. ○ The guideline development includes members from the most relevant specialties and includes other key stakeholders, such as patients, payer organisations, and public health entities, when applicable.
What are the strengths and limitations of the tool?	▪ Strengths ○ The authors suggest that G-TRUST is more stringent than AGREE II in that it stipulates an independent (i.e., non-conflicted) research analyst or methodologist be part of the process

	○ The tool is also more stringent than AGREE II in its handling of conflicts of interest (barring them rather than simply addressing them) and in broad representation on the guideline development group ○ According to the authors, a major advantage of the G-TRUST is that it assigns different levels of importance to individual items (e.g., major, minor) and arrives at a determination of overall guideline quality (useful, may not be useful, not useful). ▪ Limitations ○ No summative evaluation of the tool included in the article ○ Use of a conservative cut-off score, while preventing false positives (e.g., falsely identifying guidelines as high quality), will exclude some high-quality guidelines ○ It may be difficult for users to determine conflicts of interest and the presence of a research analyst on the guideline development group.
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	Category of evidence: Grade D.
Associated handbook(s)	N/R

Key: AGREE – Appraisal of Guideline REsearch and Evaluation; G-TRUST – Guideline Trustworthiness, Relevance, and Utility Scoring Tool; N/A – not applicable; N/R – not reported.

Table C27 Developing and testing the agency for healthcare research and quality's National Guideline Clearinghouse Extent of Adherence to Trustworthy Standards (NEATS) instrument

Publication identification	
Authors (year)	Jue et al. (2019)
Country	USA
DOI	https://doi.org/10.7326/M18-2950
Publication description	
Design	Mixed methods
Objective	To summarise the development of the NEATS instrument.
Summary/Overview	The NEATS instrument is a 15-item tool for guideline assessment, covering 8 domains: disclosure of the funding source, disclosure and management of financial conflicts of interest, GDG composition, use of a systematic review of evidence, evidence foundations for rating the strength of recommendations, specific and unambiguous articulation of recommendations, external review, and updating.
RQ1: Description of core components of clinical practice guidance	
What core components elements have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	The NEATS Instrument.
What criteria does the tool use to assess quality?	**NEATS criteria** 1. The clinical practice guideline discloses and states explicitly its funding source. 2. Financial conflict of interest of GDG members have been disclosed and managed. 3. GDG composition a. The GDG includes persons from a variety of relevant clinical specialties and other professional groups. b. The guideline states that it included a methodological expert in the GDG, and it identifies the methodologist. 4. The GDG sought the views, perspectives, and preferences of patients, patient surrogates (parents and caretakers), patient advocates, or the public, intended to represent those who have experience with the disease, its treatments, or its complications, or those who could be affected by the guideline. 5. Use of a systematic review of evidence a. The clinical practice guideline or a related companion document describes a search strategy that includes a listing of database(s) searched; a summary of search terms used; and the specific time period covered by the literature search, including the beginning date (month and year) and end date (month and year). b. The clinical practice guideline or a related companion document describes the study selection; description includes the number of studies identified, the number of studies included, and a summary of inclusion and exclusion criteria. c. The clinical practice guideline or a related companion document provides a synthesis of evidence from the selected studies; i.e., an analysis of individual studies and the body of evidence, in the form of a detailed description or evidence tables, or both.

	6. The clinical practice guideline provides a grading or rating of the level of confidence in or certainty regarding the quality or strength of the evidence for each recommendation. 7. The potential benefits and harms of recommended care are clearly described. 8. A summary of the relevant supporting evidence is explicitly linked to the recommendations. 9. The clinical practice guideline gives a rating of the strength of each recommendation that takes into account benefits and harms, available evidence, and the confidence in the underlying evidence. 10. The recommendations are specific and unambiguous, stating what action should or should not be taken in what situations and for what population groups. Where the clinical practice guideline recommendations are intentionally vague or underspecified, the clinical practice guideline clearly describes the rationale behind those recommendations. 11. The guideline has been reviewed by relevant stakeholders, including scientific and clinical experts, organisations, agencies, and patients. 12. The clinical practice guideline describes a procedure to update the guideline.
What are the strengths and limitations of the tool?	▪ Strengths ○ The tool has both good external validity and good interrater reliability across trained reviewers. ○ The NEATS instrument offers consolidated information on transparency, completeness of documentation, and rigor of development to help inform a user's judgement about a guideline's suitability for use. ○ Publicly posting clinical practice guideline assessments using the NEATS instrument gives clinicians, educators, policymakers, payers, and others insight into the degree of trustworthiness of clinical practice guidelines. ○ It is developed through a federally funded contract and hence is in the public domain. ▪ Limitations ○ No summative evaluation of the appraisal tool. ○ In assessing the tool's external validity, the authors were limited by time, budget, and the constraints of the Paperwork Reduction Act to surveying only nine persons who were not federal employees (one represented a federal clinical practice guideline developer). ○ NEATS was created for use by NGC staff and was neither designed not tested for wider application, hence its generalisability outside NGC maybe limited. However, with use of similarly trained staff and replication of a validation process, it may be suitable for use outside its original context.
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A

How is the innovation used in practice?	N/A
Notes	
Reviewer notes	Category of evidence: Grade C.
Associated handbook(s)	Agency For Healthcare Research And Quality.

Key: GDG – guideline development group; N/A – not applicable; NEATS – National Guideline Clearinghouse Extent of Adherence to Trustworthy Standards; NGC – National Guideline Clearinghouse.

Table C28 Assessing the process and outcome of the development of practice guidelines and recommendations: PANELVIEW instrument development

Publication identification	
Authors (year)	Wiercioch et al. (2020)
Country	International
DOI	https://doi.org/10.1503/cmaj.200193
Publication description	
Design	Mixed methods: Scale development methodology
Objective	To develop and validate a tool for assessing guideline panel members' perception of the appropriateness of, and satisfaction with, the process, methods and outcome of the development of a health guideline.
Summary/Overview	The PANELVIEW tool comprises 34 domains and enables guideline organisations to involve clinicians, patients and other participants in evaluating their guideline processes.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	PANELVIEW Questionnaire: Assessing the process and outcome of the development of practice guidelines and recommendations.
What criteria does the tool use to assess quality?	**PANELVIEW Questionnaire items** <u>Administration</u> ▪ The logistical support provided for organisation of the guideline project and panel meeting was appropriate (e.g., scheduling of the meeting, sharing of materials, venue/location) ▪ There was adequate preparatory work and meetings/teleconferences before the final panel meeting ▪ Adequate time was given for guideline group members to complete tasks (e.g., surveys, providing feedback) throughout the development of the guideline and to review the evidence summary and other material before the panel meeting ▪ Adequate time was allotted for the final panel meeting for all guideline questions to be discussed and recommendations to be formulated ▪ The panel meeting had a clearly defined agenda and objectives. <u>Training</u> ▪ Information was provided about the specific methodology and frameworks to ensure understanding of the overall process and steps that would be used to develop the guideline. <u>Panel chair</u> ▪ The panel chair(s) was able to provide clinical and methodological guidance during the meeting, providing direction and support for decision-making ▪ The panel chair(s) was able to manage the group process, establishing an atmosphere of support that ensured involvement of all panel members in the discussion and free expression of opinions. <u>Conflict of interest</u> ▪ There was appropriate management of potential interests (financial, academic) of guideline group members, of the organisation and in the evidence synthesis being free from bias

	▪ There was appropriate management of potential bias in panel members' interpretation of evidence and alignment with prior beliefs. Scoping the guideline ▪ The panel was given sufficient opportunity to be involved in the prioritisation of questions and scoping of the guideline ▪ The final scope of the guideline was clearly communicated to the guideline group and agreement was sought. Methodology and process ▪ The evidence synthesis was rigorous ▪ A transparent and usable summary of the evidence was made available for the discussion. Considering the evidence and contributing through expertise ▪ Appropriate consideration was given to the evidence, including all relevant types, and balanced with panel members' input and opportunity to use their experience to interpret the evidence ▪ The method or process used for decision-making with the available evidence was appropriate ▪ There was appropriate involvement and consultation with key stakeholders during the guideline development ▪ Appropriate consideration was given to patients' views, perspectives, values and preferences. Formulating the recommendations ▪ An appropriate method was used for formulating the recommendations with transparency of judgements made ▪ Appropriate consideration was given to relevant external factors (e.g., policy implications, setting-specific healthcare factors, acceptability of recommendations) in formulating the guideline recommendations ▪ The consensus method used by the panel was appropriate, allowing ability to reach consensus ▪ The wording of the guideline recommendations formulated was clear and actionable ▪ There was transparency in going from the panel's recommendations to the final recommendations that appear in the guideline report, and notice was given about any changes made. Group composition ▪ There was diversity in membership and adequate representation of backgrounds, specialties and balance of expertise in the panel composition ▪ The panel size was appropriate. Group roles ▪ The required commitment was at an appropriate level for the guideline group members ▪ The contributions of the guideline group members were valued and appropriate credit was given. Group interaction ▪ There was mutual respect between guideline group members, with friendly and professional conduct. Implementation and dissemination plan ▪ Appropriate consideration was given to the discussion of research gaps and needs for future research ▪ Appropriate consideration was given for the planning of dissemination and implementation of the guideline. Writing guideline ▪ The writing of the guideline was well planned, with agreement on the format(s) and opportunity for panel members to provide input and review the guideline draft. Incentive ▪ I felt that my involvement in the guideline will have an impact on the health of people.

	Overall satisfaction ▪ Overall, I was satisfied with the guideline development process ▪ I would participate in this guideline development process again.
What are the strengths and limitations of the tool?	Strengths ▪ Existing instruments for assessing guideline credibility rely on the guideline authors' report, which may describe the process as planned but not as implemented or as viewed by all group members, and may not reflect all relevant nuances of the process that affect the trustworthiness of recommendations. The PANELVIEW tool focuses on these important nuances and on the transparency of the guideline development process, allowing organisations responsible for guideline development to inform their quality improvement efforts. ▪ Best practice for instrument development was followed, including reviewing the literature, contacting key informants at guideline organisations and surveying panellists about key factors affecting guideline development. ▪ The tool was tested successfully with panels from international guideline organisations. Limitations ▪ There was no summative evaluation of the appraisal tool ▪ A potential limitation is that the authors did not conduct systematic searches of the nonmedical literature in the areas of business, education and policy-making for relevant items. At each step involving key informants, convenience sampling was used, which may introduce sampling bias. To address this, the authors drew on a broad representation of working guideline panellists, with varying levels of experience, as well as guideline development experts from organisations representing a wide range of processes and methods ▪ The eight guideline groups involved in field testing the PANELVIEW tool were recruited through key informants, and, for some aspects of development, the groups used similar methods (e.g., using the GRADE approach for assessing quality of evidence and strength of recommendations) and involved experienced group chairs. Despite this, variability in scores within the groups was observed, which would allow guideline developers to identify whether individual panellists viewed the process and specific aspects of the process as more or less appropriate.
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? OR What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: GRADE – Grading of Recommendations, Assessment, Development, and Evaluation**;** N/A – not applicable; N/R – not reported.

Supplementary tools that support quality appraisal of clinical practice guidance (evaluated)

Table C29 The AGREE Reporting Checklist: a tool to improve reporting of clinical practice guidelines

Publication identification	
Authors (year)	Brouwers et al. (2016)
Country	Canada
DOI	https://doi.org/10.1136/bmj.i1152
Publication description	
Design	Mixed methods
Objective	To describe the development of the AGREE Reporting Checklist, which was designed to improve the quality of practice guideline reporting and aligns with AGREE II in its structure and content.
Summary/Overview	The checklist maintains AGREE II's structure of six quality domains and 23 key items, providing a systematic and logical process for reporting essential information. For each of the 23 items, a summary statement and a bulleted list of specific reporting criteria are provided.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	AGREE Reporting Checklist.
What criteria does the tool use to assess quality?	**AGREE Reporting Checklist items** **Domain 1: scope and purpose** 1. Objectives ○ Report the overall objective(s) of the guideline. The expected health benefits from the guideline are to be specific to the clinical problem or health topic. ▪ Health intent(s) (i.e., prevention, screening, diagnosis, treatment, etc.) ▪ Expected benefit(s) or outcome(s) ▪ Target(s) (e.g., patient population, society). 2. Questions ○ Report the health question(s) covered by the guideline, particularly for the key recommendations ▪ Target population ▪ Intervention(s) or exposure(s) ▪ Comparisons (if appropriate) ▪ Outcome(s) ▪ Health care setting or context. 3. Population ○ Describe the population (i.e., patients, public, etc.) to whom the guideline is meant to apply. ▪ Target population, sex and age ▪ Clinical condition (if relevant)

	▪ Severity/stage of disease (if relevant) ▪ Comorbidities (if relevant) ▪ Excluded populations (if relevant). **Domain 2: stakeholder involvement** 4. Group membership o Report all individuals who were involved in the development process. This may include members of the steering group, the research team involved in selecting and reviewing/rating the evidence and individuals involved in formulating the final recommendations. ▪ Name of participant ▪ Discipline/content expertise (e.g., neurosurgeon, methodologist) ▪ Institution (e.g., St. Peter's hospital) ▪ Geographical location (e.g., Seattle, Washington State) ▪ A description of the member's role in the guideline development group. 5. Target population preferences and views o Report how the views and preferences of the target population were sought/considered and what the resulting outcomes were. ▪ Statement of type of strategy used to capture patients'/publics' views and preferences (e.g., participation in the guideline development group, literature review of values and preferences) ▪ Methods by which preferences and views were sought (e.g., evidence from literature, surveys, focus groups) ▪ Outcomes/information gathered on patient/public information ▪ How the information gathered was used to inform the guideline development process and/or formation of the recommendations. 6. Target users o Report the target (or intended) users of the guideline. ▪ The intended guideline audience (e.g., specialists, family physicians, patients, clinical or institutional leaders/administrators) ▪ How the guideline may be used by its target audience (e.g., to inform clinical decisions, to inform policy, to inform standards of care). **Domain 3: rigour of development** 7. Search methods o Report details of the strategy used to search for evidence. ▪ Named electronic database(s) or evidence source(s) where the search was performed (e.g., MEDLINE, EMBASE, PsychINFO, CINAHL) ▪ Time periods searched (e.g., January 1, 2004 to March 31, 2008) ▪ Search terms used (e.g., text words, indexing terms, subheadings) ▪ Full search strategy included (e.g., possibly located in appendix). 8. Evidence selection criteria

	○ Report the criteria used to select (i.e., include and exclude) the evidence. Provide rationale, where appropriate. ▪ Target population (patient, public, etc.) characteristics ▪ Study design ▪ Comparisons (if relevant) ▪ Outcomes ▪ Language (if relevant) ▪ Context (if relevant). 9. Strengths and limitations of the evidence ○ Describe the strengths and limitations of the evidence. Consider from the perspective of the individual studies and the body of evidence aggregated across all the studies. Tools exist that can facilitate the reporting of this concept. ▪ Study design(s) included in body of evidence ▪ Study methodology limitations (sampling, blinding, allocation concealment, analytical methods) ▪ Appropriateness/relevance of primary and secondary outcomes considered ▪ Consistency of results across studies ▪ Direction of results across studies ▪ Magnitude of benefit versus magnitude of harm ▪ Applicability to practice context. 10. Formulation of recommendations ○ Describe the methods used to formulate the recommendations and how final decisions were reached. Specify any areas of disagreement and the methods used to resolve them. ▪ Recommendation development process (e.g., steps used in modified Delphi technique, voting procedures that were considered) ▪ Outcomes of the recommendation development process (e.g., extent to which consensus was reached using modified Delphi technique, outcome of voting procedures) ▪ How the process influenced the recommendations (e.g., results of Delphi technique influence final recommendation, alignment with recommendations and the final vote). 11. Consideration of benefits and harms ○ Report the health benefits, side effects, and risks that were considered when formulating the recommendations. ▪ Supporting data and report of benefits ▪ Supporting data and report of harms/side effects/risks ▪ Reporting of the balance/trade-off between benefits and harms/side effects/risks ▪ Recommendations reflect considerations of both benefits and harms/side effects/risks. 12. Link between recommendations and evidence ○ Describe the explicit link between the recommendations and the evidence on which they are based. ▪ How the guideline development group linked and used the evidence to inform recommendations

	▪ Link between each recommendation and key evidence (text description and/or reference list) ▪ Link between recommendations and evidence summaries and/or evidence tables in the results section of the guideline. 13. External Review o Report the methodology used to conduct the external review. ▪ Purpose and intent of the external review (e.g., to improve quality, gather feedback on draft recommendations, assess applicability and feasibility, disseminate evidence) ▪ Methods taken to undertake the external review (e.g., rating scale, open-ended questions) ▪ Description of the external reviewers (e.g., number, type of reviewers, affiliations) ▪ Outcomes/information gathered from the external review (e.g., summary of key findings) ▪ How the information gathered was used to inform the guideline development process and/or formation of the recommendations (e.g., guideline panel considered results of review in forming final recommendations). 14. Updating procedure o Describe the procedure for updating the guideline ▪ A statement that the guideline will be updated ▪ Explicit time interval or explicit criteria to guide decisions about when an update will occur ▪ Methodology for the updating procedure. **Domain 4: clarity of presentation** 15. Specific and unambiguous recommendations o Describe which options are appropriate in which situations and in which population groups, as informed by the body of evidence. ▪ A statement of the recommended action ▪ Intent or purpose of the recommended action (e.g., to improve quality of life, to decrease side effects) ▪ Relevant population (e.g., patients, public) ▪ Caveats or qualifying statements, if relevant (e.g., patients or conditions for whom the recommendations would not apply) ▪ If there is uncertainty about the best care option(s), the uncertainty should be stated in the guideline. 16. Management options o Describe the different options for managing the condition or health issue. ▪ Description of management options ▪ Population or clinical situation most appropriate to each option. 17. Identifiable key recommendations o Present the key recommendations so that they are easy to identify. ▪ Recommendations in a summarised box, typed in bold, underlined, or presented as flow charts or algorithms ▪ Specific recommendations grouped together in one section.

	Domain 5: applicability 18. Facilitators and barriers to application ○ Describe the facilitators and barriers to the guideline's application. ▪ Types of facilitators and barriers that were considered ▪ Methods by which information regarding the facilitators and barriers to implementing recommendations were sought (e.g., feedback from key stakeholders, pilot testing of guidelines before widespread implementation) ▪ Information/description of the types of facilitators and barriers that emerged from the inquiry (e.g., practitioners have the skills to deliver the recommended care, sufficient equipment is not available to ensure all eligible members of the population receive mammography) ▪ How the information influenced the guideline development process and/or formation of the recommendations. 19. Implementation advice/tools ○ Provide advice and/or tools on how the recommendations can be applied in practice. ▪ Additional materials to support the implementation of the guideline in practice. For example: • Guideline summary documents • Links to checklists, algorithms • Links to how-to manuals • Solutions linked to barrier analysis (see Item 18) • Tools to capitalise on guideline facilitators (see Item 18) • Outcome of pilot test and lessons learned. 20. Resource implications ○ Describe any potential resource implications of applying the recommendations. ▪ Types of cost information that were considered (e.g., economic evaluations, drug acquisition costs) ▪ Methods by which the cost information was sought (e.g., a health economist was part of the guideline development panel, use of health technology assessments for specific drugs, etc.) ▪ Information/description of the cost information that emerged from the inquiry (e.g., specific drug acquisition costs per treatment course) ▪ How the information gathered was used to inform the guideline development process and/or formation of the recommendations. 21. Monitoring/auditing criteria ○ Provide monitoring and/or auditing criteria to measure the application of guideline recommendations. ▪ Criteria to assess guideline implementation or adherence to recommendations ▪ Criteria for assessing impact of implementing the recommendations

	▪ Advice on the frequency and interval of measurement ▪ Operational definitions of how the criteria should be measured. **Domain 6: editorial independence** 22. Funding body o Report the funding body's influence on the content of the guideline. ▪ The name of the funding body or source of funding (or explicit statement of no funding) ▪ A statement that the funding body did not influence the content of the guideline. 23. Competing interests o Provide an explicit statement that all group members have declared whether they have any competing interests. ▪ Types of competing interests considered ▪ Methods by which potential competing interests were sought ▪ A description of the competing interests ▪ How the competing interests influenced the guideline process and development of recommendations.
What are the strengths and limitations of the tool?	**Strengths:** ▪ The structure and design of the checklist aligns with AGREE II. ▪ The checklist can be used prospectively in the drafting and editing stages as well as retrospectively as a quality assurance step. **Limitations:** ▪ Not reported.
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: AGREE – Appraisal of Guideline REsearch and Evaluation; CINAHL – Cumulated Index to Nursing and Allied Health Literature; EMBASE – Excerpta Medica database; N/A – not applicable; N/R – not reported.

Table C30 Development and validation of a tool to assess the quality of clinical practice guideline recommendations

Publication identification	
Authors (year)	Brouwers et al. (2020)
Country	International
DOI	https://doi.org/10.1001/jamanetworkopen.2020.5535
Publication description	
Design	Mixed methods: Tool development and descriptive cross-sectional validation study
Objective	To describe the development and validation of AGREE-REX, a tool designed to evaluate the quality of clinical practice guideline recommendations.
Summary/Overview	This study describes the development and validation of AGREE-REX, a tool designed to assess the quality of clinical practice guideline recommendations. The study involved 322 international stakeholders and resulted in the creation of an 11-item tool with three response scales and two overall items. The tool was found to be easy to understand and apply, and may have implications for the implementation of clinical practice guideline recommendations in healthcare settings.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	AGREE-REX, a tool to evaluate the quality of clinical practice guideline recommendations. It is a complement to the AGREE II tool.
What criteria does the tool use to assess quality?	To ensure that guideline recommendations are of high quality, the AGREE-REX addresses three factors that must be considered: ▪ Clinical credibility of the recommendations, based on the available evidence and its appropriateness for the target users, context, and patients/populations; ▪ Consideration of values of all relevant stakeholders in the formulation of the recommendations; ▪ Implementability of the recommendations. **AGREE-REX items** ▪ **Evidence** o The guideline assesses any risk of bias related to the study designs of the supporting evidence o The guideline describes the consistency of the results (i.e., similarity of results across studies) o The guideline addresses the directness of the evidence (i.e., addresses the exact interventions, populations, and outcomes of interest) to the clinical/health problem o The guideline indicates the precision of the results (e.g., width of confidence intervals of individual studies or meta-analyses) o The guideline describes the magnitude of the benefits and harms o The guideline assesses the likelihood of publication bias o The guideline addresses the possibility of confounding factors (if applicable) o The guideline indicates the dose-response gradient (if applicable).

	▪ **Applicability to target users** o The guideline addresses a clinical/health problem that is relevant to the intended target users o There is an alignment between: ▪ Target user's scope of practice and targeted patients/populations ▪ Target user's scope of practice and recommended actions ▪ The direction of the recommendations (i.e., in favour or against particular action) and the trade-offs between harms and benefits ▪ The definitiveness or strength of the recommendations and the trade-offs between harms and benefits. ▪ **Applicability to patients or populations** o The guideline includes outcomes that are relevant to the targeted patients/populations. These outcomes are often referred to as patient-important outcomes, patient-centred outcomes, patient-reported outcomes, or patient experience ▪ Relevant outcomes were considered in the development of the evidence base ▪ Recommended actions have the potential to affect outcomes relevant to patients/populations (e.g., improve desirable patient-relevant outcomes, mitigate undesirable patient-relevant outcomes). o The guideline reports how the importance of outcomes to patients was determined o The guideline describes how to tailor recommendations for application to individual (or subsets of) patients or populations (e.g., based on age, sex, ethnicity, comorbidities). ▪ **Values and preferences of target users** o Values and preferences of guideline target users, as they relate to the recommended actions, have been sought and considered o Factors related to target user acceptability of the recommended actions have been considered (e.g., the acceptability of learning new clinical skills or the need to adapt current routine) o The guideline differentiates between recommended actions for which clinical flexibility and individual patient tailoring are more appropriate in the decision-making process and those for which they are less appropriate o The guideline describes the range of recommended actions that are acceptable to the clinical community, including the preferred option (if relevant), and describing why it is the preferred choice. ▪ **Values and preferences of patients/populations** o Values and preferences of the target population (including patients, family and caregivers, if appropriate) have been sought and considered o Factors related to patient/population acceptability of the recommended actions have been considered (e.g., motivation, ability to achieve outcomes, expectations, perceived effectiveness) o The guideline differentiates between recommended actions for which patient choice and/or values are likely to play a large part in the decision-making process and those for which they are likely to play a small role o The guideline states whether tools to assist in patient decision-making would be beneficial. ▪ **Values and preferences of policy/decision-makers**

	○ Information about the needs of policy and decision-makers has been sought and considered in the formulation of the recommendations ○ The effect of the recommendations on policy and system-level decision-making has been considered in the formulation of the recommendations ○ The effect of the recommendations on health equities has been considered in the formulation of the recommendations ○ The guideline describes where changes to policy should be made to align with the recommendations. ▪ **Values and preferences of guideline developers** ○ There is a clear description of the values and preferences that guideline developers brought to the development process ○ There is a clear description of how guideline developer values and preferences influenced their interpretation of the balance between benefits and harms ○ The method used to integrate values and preferences, including when they differ between stakeholders (e.g., target users, patients/population, policy makers), is described. ▪ **Purpose** ○ The guideline recommendations align with the implementation goals of the guideline (e.g., for advocacy or policy change) ○ The anticipated effects of recommendation adoption on individuals (e.g., patients, populations, target users), organisations, and/or systems are described. ▪ **Local application and adoption** ○ The guideline describes the types and degree of change required from current practice ○ The guideline differentiates between recommendations for which local adaptation may be more or less relevant ○ The guideline articulates relevant factors important to its successful dissemination ○ The guideline developers considered the issues that can influence the adoption of the recommendations and provided tools and/or advice for guideline implementers related to: ▪ How to tailor recommendations for the local setting ▪ Resource considerations needed to implement the recommendations (e.g., human resources, equipment) and their associated costs ▪ Economic analysis (e.g., cost-effectiveness or cost-utility) of recommended actions (if appropriate) ▪ Competencies and/or training of personnel required to implement the recommended actions ▪ Data required to implement and monitor the adoption of recommended actions ▪ Strategies to overcome barriers related to healthcare professional acceptability and/or patient/population and/or policy acceptability of the recommended actions ▪ Criteria that can be used to measure recommendation implementation and quality improvement.
What are the strengths and limitations of the tool?	**Strengths of the AGREE-REX tool** ▪ The authors state that AGREE-REX is a usable, reliable, and valid tool to evaluate CPG recommendations

	▪ Methodologic standards of measurement design were used in its development ▪ The use of multidisciplinary literature as a basis for the concepts underpinning AGREE-REX ▪ Its development by a multidisciplinary international research team and engagement of 322 internationally representative participants involved in clinical practice guidelines. The participants reaffirmed the need for the tool, and their participation was seen as vital to ensure that the resource was tailored to the needs of the international clinical practice guidelines communities. **Limitations** ▪ There was no summative evaluation of the appraisal tool ▪ The measurement properties and usability surveys were performed with the penultimate draft version of the tool and not the final version ▪ The criteria used to select the clinical practice guidelines (<50 pages, English language only) and the application of the tool to the whole set of recommendations in each report, as the criteria and unit of recommendation may affect the perceptions of the tool and its measurement properties.
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	N/A
What are the core elements of the key innovation?	N/A
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	N/A
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/A
How is the innovation used in practice?	N/A
Notes	
Reviewer notes	Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: AGREE-REX – Appraisal of Guidelines Research and Evaluation–Recommendations Excellence; N/A – not applicable; N/R – not reported.

RQ3

Innovations in evidence and or guidance translation (not evaluated)

Table C31 GRADE Evidence to Decision (EtD) frameworks for adoption, adaptation, and de novo development of trustworthy recommendations: GRADE-ADOLOPMENT

Publication identification	
Authors (year)	Schünemann et al. (2017)
Country	International
DOI	https://doi.org/10.1016/j.jclinepi.2016.09.009
Publication description	
Design	Mixed methods: Description of model/framework development
Objective	To describe a potentially efficient model for guideline production based on adoption, adaptation, and/or de novo development of recommendations utilising the GRADE EtD frameworks.
Summary/Overview	The study describes a methodology that combines the advantages of adoption, adaptation, and de novo development of recommendations ('GRADE-ADOLOPMENT') based on GRADE EtD frameworks. GRADE-ADOLOPMENT allows for creation of context appropriate recommendations.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	**GRADE-ADOLOPMENT framework** GRADE-ADOLOPMENT combines the advantages of adoption, adaptation and de-novo development of recommendations based on the GRADE EtD frameworks. The structure of the GRADE EtD frameworks and the criteria that determine the direction and strength of a recommendation allows adolopers to create recommendations appropriate for their context.
What are the core elements of the key innovation?	The cornerstones of the adolopment approach are to: ▪ Identify and prioritise credible existing guidelines or evidence syntheses of interest and relevance after or before priorities are set by a guideline group. This step should involve the relevant stakeholders. It involves deciding to accept or modify whole guidelines or their specific recommendations by considering whether they are credible, up to date, acceptable, and applicable given the cultural and organisational context. ▪ Evaluate and complete GRADE EtD frameworks for either a matched recommendation or a new recommendation. This step involves identifying and reviewing information of existing EtD frameworks or identifying information that informs the EtD criteria and completing a new EtD for the adoloped recommendation. ▪ Final adoption, adaptation, or de novo creation of recommendations based on the extent of changes made to the original recommendation or degree of work involved.

What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	Considerations that guide the choice of the guideline development approach include the availability of monetary and nonmonetary resources, credibility, maximisation of uptake, the benefits of sharing information widely, and the avoidance of duplication of efforts. Organisations that produce guidelines will need to decide on the best approaches to develop guidelines and to design detailed strategies and build capacity to implement them. Previous work with international organisations and health authorities on guideline development has addressed the need to compile and update evidence in sharable formats while allowing for consideration of context-specific factors. The GRADE EtD tables and frameworks are increasingly used to produce guidelines. The EtD frameworks provide information on criteria that are relevant to guideline recommendations (e.g., health benefits, harms, certainty in the best available evidence, cost, feasibility) and how the panellists judge the effect of this information on the final recommendation. The EtD frameworks may facilitate the adoption or adaptation of guidelines to the setting, context, and culture of a specific jurisdiction or country. The authors developed and tested an approach for adoption, adaptation, and de novo guideline development based on the GRADE EtD frameworks. To complete this work, the authors applied prior work on adaptation of guidelines to address the challenges guideline developers face.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	GRADE-ADOLPMENT was developed as a result of establishing a new national guideline programme by the Ministry of Health in the Kingdom of Saudi Arabia (KSA). Twenty-two guidelines/topics that aligned with topics of interest to the Saudi Centre for Evidence-Based Health Care were identified for adaptation by the KSA. The effective time needed to complete ten guidelines with 80 recommendations during one phase was approximately 4 months, and to complete 12 guidelines with 146 recommendations during a second phase was approximately 6 months. Examples of the adoption, adaptation or creation of de novo recommendations from this process are outlined below. **Adoption of recommendations** ▪ The GRADE EtDs ask guideline panels to consider criteria that influence the direction and strength of a recommendation as well as its implementation. The guideline panel should evaluate the evidence, judgements, and decisions of the original recommendation from the source guideline and address agreement and disagreement with these judgements. If judgements do not differ sufficiently to change the direction and strength of a recommendation, panel members will adopt the recommendation as is. If judgments differ, panel members will want to change, i.e., adapt, the recommendation. Whether they adopt or adapt the recommendation, the EtD framework helps the guideline panel to consider criteria that address implementation and possible research gaps, even those specific to the setting. ▪ The EtD framework in one of the KSA Ministry of Health guidelines details an adopted recommendation from an existing guideline for initiating dialysis in adult patients. This recommendation was adopted from an existing guideline after considering all criteria in the EtD framework, local evidence, and additional considerations in the EtD. Adoption was facilitated by understanding the judgements that led to formulating the original recommendation. **Adaptation of recommendations** ▪ A guideline panel following the EtD framework may decide that their judgements differ from those of the original guideline panel and, thus, they may provide a recommendation that differs from the original one.

	▪ The guideline panel for one of the KSA Ministry of Health guidelines adapted a recommendation from the Canadian Task Force guideline on breast cancer screening. The Canadian guideline provided a weak recommendation against screening in 40- to 50-year old women, whereas the KSA guideline made a conditional recommendation in favour of breast cancer screening in 40- to 50-year old women because of the presumed higher baseline risk (affecting the problem criterion in the EtD and the absolute risk reduction in the benefits and harms criterion) in younger women in the KSA. Constructing the EtD framework and extracting information from the original recommendation facilitated understanding and explanation of the reasons for disagreement, in this case presumed different baseline risks. **De novo development** ▪ The KSA guideline panel created a de novo recommendation that addressed the question "Should multi-vessel versus culprit vessel only percutaneous coronary interventions be used in patients with acute ST-wave elevation myocardial infarction and multi-vessel coronary artery disease be used" identified two new trials compared with an evidence synthesis used for a NICE guideline. The number of trial participants increased from approximately 200 (two trials) to 1,000 (four trials). Although the NICE guideline panel refrained from developing a recommendation because of paucity of the evidence, the KSA panel, through updating the search, developed a new recommendation. Although the balance based on the clinical evidence for effects favoured benefits over harms for patients, consideration of factors such as the local baseline risk and feasibility of administering that intervention in the local healthcare setting impacted on the direction and the strength (i.e., weak/conditional or strong recommendation) of the recommendation (The panel suggests multi-vessel Primary Percutaneous Coronary Intervention over culprit-only Percutaneous Coronary Intervention for patients with multi-vessel coronary artery disease undergoing Primary Percutaneous Coronary Intervention [conditional recommendation; low-quality evidence]). **Required resources** ▪ Although the time and resources required for the adoption or adaptation of recommendations were less than that of developing all guidelines de novo, the approach still required specific expertise in guideline development and evidence synthesis, having a designated methodology lead for each guideline developed by the KSA Ministry of Health, research librarian support for updating literature searches, methodological expertise for updating evidence syntheses and analyses, and experience in facilitation of panel meetings.
Notes	
Reviewer notes	RQ3: Innovation of GRADE –ADOLOPMENT framework to inform context appropriate guideline recommendations. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: EtD – evidence to decision; GRADE – Grading of Recommendations Assessment, Development and Evaluation; KSA – Kingdom of Saudi Arabia; N/A – not applicable; NICE – National Institute for Health and Care Excellence; N/R – not reported.

Table C32 Current practices and challenges in adaptation of clinical guidelines: a qualitative study based on semi-structured interviews

Publication identification	
Authors (year)	Song et al. (2021)
Country	International
DOI	http://dx.doi.org/10.1136/bmjopen-2021-053587
Publication description	
Design	Qualitative study
Objective	To better understand the current practice of clinical guideline adaptation and identify challenges encountered in this process.
Summary/Overview	The study describes the practice of clinical guideline adaptation, which is increasingly used to develop clinical guidelines. The core steps for the adaptation process are 1) selection of scope and source guideline, 2) assessment of source materials (guidelines, recommendations and evidence level), 3) decision-making process, and 4) external review and follow-up process. This study is part of the RIGHT-Ad@pt project, which aimed to develop a reporting checklist for clinical guideline adaptation.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	CG adaptation
What are the core elements of the key innovation?	Qualitative study participants worked in nine different organisations from seven countries, the majority being from high-income countries (60%; 6/10). Most participants had over five years of experience in CG adaptation (70%; 7/10). Most of the included organisations were research/knowledge-producing centres (67%; 6/9), had over five years of experience in CG adaptation (78%; 7/9), had a working group size that ranged from 6 to 20 members (78%; 7/9) and spent less than two years to complete their adaptation process (78%; 7/9). Most of these organisations had funding sources from government, medical association operation fees, national/international foundations, or the combination of those above (78%; 7/9). **Main steps in the adaptation of CG according to findings from the qualitative study** **Selection of the scope and source guideline(s):** ▪ CG adaptation groups defined or identified CG topic, scope and key questions before or after the selection of source CGs. The screening criteria of source CGs for a further appraisal at this preliminary stage were: (1) stakeholders' preferences of CG topic; (2) a good reputation of the CG developers; (3) methodological quality

	of the source CGs; (4) clinical relevance to the target context and (5) conflicts of interest management and funding independence of the source CGs. **Assessment of source materials:** ▪ CG adaptation groups reviewed and assessed source CGs. The authors stratified this step into three levels based on participants' reported practice: o Guideline level: The guideline quality, trustworthiness, transparency of the process, value and relevance to clinical practice, resource availability and inclusion of latest evidence (up to date) were assessed. ▪ Quality assessment ▪ Checking publication date ▪ Trustworthiness ▪ Applicability. o Recommendation level: The recommendation content, the formulation process of source recommendations (e.g., how the net benefit, resources, patients' values and other criteria were considered), as well as the strength of the recommendation were reviewed. ▪ Recommendation content ▪ Recommendation consistency ▪ Evidence interpretation ▪ Evidence to decision process. o Evidence level: The certainty of the evidence of the source recommendations was reviewed. ▪ Re-rating certainty of evidence ▪ Updating search of source guideline(s) ▪ Supplementing with new evidence. **Decision-making process:** ▪ CG adaptation groups reviewed the summarised evidence and decided whether to adapt (with modifications) or adopt (without modifications) the source recommendations. o Reviewing the summarised evidence/assessment results o Making decisions on whether to adapt or adopt source recommendations. **External review and follow up:** ▪ Following the decision-making process, an external review or a peer review process was conducted. A follow-up process was scheduled, including the plan for dissemination, monitoring and updating. o External review or a peer review process o Dissemination o Monitoring o Updating.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	The main reasons for adapting clinical guidelines according to the study participants were: ▪ To develop their own CGs ▪ To implement or endorse source CGs ▪ To update an existing CG and ▪ To analyse conflicting recommendations from different source CGs.

	The most common reason to adapt given by study participants was to develop CGs for their intended setting based on other existing CGs, by retrieving and adapting existing CGs that could potentially answer their questions, saving resources and time and avoiding duplication of efforts.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	▪ **Addressing context differences between source CG(s) and adapted CG** o Differences between source CGs and target context were addressed mainly through panel discussion and experts' opinions o CG adaptation groups could address these differences at multiple levels: (1) at CG level, by prioritising source CGs according to different criteria or discarding the entire source CG if the difference between the source CG and target context was large enough; (2) at recommendation level, by modifying the strength of recommendations due to differences between the source CG and target context after considering the balance of the benefits and harms, other factors (e.g., acceptability or feasibility) or formulating new recommendations (e.g., new recommendations for subgroup population) and (3) at evidence level, by supplementing with new evidence (e.g., local data). ▪ **Addressing inconsistencies between recommendations from different source CG(s)** o Inconsistencies between recommendations were addressed by prioritising those source CGs that (1) had good quality or rigorous development process, (2) were relevant to the target context, (3) were most up to date, and (4) were considered trustworthy. o The reasons behind the inconsistency were also assessed at the recommendation and evidence level. At the recommendation level, this involved assessing whether (1) the inconsistency was due to a different target population, (2) the evidence was sufficient or up to date and (3) the evidence was appropriately interpreted. At the evidence level, this involved assessing whether the source evidence was appropriately assessed. ▪ **Updating source evidence** o CG adaptation groups sometimes used evidence that was more recent or relevant in addition to the source evidence. To identify new evidence, guideline developers relied on literature searches, including a full de novo search or pragmatic search (e.g., PubMed, local databases or Cochrane database), updating the source search or experts' suggestions. o If the evidence base of the source CGs was unclear or did not answer the clinical questions, developers conducted a de novo CG development process, discarded the recommendation or formulated recommendations based on the guideline panel discussion.
Notes	
Reviewer notes	RQ3: Innovation of adaptation of clinical guidelines. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: CG – clinical guideline; N/A – not applicable; N/R – not reported; RIGHT – Reporting Items for practice Guidelines in HealThcare.

Innovations in evidence synthesis (not evaluated)

Table C33 Qualitative Evidence Synthesis (QES) for Guidelines: Paper 1 – Using qualitative evidence synthesis to inform guideline scope and develop qualitative findings statements

Publication identification	
Authors (year)	Downe et al. (2019)
Country	International
DOI	https://doi.org/10.1186/s12961-019-0467-5
Publication description	
Design	Consensus methodology – nominal group technique: The WHO convened a group of methodologists involved in developing recent (2010–2018) guidelines that were informed by QES. Using a pragmatic and iterative approach that included feedback from WHO staff and other stakeholders, the group reflected on, discussed and identified key methods and research implications from designing QES and using the resulting findings in guideline development.
Objective	To describe and discuss methods for conducting a QES in the context of developing a guideline, so that QES findings can (1) inform the scope of a guideline and (2) be used to develop findings for key guideline decision-making criteria.
Summary/Overview	Following on from the WHO addition of a chapter on the use of evidence from qualitative research to develop WHO guidelines to the *WHO Handbook for Guideline Development* in 2015, the authors describe the methods for conducting a QES in guideline development, with the examples of the WHO antenatal and intrapartum care guidelines.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	Methods for conducting a QES in the context of developing a guideline, so that QES findings can (1) inform the scope of a guideline and (2) be used to develop findings for key guideline decision-making criteria.
What are the core elements of the key innovation?	**Overview of how to conduct a QES in the context of guideline development** ▪ Qualitative review protocol ○ As in other qualitative systematic reviews, the protocol should include the objective of the review, criteria for including studies (types of studies, participants, settings, interventions and phenomena of interest), the search strategy, data collection and analysis, and a reflexivity statement. ▪ Reflexivity statement ○ Expresses the *a priori* views, values and beliefs of the review authors about the subject of interest. It is intended to provide some transparency and give readers an insight into the lens through which the authors have viewed their data. ▪ Search methods ○ Ideally, an initial scoping search should be conducted prior to the framing of the guideline parameters to identify potential concepts, e.g., values and associated outcomes that may be

	important to the population under investigation. Where this has been done, the findings from the scoping review may guide the subsequent QES search criteria. ▪ Preparing an effective search strategy o Characteristics include: database selection, date range, types of publications, language. ▪ Study selection o Unlike the techniques used to identify quantitative studies for systematic reviews or meta-analyses, the authors suggest that it is not essential to identify and include every available relevant study. The purpose of QES is interpretive rather than predictive. Important, transferable concepts (or themes) are unlikely to change substantially in subsequent studies once they are consistently found in a body of papers from a wide range of participants and contexts. The number of studies included in any specific QES will therefore depend on the variety of concepts identified, the range of sociocultural contexts of interest to the guideline, and the degree of agreement between studies on the emerging concepts and themes. ▪ Quality assessment o There is as yet no standardised tool for the quality appraisal process. ▪ Sampling o Reviewers should seek to ensure that no one sampling system affects the overall quality of the review by introducing reviewer bias. There are a number of sampling methods as well as a variety of approaches, and reviewers should be aware of the different techniques before deciding which to use. ▪ Demonstrating rigor in study selection o A Standard Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram should be presented to demonstrate the decisions that led to the final study inclusion. **QES to develop summary of findings statements for EtD Frameworks** Whereas *a priori* scoping reviews are broad and conceptual, qualitative reviews to develop findings for key EtD framework criteria within the guideline protocol are directed by the types of interventions that are being examined. ▪ Data analysis o The main purpose of an EtD-orientated QES is to generate a series of findings from the included data, which are directly focused on interventions addressed in the guideline, assessed for confidence and tailored towards acceptability, feasibility and equity, and the values that stakeholders attribute to the outcomes associated with the intervention. o The findings are then added to the guideline EtD frameworks, prior to guideline panel consideration. ▪ Tailoring QES findings statements for EtD frameworks o Once the review findings for a QES have been generated, reviewers should start drafting short statements that describe the findings data. The statements associated with each finding need to be framed with end-users and key stakeholders in mind, and the review team should consider what these potential users would want to know. o Each finding statement should be clear and concise and accurately capture the meaning of the underlying data that contribute to it. Each one should include an assessment of confidence in the

	contributing evidence. A finding statement should be developed iteratively so that key concepts can be clarified and explored, but it should be no more than a few sentences in length. ▪ Demonstrating rigour in study analysis (for both types of QES, that is, to inform the scope of a guideline and to develop findings for key guideline decision-making criteria) o As for all systematic reviews, the characteristics and quality assessment of the included studies should be presented in the review, along with a summary of the reasons for excluding studies. o A table listing the review themes and/or a Summary of Findings should be included, which lists the qualitative codes from the included studies that contributed to each theme or finding. o For each finding included in the Summary of Findings, the CERQual rating should also be listed, with reasons for downgrading if this has occurred. o A review finding may be downgraded if it fails to meet any of the four appraisal components (methodological limitations, relevance, coherence and adequacy) inherent in the CERQual tool. This can also be done for thematic findings.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	There is increasing interest in the use of qualitative evidence to inform decisions in a wide variety of sectors, including health and social care, prison care, and education. However, until recently, the decisions made by guideline panels about criteria such as values and preferences, acceptability, feasibility and equity implications have been largely based on the expert opinion of guideline development groups at WHO and/or on evidence that they happen to know about or that has been collected ad hoc, rather than on a systematic review of relevant research. Evidence from QES can be used alongside effectiveness evidence to inform all stages of developing a guideline, including identifying the relevant interventions and outcomes at the scoping stage, synthesising and evaluating evidence, formulating recommendations, and developing implementation considerations. QES reviews conducted at the scoping stage, before the guideline protocol is finalised, can identify broader concepts that can shape the overall scope of the guideline. Once the protocol is finalised, QES reviews designed to inform EtD frameworks are tailored to identify the acceptability, feasibility and/or equity of a specific intervention within the guideline, and/or to inform judgements about how much stakeholders might value the outcomes associated with the intervention.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	**Example of the use of QES** ▪ Summary of qualitative synthesis methods used for the QES at the guideline scoping stage for WHO guidelines on antenatal and intrapartum care o Step 1. The included papers were examined, and an index paper was selected that best reflected the focus of the review. o Step 2. The themes and findings identified by the authors of the index paper were entered onto a spreadsheet, to develop an initial thematic framework. o Step 3. The findings of all the remaining papers were then mapped into this framework, which continued to develop as the data from each paper were added. This process included looking for what was similar between papers ('reciprocal analysis') and what contradicted ('disconfirms') the emerging findings ('refutational analysis'). For the refutational process, as each paper was added to the analysis, the authors consciously looked for data that could disconfirm the emerging themes or prior beliefs related to the topic of the review. If any disconfirming data were found, the themes

	were amended, so that they continued to capture all the data from the papers that had already been analysed as well as taking account of the new insights. ○ Step 4. All the themes were translated (or synthesised) into a 'line of argument synthesis', based on theoretical concepts that explained the data at a conceptual level. ▪ Findings of scoping reviews to inform the WHO guidelines on antenatal and intrapartum care ○ The themes emerging from the data in the scoping reviews led to 'line of argument' syntheses. A robust line of argument is more than the sum of the parts of the review. The authors suggest that it has high theoretical transferability beyond the specific included studies, and as such is likely to be applicable in a wider range of settings and circumstances. For both the antenatal and intrapartum care guidelines, the lines of argument derived from the scoping reviews were used to inform and direct the philosophical framing of the guideline recommendations. ○ The findings illustrated that what matters to women around the world in relation to both pregnancy and childbirth is both safety (physical, clinical, psychological and emotional) and a positive experience. ○ These components were then summarised into a single composite outcome for each review, termed 'positive pregnancy experience' and 'positive childbirth experience', respectively. The positive experience concept captures factors that are part of the standard outcomes dataset for maternity care effectiveness reviews, i.e., mortality and morbidity. However, the concept also encompasses factors such as psychosocial and emotional outcomes in both the short and longer term.
Notes	
Reviewer notes	RQ3: innovation of qualitative evidence synthesis in guideline development. Category of evidence: Grade D.
Associated handbook(s)	WHO handbook for guideline development, 2nd Edition.

Key: EtD – evidence to decision; GRADE-CERQual – Grading of Recommendations Assessment, Development, and Evaluation-Confidence in the Evidence from Reviews of Qualitative research; N/A – not applicable; N/R – not reported; PRISMA - Preferred Reporting Items for Systematic Reviews and Meta-Analyses; QES – qualitative evidence synthesis; WHO – World Health Organization.

Table C34 Qualitative Evidence Synthesis (QES) for guidelines: Paper 2 – Using qualitative evidence synthesis findings to inform evidence-to-decision frameworks and recommendations

Publication identification	
Authors (year)	Lewin et al. (2019)
Country	International
DOI	https://doi.org/10.1186/s12961-019-0468-4
Publication description	
Design	Consensus methodology – nominal group technique: The WHO convened a group of methodologists involved in developing recent (2010–2018) guidelines that were informed by QES. Using a pragmatic and iterative approach that included feedback from WHO staff and other stakeholders, the group reflected on, discussed and identified key methods and research implications from designing QES and using the resulting findings in guideline development.
Objective	To describe and discuss how findings from QES can be used to populate key EtD framework criteria for decision-making in guideline development and to inform recommendations.
Summary/Overview	The paper describes ways in which findings from qualitative evidence synthesis can be used to populate EtD framework criteria for decision making. It describes methods to find relevant qualitative evidence, things to consider while developing narrative text and allocation of findings to different criteria in the EtD framework.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	QES

What are the core elements of the key innovation?	**Identifying relevant qualitative evidence**
	▪ Findings from a QES may enter a guideline process in two ways:
	  ○ Through already-published syntheses that address the guideline questions directly or indirectly
	  ○ Through one or more syntheses commissioned for the guideline. These may include both broad QES covering multiple guideline interventions and 'mini-QES' focusing on a specific intervention.
	▪ Syntheses used in a guideline may focus on people's views regarding the interventions addressed by the guideline, such as communication interventions in labour. Syntheses may also focus on the problem or issue underlying the interventions being addressed by the guideline, for instance, the ways in which women and healthcare providers communicate during labour. Syntheses may also include evidence that is more, or less, direct or relevant, in relation to the guideline question. For example, a synthesis may focus on the views of people in a specific context, such as primary healthcare, while the guideline may include all levels of healthcare. Such differences are taken into account when assessing confidence in the evidence using the GRADE-CERQual approach.

	▪ Syntheses vary in how their findings are presented, depending on whether a more aggregative or interpretive synthesis method is used, on whether 'thick' or in-depth data underlie a synthesis finding, and on the review authors' writing style. Where a synthesis aims to provide explanations or build theory, the findings may be presented both narratively and figuratively, for example, in the form of an infographic or logic model. These infographics and logic models can be incorporated into an EtD where appropriate, for example, where they help to explain factors affecting the acceptability of an intervention. **Populating evidence-to-decision framework criteria with qualitative evidence – principles and processes** ▪ **Allocate the findings to the different criteria in the EtD frameworks** o A QES finding may be relevant to more than one EtD criterion (for instance, to both intervention acceptability and feasibility), and sometimes a pragmatic decision will need to be taken on where to place the finding. Overall, the technical team needs to ensure that the relevant findings are reported somewhere in the framework so that they can be taken into account in decision-making. o The findings from QES may be relevant to more than one framework. Such findings can either be repeated in each relevant framework or included in an overarching text linked to multiple frameworks. o Where an overarching narrative is developed, the technical team need to ensure that it is clear to the guideline panel that the qualitative evidence for several frameworks is presented in an overarching document, and each EtD needs to link to this document. While the same qualitative evidence might be relevant to different guideline questions, the guideline panel's judgements for each criterion might differ, depending on the intervention evaluated in each question. o Wider, less specific findings may need to be used in relation to an intervention where more specific findings are not available. o The technical team needs to ensure that people using the recommendations are able to understand the justification for each recommendation from the evidence presented. ▪ **Weave the individual QES findings into a narrative for each framework criterion** o Once the findings have been allocated to a specific criterion, the guideline technical team needs to weave these findings into a single, short narrative for inclusion in an EtD framework. This narrative should also include the CERQual assessments for the included findings. The following principles could be followed to develop the narrative text: ▪ The narrative should include the key points from the findings that are relevant to the decision that the framework will inform. ▪ The narrative should include enough information on the context of the findings (for instance, that participants were from remote rural communities) to reduce ambiguity and allow interpretation, including of the relevance of the evidence as assessed using CERQual. ▪ A graded entry or layered approach to presenting information may be helpful, with the most high-level information presented in the EtD framework. In a graded entry format, users can then navigate from this summary to more detailed information, for example, the full summary of qualitative findings table, and from there to the full synthesis report. ▪ Users should be able to trace back from the narrative to the individual findings that informed the narrative. Traceability can be enhanced by giving a unique code to each QES finding and including these codes in the narrative.

	▪ **Consider whether any additional considerations need to be included in each framework** o There may be circumstances in which other qualitative, or related, evidence or information needs to be included for a particular framework criterion, in addition to the findings of the contributing QES. This additional information may also be needed where no relevant evidence was found by the QES. This additional evidence might include: ▪ Descriptions of conceptual or theoretical frameworks that help in understanding the QES findings or that place these within a wider context. ▪ Findings from individual qualitative studies that provide important contextual information related to the setting of the recommendation or decision but that were not eligible for inclusion in the QES. ▪ Plausible reasons for anticipating that the intervention might or might not be acceptable to key stakeholders or might be difficult to implement, particularly where little or no evidence on acceptability or feasibility was found for an intervention. ▪ Any assumptions made in relation to the findings presented and, if relevant, the basis for those assumptions. ▪ **How people value the outcomes** o The guidance on populating an EtD framework notes that the direction of a recommendation may change where there is uncertainty about how those affected by an intervention value the outcomes of interest. Additionally, the strength of a recommendation may be affected by research evidence showing that different groups value the desirable and undesirable effects differently. o Three complementary sources for evidence on how people value outcomes in relation to an intervention or option: ▪ Utility values ▪ Studies that directly measure the choices people make when presented with the probabilities of the desirable and undesirable effects, a description of those outcomes (health states) and information about when they would occur and how long they would last ▪ Qualitative evidence from studies that explore people's views of the impacts of different health issues and interventions. ▪ **Gender, health equity and human rights approach** o There are two ways in which guideline technical teams have used qualitative evidence to populate the gender, health equity and human rights impacts section within the EtD framework; ▪ Firstly, issues may be identified directly from the findings of a QES. ▪ Secondly, where a QES undertaken for a guideline does not identify gender, health equity or human rights issues explicitly, it may be possible to infer these from the findings through discussion within the technical team or experts in the field. o A narrative summary of the issues can then be created. Where this is done, it is important to indicate to those making recommendations that these issues were hypothesised from the evidence rather than being described there explicitly and the technical team should consider including these issues under 'Additional considerations' in the EtD framework. ▪ **Acceptability and feasibility**

	○ Qualitative evidence on the acceptability and feasibility of different interventions is often linked. The technical team will often need to take pragmatic decisions on whether to report QES findings in the acceptability or feasibility sections of the EtD framework. As a recommendation is based on judgements regarding all of the evidence presented in a framework, where best to place a specific relevant QES finding is less important than ensuring it is included. ○ When the technical team starts to develop the summary narratives for the acceptability and feasibility sections of the EtD framework, they should also consider how to convey the extent to which the evidence shows similarities and differences across stakeholders and contexts. ○ In some cases, it may be appropriate to include separate narratives for different stakeholders or contexts. When no specific qualitative evidence for a particular option, stakeholder group or context is found, it may be possible for the technical team to draw inferences from findings for other options, stakeholders or contexts. Where inferences are made, this should be made clear in the relevant framework. ▪ **How qualitative evidence synthesis findings may influence guideline recommendations** ○ When making a recommendation, a guideline panel should take into account all of the evidence presented in the EtD framework. ○ The extent to which the qualitative evidence included in a framework influences or drives a decision regarding a particular recommendation will vary across the questions considered by a guideline – in some cases, a decision may be driven by other information presented in the framework. Regardless, all judgements should be supported by a clear justification that refers to the key criteria that drove the decision. **Criteria of the GRADE evidence-to-decision framework and where qualitative evidence might be useful in relation to these criteria** ▪ How large are the positive (desirable) effects of the intervention? ○ Not applicable ▪ How large are the negative (undesirable) effects of the intervention? ○ Not applicable ▪ What is the overall certainty of the evidence of effects? ○ Not applicable ▪ Is there important uncertainty about or variability in how much people value the outcomes and/or interventions? ○ QES at the scoping stage of the guideline or decision process ▪ What is the overall balance of effects? ○ QES findings in how the key stakeholder groups, including citizens, service users and service providers, value different outcomes ▪ How large are the resource requirements? ○ Not applicable ▪ What would be the impacts on gender, health equity and human rights? ○ QES findings on equity issues such as barriers and facilitators to accessing the option ▪ Is the option acceptable to key stakeholders? ○ QES findings on the acceptability of the option ▪ Is the option feasible to implement?

	○ QES findings on the feasibility of the option ▪ What are the implementation considerations? ○ QES findings that informed the other framework criteria can be used to develop or infer implementation considerations **Commissioning QES to inform a guideline** ▪ Stages: ○ Identifying the areas and topics for a QES – the guideline technical team identifies the broad areas or topics for which a QES will be needed; this could include a QES to inform the scoping of the guideline or a QES to inform specific criteria that are part of an EtD framework (such as the acceptability and feasibility of an intervention). ○ Identify synthesis leads and teams – ideally teams should include at least one person with extensive experience in qualitative evidence synthesis and a person with content area expertise in relation to the guideline topic. ○ Discussion of the scope of each synthesis – where more than one synthesis is being commissioned for a guideline, it may be helpful to hold a meeting of the guideline technical team and the synthesis lead authors to consider the scope and objectives of each synthesis. This discussion should include the range of questions that the synthesis will consider, in relation to the EtD criteria used for the guideline. For example, should the synthesis consider equity and human rights issues and resource use issues, in addition to intervention acceptability and feasibility? The discussion should also cover which synthesis approach/es to use, based on which would be most appropriate for addressing the synthesis objectives, how the QES findings will be used within the EtD frameworks, and how best to tailor the synthesis to address the specific needs of a guideline process ○ Preparing the terms of reference – this would include which databases will be searched; how the synthesis findings will be prepared for the guideline, including the types of information and data that will be included in the CERQual Qualitative Evidence Profiles and Summary of Qualitative Findings tables; how an assessment of confidence in the evidence will be made; the content of the final manuscript; and how the technical team and synthesis leads will communicate during the process of producing the syntheses. ○ Develop a protocol for each synthesis – where more than one synthesis is commissioned for a guideline, it may be helpful to ensure (as far as possible) that the synthesis processes are standardised across protocols and make sense in relation to the synthesis objectives. Where possible, the protocol/s should be made publicly available (through, for example, registering the synthesis with Cochrane EPOC, Prospero etc.). ○ A budget for the review should be estimated. In addition to time to conduct the review, person-time needs be included for undertaking a CERQual assessment; several rounds of discussion of the review findings between the synthesis team and the guideline technical team, to ensure that the findings are written as clearly as possible and are congruent with the underlying data; reviewing any summarised findings prepared for different domains of the EtD frameworks; and preparing the synthesis for publication. A qualitative evidence synthesis is labour intensive process and the additional stages needed to prepare the findings for a guideline process generally add additional person-time to the process.

What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	To address EtD framework criteria such as the acceptability and feasibility of interventions, guideline producers are now exploring the use of qualitative evidence. This has led to growing interest in systematic reviews of qualitative studies (also known as QES) – an approach for synthesising the findings from multiple primary qualitative studies. Like systematic reviews of the effectiveness of interventions, QES can provide key evidence for informing guideline recommendations and other decisions.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	**Implications for practice** ▪ Guideline technical teams ideally need to include, or have access to, people with skills in QES, GRADE-CERQual and in populating and using EtD frameworks. This has implications for the resources required to undertake a guideline development process. ▪ The scoping phase of guideline development is critical for identifying the interventions, stakeholders and contexts relevant to the guideline questions. Decisions on these aspects will shape the scope of the QES undertaken for the guideline and adequate time needs to be allowed for this process, including for interactions with the QES teams. ▪ Technical teams should be aware that the findings of scoping and other QES conducted for a guideline may impact on the range and scope of effectiveness reviews for the guideline. QES findings regarding which interventions are seen as important by stakeholders and how people value different outcomes may need to be fed back into the scoping process for effectiveness reviews commissioned for a guideline. ▪ As the number of published QES increases, it is more likely that an existing QES may be found that addresses some or all of the guideline questions. Searches for existing QES should be done before a new QES is commissioned. ▪ A technical team may need to commission both broad QES that cover multiple guideline interventions as well as 'mini-QES' that focus on one specific intervention. It can sometimes be useful to use rapidly conducted 'mini-QES' to address important gaps in the evidence available for a guideline. ▪ Close collaboration between the QES authors and the guideline technical team responsible for populating the EtD framework may help to ensure that the QES findings are developed and tailored to each EtD framework, and relevant criteria within these frameworks. Close collaboration may also help to ensure congruence between the findings in the published QES and those included in the frameworks. ▪ Users of EtD frameworks need to be able to easily identify the sources of qualitative and other evidence presented in a framework. This traceability requires careful attention to documenting how evidence moves from primary studies, to a QES, and then into a framework. ▪ Technical teams should consider the information and training needs of groups making recommendations in relation to qualitative evidence, and in the use of this evidence in guidelines. Information sessions or training for these groups may be needed in advance of formal meetings of these groups.
Notes	
Reviewer notes	RQ3: Innovation of qualitative evidence synthesis for guideline development. Category of evidence: Grade D.
Associated handbook(s)	WHO Handbook for Guideline Development, 2nd Edition.

Key: EPOC – Effective Practice and Organisation of Care; EtD – evidence to decision; GRADE-CERQual – Grading of Recommendations Assessment, Development, and Evaluation-Confidence in the Evidence from Reviews of Qualitative research; N/A – not applicable; N/R – not reported; QES – qualitative evidence synthesis. WHO – World Health Organization.

Table C35 Qualitative Evidence Synthesis (QES) for guidelines: Paper 3 – Using qualitative evidence synthesis to develop implementation considerations and inform implementation processes

Publication identification	
Authors (year)	Glenton et al. (2019)
Country	International
DOI	https://doi.org/10.1186/s12961-019-0450-1
Publication description	
Design	Consensus methodology – nominal group technique: The WHO convened a group of methodologists involved in developing recent (2010–2018) guidelines that were informed by QES. Using a pragmatic and iterative approach that included feedback from WHO staff and other stakeholders, the group reflected on, discussed and identified key methods and research implications from designing QES and using the resulting findings in guideline development.
Objective	To describe how members of the guideline technical teams have used the findings from QES to develop implementation considerations for WHO guidelines.
Summary/Overview	The paper provides ways in which findings from qualitative evidence synthesis can be used to develop implementation considerations for clinical practice guidelines. The key steps are: 1) Determine the scope of the guideline, 2) Search for relevant QES or commission QES, 3) Assess relevancy of the findings, 4) Summarise the findings or use the findings to prepare draft implementation considerations, 5) Guideline panel agrees on the recommendations or the panel makes comments and suggestions to the drafted implementation considerations, and 6) Formal approval of the recommendations by the WHO and dissemination.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	Use of qualitative evidence synthesis in implementation considerations in clinical practice guidelines.
What are the core elements of the key innovation?	**Moving from qualitative evidence to implementation considerations – the work process** **Step 1**: ▪ The WHO determines the scope of the guideline. **Step 2**: ▪ The technical team searches for existing QES or commissions QES that explore intervention acceptability, feasibility or equity implications or information about the value people place on different outcomes. **Step 3:** ▪ Pathway 1: The technical team assesses which QES findings are relevant for the acceptability, feasibility or equity of the intervention or the value people place on different outcomes.

	▪ Pathway 2: The technical team assesses which QES findings represent a factor that might affect the implementation of the intervention. **Step 4:** ▪ Pathway 1: The technical team summarises the findings and presents these in the Evidence-to-decision framework, together with their CERQual assessment. ▪ Pathway 2: The technical team uses these findings to draft implementation considerations and presents these in the EtD framework. **Step 5:** ▪ Pathway 1: The guideline panel agrees on recommendations. ▪ Pathway 2: Where interventions are recommended, the panel makes comments and suggestions to the drafted implementation considerations and suggests additional considerations. **Step 6**: ▪ The recommendations are formally approved by the WHO, and disseminated together with the implementation considerations.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	Systematic reviews of randomised trials are commonly acknowledged as the best source of evidence when assessing intervention effectiveness. Systematic reviews of qualitative research, also known as QES, are, however, better suited for questions of acceptability and feasibility. Although the global guidelines of the WHO generally do not include implementation plans, EtD frameworks encourage guideline panels to list broader implementation considerations. These are not intended to serve as technical manuals or detailed implementation plans. Instead, they are probes, prompts, suggestions or requirements that implementers should consider when developing their local plans.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	**Implications for practice** ▪ The authors suggest that technical teams should consider using health systems frameworks at an early stage of the process, not only to organise implementation considerations but also to identify them. ▪ To increase access to qualitative evidence about higher levels of the health system, technical teams should consider the scope of commissioned QES and should also consider how they can gather direct input from higher-level stakeholders, for instance, through an increased use of key informant interviews or surveys. ▪ Implementation considerations may also be influenced by overarching ideals and principles held by the guideline producing body, including principles tied to human rights. Guideline commissioners, technical teams and panels should consider adopting a more reflexive and transparent approach early on in the guideline process, where they identify these overarching principles. ▪ Where confidence in the qualitative evidence is low or very low, technical teams should consider formulating implementation considerations based on this evidence cautiously, for instance, as questions, prompts or suggestions. Alternatively, they should consider basing the implementation considerations primarily on those qualitative findings that have been assessed as being of moderate or high confidence.
Notes	
Reviewer notes	RQ3: Innovation of qualitative evidence synthesis to inform guideline implementation.

	Category of evidence: Grade D.
Associated handbook(s)	WHO Handbook for Guideline Development, 2nd Edition.

Key: CERQual – Confidence in the Evidence from Reviews of Qualitative Research; EtD – evidence to decision; N/A – not applicable; N/R – not reported; QES – qualitative evidence synthesis; WHO – World Health Organization.

Table C36 Evidence informed decision making: the use of "colloquial evidence" at NICE

Publication identification	
Authors (year)	Sharma et al. (2015)
Country	Denmark, UK
DOI	https://doi.org/10.1017/S0266462314000749
Publication description	
Design	Systematic review
Objective	To understand the types of colloquial evidence (CE) used across the guidance producing teams at NICE; the extent of its use and how it has been incorporated within their deliberative processes.
Summary/Overview	The study identified that different forms of evidence including CE were used in different ways and for different reasons for guidance development at NICE. On the whole, three broad sources of CE were identified at NICE: evidence from experts and patients/carers, evidence from grey literature and evidence from all stakeholders through public consultation. The study also proposed using the SART system for critically appraising the colloquial evidence.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	The use of colloquial evidence, defined as "evidence about resources, opinion, political judgment, values, culture, and the particular pragmatics of a situation", to complement the scientific evidence, including information that comes informally from different stakeholders such as practitioners and patients, in guideline development at NICE.
What are the core elements of the key innovation?	Three sources of CE ▪ CE1: Evidence from experts (professionals/clinicians) and patients/carers ▪ CE2: Evidence from grey literature ▪ CE3: Evidence from all stakeholders through public consultation Proposed key areas of questioning for critical appraisal of colloquial evidence using the SART system <u>Source</u> ▪ Is the source of the CE credible? o by whom was it written/spoken and what was its purpose o determine to what extent one could trust the author(s) o are there certain conflicts of interests to be noted o did it appear to be objective or were there any potential biases present o who was the sponsor/funder of the source o did the information only present one point of view.

	Accuracy ▪ Is the information accurate? o consider how plausible the claims are, in terms of whether they could be verified by other sources or testimonies o if they contain any footnotes or state any references that can be checked and how reliable are the sources cited o are the aims available and information presented clearly o are the methods used for collecting that information transparent and clearly presented o are there any inconsistencies present in the content o does it highlight any areas of uncertainty. Relevance ▪ Is the evidence relevant with respect to the scope of the guidance? o consider how applicable the information really is to the question at hand o is all the information present or does it seem to be incomplete with respect to the question (is it discussing only part of the equation) o is the setting and context clearly defined and how well does it fit with the research question? Timeliness ▪ How timely is the information? o how current is the information o are any dates associated with the information (e.g. the last date when the website was updated if an electronic source, or recent or old experiences if testimonies etc.) o how regularly is the information updated o Is there any more recent research or information available.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	▪ There is often a need to contextualise scientific evidence and to understand how it should be implemented in healthcare practice, as without contextualisation, guidance and policies may fail to produce the desired results. ▪ This shift from evidence-based to evidence-informed decision making has been reflected in the definition of evidence and methodological practices of leading guidance producing organisations such as the Health Evidence Network of the WHO that define evidence as "findings from research and other knowledge that may serve as a useful basis for decision making in public health and health care".
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	The guidance development processes at NICE can be broadly grouped into five stages, namely scoping, evidence review and economic modelling, deliberative process, stakeholder consultation and implementation tool development. CE is used throughout these stages. ▪ CE1: Evidence from experts (professionals/clinicians) and patients/carers o In the CG programme, advice from clinical experts and patients/carers is obtained throughout the guidance development process through GDGs that include health professionals and patient/carer representatives with relevant expertise and experience of the specific guideline topic at hand. o The GDGs for the CG programme have topic specific membership but can still have further co-opted experts (that is, specialists, professionals, relevant commissioners and patients/carers) if required for

	the deliberative processes. These deliberations are summarised within the "considerations" or "evidence to recommendations" sections (depending on the programme) of the final guidance and act as a primary direct source of CE. ○ Additionally, all guidance production at NICE follows a patient and public involvement policy, which sets the platform for the contribution of lay people, and organisations representing their interests, to the work of NICE. ▪ CE2: Evidence from grey literature ○ Across the programmes, if data for all the parameters of an economic model is not available through the scientific published literature, CE in the form of data from various electronic sources can also be used. Data from the Health and Social Care Information Centre website for data on Hospital Episodes Statistics and Healthcare Resource Groups is commonly used for this purpose. ○ In the CG programme, evidence from websites (such as 'health talk online', https://healthtalk.org/) is also routinely used, where websites are searched manually for any additional relevant patient experience information. ▪ CE3: Evidence from all stakeholders through public consultation ○ All programmes have a public consultation of their draft scope, and draft guidance, where registered stakeholders (which include professional groups and societies, patient groups and charities, other NICE staff, the NHS etc.) and industry are able to comment and submit their views on the questions (draft scope) or recommendations (draft guidance) proposed. ○ The technical teams present these comments to the standing committees or temporary advisory bodies, who then consider them through their deliberations, for the final document, to ensure that all comments have been taken into consideration. Therefore, the stakeholder comments can have a direct impact on the final recommendations of any guidance. ▪ Critical Appraisal of Colloquial Evidence ○ All programmes use formal critical appraisal techniques for considering scientific evidence but none had an explicit appraisal checklist for reviewing CE. The mechanism of appraisal of CE at NICE was informal and through deliberative consideration and through stakeholder consultation of those considerations. It could be argued that formal checklists may be less helpful with CE, as it is conceptually different from other forms of evidence. There may, however, be the need for development of some form of evaluation of CE and its contribution to the deliberative process. ○ The authors suggest that the SART instrument is a crude tool that needs to be developed further and tested for robustness. To create a formal checklist, a thorough process such as an expert elicitation Delphi Panel or other formal consensus method should be undertaken. Moreover, as these concepts are solely derived from the limited literature identified by one scoping literature review, this may not represent all the factors that need to be considered and therefore, any interpretations should be made very cautiously.
Notes	
Reviewer notes	RQ3: Innovation of use of colloquial evidence in guideline development. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: CE – colloquial evidence; CG – clinical guidelines; GDG – guidance development group; N/A – not applicable; NICE – National Institute for Health and Care Excellence; N/R – not reported; SART – Source, Accuracy, Relevance, Timeliness; WHO – World Health Organization.

Table C37 Guidance on guidelines: Understanding the evidence on the uptake of health care guidelines

Publication identification	
Authors (year)	Brennan et al. (2016)
Country	UK
DOI	https://doi.org/10.1111/jep.12734
Publication description	
Design	Realist review
Objective	To assess how reviews of the evidence base should be conducted to inform barriers and facilitators to guideline implementation.
Summary/Overview	This paper compares two review approaches – thematic/narrative and realist – to assess how best to conduct reviews to offer practical advice for both the intended recipients of guidelines (for example, practitioners) and those with an interest in whether they are followed (for example, managers, commissioner, and policy makers).
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	**Realist synthesis** Realist synthesis promotes an explanatory role for systematic reviews and seeks to explain why an intervention might work (or flounder) and to uncover the many contingencies that generate success (or failure). The approach assumes heterogeneity in the implementation of and response to any intervention and seeks transferable lessons by focusing the review on "programme theories," which are common to all interventions.
What are the core elements of the key innovation?	Programme theories represent the ideas and assumptions underlying how and why an intervention is expected to work. A wider range of evidence may be drawn into the review, explored in a research design, which extracts, formalises, and tests the programme theories that underpin interventions.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	Major dilemmas that confront guideline use include the following: 1. The tension in using simple guidelines for complex comorbidity; 2. The tension between (inter)national credibility of and local control over guidelines.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	**Realist review example: Development of guidelines to address comorbidity** ▪ Use of alternative methods such as realist synthesis and programme theories to review the primary research on guideline use and adherence in clinical practice

	▪ **Solutions to address the complexity of comorbidity in guidelines** o To increase the comorbid patient's exposure to multiple guidelines, guidelines should: ▪ provide information to enable clinicians to more effectively apply the guideline to patients with multimorbidity; for example, by detailing the percentage of patients with comorbid conditions included in the original trials and the extent to which comorbid conditions may modify treatment effects ▪ shift from "disease specific guidelines" to "patient-centred guidelines" ▪ focus much more on choosing and prioritising treatment and so in theory reduce the tension inherent in following multiple combinations of condition-specific guidance. To work in practice, this theory rests on the idea that clinicians can take account of multiple patient factors in adjusting guideline recommendations to the patient in front of them. o To increase guideline uptake: ▪ adapt guidelines to kindle interest rather than to impart new knowledge. The more local the adaptation the greater the number of interested parties who will be drawn directly into implementing and acquiescing with the scheme—thus increasing the chances of the guideline being followed.
Notes	
Reviewer notes	RQ3: innovation of realist review to explore barriers and facilitators to guideline implementation within complex adaptive health systems. Category of evidence: Grade D.
Associated handbook(s)	N/R

Key: N/A – not applicable; N/R – not reported.

Rapid or living guidance (not evaluated)

Table C38 Developing trustworthy recommendations as part of an urgent response (1–2 weeks): a GRADE concept paper

Publication identification	
Authors (year)	Akl et al. (2021)
Country	International
DOI	https://doi.org/10.1016/j.jclinepi.2020.09.037
Publication description	
Design	Mixed methods
Objective	To propose an approach for developing trustworthy recommendations as part of urgent responses (1–2 weeks) in the clinical, public health, and health systems fields.
Summary/Overview	The study proposes an approach for developing trustworthy recommendations as part of an urgent response (1-2 weeks). The approach offers the alternatives of using existing guidelines to adopt or adapt recommendations; using existing systematic reviews to develop new recommendations; and, when the previous options are not possible, relying on expert panel input to develop new recommendations.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	An approach for developing trustworthy recommendations as part of urgent responses (1–2 weeks) in the clinical, public health, and health systems fields.
What are the core elements of the key innovation?	Steps of the urgent response approach. These assume that existing and adequate guidelines and systematic reviews, but not new rapid reviews, are being considered as evidence. ▪ Assess the level of urgency ▪ Assess feasibility ▪ Set up the organisational logistics ▪ Specify the question(s) ▪ Collect the information (evidence) needed ▪ Assess the adequacy of identified information (evidence) o Adequate information found ▪ Adopt existing recommendation ▪ Adapt existing recommendation. o Adequate systematic review found ▪ Develop new recommendation using existing systematic review. o No adequate recommendation found; no adequate systematic review found

	▪ Develop new recommendation using panel input. ○ Criteria to assess adequacy of information ▪ Relevance or directness ▪ Credibility ▪ Currency or recency. ▪ Develop the recommendations. **Steps for developing the recommendations** ▪ Adopt versus adapt existing recommendations ○ To decide on adopting versus adapting an original recommendation, the expert panel group needs to assess whether the direction or the strength of a recommendation may be affected by any difference in the following factors (ideally using a decision-making framework, e.g., EtD table): ▪ Rating of the importance of outcomes ▪ Indirectness ▪ Rating of evidence ▪ Baseline risks ▪ Perspective ▪ Contextual factors. ○ For each of the above factors, the judgment (between the urgent guidance setting and the original one) needs to be sufficiently different for panel members to adapt the recommendation (i.e., modify its direction and/or strength). The recommendation statement wording may be edited to enhance the usability for the intended target group. ▪ Develop new recommendation using existing adequate systematic review ○ One other reasonable scenario would be to find an adequate systematic review that would jump-start the development of de novo recommendation(s). The development would follow the standard GRADE evidence assessment and recommendation development process with some potential shortcuts to ensure the process is completed within the desired timeframe. ▪ Develop new recommendation using expert panel input ○ When no adequate recommendations are available for adopting or adapting, and no adequate systematic reviews are available to inform a de novo recommendation, using expert panel input as the sole source of evidence is a feasible alternative for an urgent response. It is important to clarify that panel input is still important when developing a recommendation based on existing recommendations or systematic reviews. ▪ Consider an updating plan ○ Urgent situations are typically associated not only with scarcity of data but also with a rapidly developing evidence base and contextual information. This raises the consideration of establishing an updating plan, ideally through a living process. This is particularly relevant when emerging data can potentially lead to a change in the recommendation. The updating plan would need to define the frequency of reassessment of the recommendation (e.g., weekly, monthly), and to be adjusted to the speed of development of the urgent situation and of the emergence of the evidence.

What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	▪ Some situations raise the need for urgent recommendations to support the interventions of clinicians, public health practitioners, and policymakers, for example, to address the COVID-19 pandemic. ▪ Under such circumstances, developing trustworthy recommendations in a sufficiently short timeframe is essential but can be challenging in rapidly changing contexts. Groups producing guidelines specifically need to balance the need for developing a timely response with the need to ensure the trustworthiness of their advice. ▪ The WHO defined two types of guidelines developed in response to an emergency or urgent need: emergency (rapid response) guidelines (produced within hours to days) and rapid advice guidelines. ▪ Thayer and Schunemann (2016)[(2)] defined four levels of urgency for developing recommendations: ultra-short emergency response (1–2 hours), urgent response (1–2 weeks), rapid response (1–3 months), and routine response (more than 3 months). Although more detailed advice exists for routine and rapid responses, so far there is no formal guidance on how to apply GRADE in situations requiring urgent responses.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	**Use of steps for developing recommendations as part of an urgent response in practice** ▪ Adopt vs. adapt existing recommendations o WHO developed interim guidance for the "clinical management of severe acute respiratory infection when COVID-19 infection is suspected" in consultation with the International Forum for Acute Care Trialists, International Severe Acute Respiratory and Emerging Infection Consortium, and Surviving Sepsis Campaign. Guideline developers originally adapted this interim guidance from "Clinical management of severe acute respiratory infection when Middle East respiratory syndrome coronavirus (MERS-CoV) infection is suspected: interim guidance". In the first edition of the COVID-19 guidance, developers adopted many document sections, research questions, and applicable guidance verbatim, adapting some directives to reflect underlying uncertainty about the microbiological profile of COVID-19 when informed by the indirect information from SARS (caused by SARS-CoV1) and MERS (caused by MERS-CoV) cases. ▪ Develop new recommendation using existing adequate systematic review o The Infectious Diseases Society of America developed rapid recommendations on the treatment and management of hospitalised patients with COVID-19, including a recommendation on corticosteroid treatment for hospitalised patients with acute COVID-19. At the time of the first iteration of the guideline, the review team did not identify any direct evidence to inform this recommendation; however, they identified a systematic review reporting on corticosteroid use among patients with SARS-CoV-1 or MERS-CoV. The guideline panel determined this existing review to be direct enough to inform their recommendation. ▪ Develop new recommendation using expert panel input o The panellists are asked to review the literature they are provided with and then fill out a "panellist EtD" table. In that table, and in lieu of systematically collected evidence, the panellists provide a description of their "expert evidence", consisting of their observations and experiences (equivalent to case reports and case series); these are expected to reflect "facts" (as opposed to opinions). In the next step, the steering group collates the input from all the panellists and populates the EtD that will be used as the basis for panel discussion and consensus building. Typically, the panel chair builds consensus with the panellists through discussion, and if needed through (iterative) voting. Alternatively, and for efficiency purposes,

	the chair can make suggestions and have panellists agree or disagree with them; however, caution is needed to ensure a broad range of perspectives is considered. ▪ Consider an updating plan o In the case of urgency related to an environmental exposure, there might be a need to continuously monitor the level of exposure within the population of interest for the purpose of triggering or updating the recommendations. For example, in the event of a nuclear incident, regular collection of environmental radioactive iodine levels from the field would be needed to reverse a recommendation for mass evacuation.
Notes	
Reviewer notes	RQ3: Innovation of criteria for developing trustworthy recommendations during urgent responses. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: EtD – evidence to decision; GRADE – Grading of Recommendations Assessment, Development and Evaluation; MERS – Middle East respiratory syndrome; N/A – not applicable; N/R – not reported; SARS – severe acute respiratory syndrome; WHO – World Health Organization.

Table C39 Trading certainty for speed - how much uncertainty are decision makers and guideline developers willing to accept when using rapid reviews: an international survey

Publication identification	
Authors (year)	Wagner et al. (2017)
Country	International
DOI	https://doi.org/10.1186/s12874-017-0406-5
Publication description	
Design	Descriptive cross-sectional study: An international web-based survey in English, German, and Spanish targeting decision makers and guideline developers who might commission evidence syntheses to inform their decisions and recommendations.
Objective	To determine how much incremental uncertainty about the correctness of an answer guideline developers and health policy decision makers are willing to accept in exchange for a rapid evidence synthesis.
Summary/Overview	The study explored the risk of getting an incorrect answer that healthcare decision-makers and guideline developers were willing to accept as a trade-off for using rapid reviews. The study found that the survey participants, on average, viewed 10% as the maximum tolerable risk of getting an incorrect answer from a rapid review. Respondents of the survey expected rapid reviews to provide answers similar to systematic reviews in at least nine out of ten cases.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	Use of rapid reviews for guideline development.
What are the core elements of the key innovation?	**The risk of getting an incorrect answer that healthcare decision makers and guideline developers are willing to accept as a trade-off for using rapid reviews.** Participants (n = 325) of the survey, on average, viewed 10% as the maximum tolerable risk of getting an incorrect answer from a rapid review. Survey respondents expected rapid reviews to provide answers similar to systematic reviews in at least nine out of ten cases. Across all three scenarios (clinical treatment scenario, public health scenario and clinical prevention scenario), the median acceptable incremental risk of getting an incorrect answer from a rapid review was 10% (interquartile range [IQR] 5.0–15.0). Individual responses for the three scenarios, however, varied widely and ranged from 0% (rapid reviews have to be as reliable as systematic reviews) to 50% (rapid reviews are still useful, even if they provide incorrect answers in 5 out of 10 reviews).

	Regarding individual scenarios, respondents' willingness to accept risks for incorrect results was the same for the clinical treatment (scenario 1, n=313) and the public health (scenario 2, n=320) scenarios with a median acceptable risk of 10% (IQR 5.0% to 15% for both). For the clinical prevention scenario (scenario 3, n=312), the median acceptable risk was 6.5% (IQR 5.0% to 10.5%). Across all three scenarios, male participants (n=158) were more cautious in their willingness to accept incorrect answers than female (n=165) participants (median 5% [IQR 5.0–12.0] vs. 10% [IQR 5.0–15.0]). The overall acceptable risk for getting an incorrect answer was, in general, similar across types of evidence users.
What is the rationale behind the methodology? OR What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	The increased uncertainty in rapid reviews that results from streamlining methods might affect estimates of probabilities of individual outcomes and cause ambiguities. Rapid reviews do not affect, however, uncertainty due to the complexity of a medical question or healthcare situation. When deciding whether or not to favour a rapid review as an alternative to a systematic review, decision makers have to weigh the celerity to complete a decision support review against the potential risk of incorrect answers for some outcomes of interest.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	Rapid reviews offer the potential to speed up the process of evidence synthesis for guideline development. The findings of this study suggest that decision makers and guideline developers are willing to accept some trade-off in validity in exchange for a rapid synthesis of the evidence. Nevertheless, they expect the validity of rapid reviews to come close to that of systematic reviews.
Notes	
Reviewer notes	RQ3: Innovation of rapid review to inform guideline development. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: IQR – interquartile range; N/A – not applicable; N/R – not reported.

Table C40 Developing WHO rapid advice guidelines in the setting of a public health emergency

Publication identification	
Authors (year)	Garritty et al (2017)
Country	Canada, Croatia, Switzerland
DOI	https://doi.org/10.1016/j.jclinepi.2016.08.010
Publication description	
Design	Mixed methods: Based upon an existing rapid review approach, which was modified to meet WHO needs and to allow integration with the organisation's existing approach to developing standard guidelines, in addition to discussions with WHO staff involved in emergency response.
Objective	To describe newly established guidance for guideline developers at the WHO on the process and procedures for developing a rapid advice guideline in the context of a public health emergency.
Summary/Overview	This paper describes the considerations that are relevant to deciding if a rapid advice guideline should be developed in the context of a public health emergency and outlines the processes and methods for developing such guidelines. The principles underlying rapid advice guidelines are the same as for standard guidelines.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	**Rapid advice guidelines** ▪ Purpose: to meet an emergent or urgent public health need when the short timeline mandates a modified process ▪ Scope: focused ▪ Developer: WHO technical staff ▪ New or existing recommendations: usually new; may contain existing recommendations if they have been evaluated and updated as appropriate ▪ Development period: usually one to three months.
What are the core elements of the key innovation?	**Steps in the development of rapid advice guidelines** **Phase 1 (Planning)** Primary contributor: Member State, WHO country office, or public/private entity 1. Request for guidance on a topic: ▪ The request is in the context of a public health emergency. Primary contributor: WHO technical unit 2. Determine if a guideline is needed, review existing WHO and external guidelines.

	▪ The technical unit must determine if a rapid advice guideline is needed or if a standard or interim guideline would be more appropriate. 3. Discuss the process with GRC Secretariat and with other WHO staff with experience of developing guidelines. ▪ The planned guideline is discussed with the Secretariat when it first becomes a possibility. 4. Form the Steering Group ▪ All relevant departments at WHO headquarters and in the regional offices must be involved. 5. Identify sufficient resources. 6. Determine the timeline. Primary contributor: Steering Group 7. Draft the scope of the guideline. ▪ The literature is scoped through a brief review. The guideline's scope must be narrow and feasible. 8. Begin preparing the planning proposal. 9. Identify potential members of the GDG and the chair ▪ Issue invitations early; involve the GDG in determining the scope and key questions. 10. Obtain DOIs and manage any COIs among potential GDG members. ▪ The process for rapid advice guidelines and standard guidelines is identical. Primary contributor: Steering Group and the GDG 11. Formulate key questions in PICO format. ▪ Key questions (in PICO format) include only those of the highest priority and must be focused and narrow. Background questions are not addressed in a rapid advice guideline. Primary contributor: WHO Steering Group 12. Finalise the guideline planning proposal ▪ The process is the same as for a standard guideline. Primary contributor: GRC 13. Review and approve the planning proposal ▪ The GRC uses an accelerated process for review and disposition. **Phase 2: Development** Primary contributor: SR team 14. Perform SRs of the evidence for each key question with the potential of abbreviating the SR process (i.e., perform a RR) ▪ The contractor needs to be identified from the outset and involved in the scoping and development of key questions: they can advise on what is feasible in the given time frame. 15. Evaluate evidence quality for each important outcome, using GRADE as appropriate. ▪ The process is the same as for a standard guideline. Primary contributor: Steering Group 16. Convene a meeting of the GDG ▪ Meeting place and participants need to be identified at the beginning of the development process. The meeting has a similar format and agenda as for the development of a standard guideline. Primary contributor: GDG 17. Formulate recommendations using the GRADE framework

	▪ The general methods are the same as for a standard guideline. The evidence may be sparse, so other factors that inform the recommendations must be transparent and based on indirect evidence when possible, and on equity, human rights and gender considerations. Primary contributor: Steering Group 18. Draft the guideline document ▪ The document should be concise and tailored to the end user. Primary contributor: External review group 19. Conduct targeted external peer review ▪ External peer review is recommended for rapid advice guidelines but may not be feasible in some situations. **Phase 3: Publishing and updating** Primary contributor: Steering Group and editors 20. Finalise the guideline document. Perform copy editing and technical editing. Submit the final guideline to the GRC for review and approval ▪ This step will have to be performed in an accelerated manner. Editorial staff need to be identified early in the process. Primary contributor: WHO Guidelines Review Committee 21. Review and approve the final guideline ▪ The GRC uses an accelerated process for review and disposition. Primary contributor: Steering Group and editors 22. Finalise the layout. Proofread ▪ This step needs to be accelerated and perhaps abbreviated from the standard processes. 23. Publish (online and in print, as appropriate) Primary contributor: WHO technical unit and program manager 24. Disseminate, adapt, implement, evaluate Primary contributor: WHO technical unit 25. Update ▪ From the outset, the technical unit must consider the likely shelf life of the rapid advice guideline and whether a standard guideline will follow and when. **Approaches to the rapid review** ▪ Approach 1: If only one exceptional, high quality systematic review is identified, summarise the findings ▪ Approach 2: If multiple, high quality systematic reviews are identified, assess the rigour of each systematic review, any overlap among included studies and the comparability of the findings ▪ Approach 3: Update an existing, high quality systematic review ▪ Approach 4: Include primary studies that have strong(er) study designs and are of high quality (if available).
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	If a public health emergency continues and as response efforts evolve into recovery and rebuilding, guidelines are needed that are developed using more rigorous methods and generally with a somewhat longer timeline: perhaps one to three months. These are termed as “rapid advice guidelines”.

What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	An example of the development of a rapid advice guideline: Personal protective equipment in the context of filovirus disease outbreak response: Rapid advice guideline (2014). World Health Organization. Context ▪ A public health emergency of international concern. Issue ▪ Healthcare workers caring for individuals with Ebola were at an increased risk of contracting Ebola virus disease during the outbreak in West Africa starting in 2013. ▪ There was uncertainty in the field as to the most effective types of personal protective equipment. Development of a WHO rapid advice guideline ▪ A RR was conducted over 7 weeks to inform the recommendations. ▪ Initially, the RR focused on the comparative effectiveness and disadvantages of personal protective equipment (gloves, gowns, and face protection) for healthcare workers working with Ebola patients. However, only non-comparative studies were identified. ▪ Concurrent with the RR, a survey of values and preferences was administered to expatriated healthcare workers over a 3-week period, which helped to inform recommendations. ▪ The non-comparative data from the RR, the survey data, and information from experts in virology and blood-borne pathogens and materials science formed the basis for the recommendations, which were formulated at an expert meeting. Significance ▪ Produced over a 12-week time frame, this marked the first rapid advice guideline produced by WHO following the approaches outlined in this paper.
Notes	
Reviewer notes	RQ3: Innovation of rapid guideline development. Category of evidence: Grade C.
Associated handbook(s)	WHO handbook for guideline development, 2nd Edition.

Key: COI – conflict of interest; GDG – guideline development group; GRADE - Grading of Recommendations, Assessment, Development, and Evaluation; GRC – guidelines review committee; N/A – not applicable; N/R – not reported; PICO – Population/patient, Intervention, Comparison, Outcome; RR – rapid review; SR – systematic review; WHO – World Health Organization**.**

Table C41 Designing a rapid response program to support evidence-informed decision-making in the Americas region: using the best available evidence and case studies

Publication identification	
Authors (year)	Haby et al. (2016)
Country	International
DOI	https://doi.org/10.1186/s13012-016-0472-9
Publication description	
Design	Mixed methods: rapid reviews, literature (narrative) reviews (objective 1 and 3), case studies (objective 3).
Objective	To inform the design of a rapid response programme to support evidence-informed decision-making in the Americas region. 1. What are the best methodological approaches for rapid reviews of the research evidence (the product)? 2. What other strategies are needed to facilitate evidence-informed decision-making in health policy and practice? 3. How best to operationalise the programme?
Summary/Overview	The study describes the design of a rapid response programme for health policy and practice that can be applied in the Americas region. A rapid response programme provides rapid reviews of the results of high-quality research evidence and contextualises and targets it to the needs of decision-makers. It has a fast turn-around time. Additionally, the study identified shortcuts that could be considered to reduce the time needed to complete the reviews.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	A rapid response programme that provided rapid reviews of the results of high-quality research evidence, which was contextualised and targeted to the needs of decision-makers, with a fast turn-around time, and that included interaction between researchers and decision-makers, could overcome some of the barriers and facilitate the uptake of research into policy and practice.
What are the core elements of the key innovation?	**Methodological approach** Systematic reviews included in the review of methodologies of rapid reviews and the AMSTAR questions were used to make a list of all possible areas. To determine the potential impact of the shortcut on the validity of the results (see Table C41a), primary studies included in reviews by Ganann et al.[3] and Cameron et al.[4] were used and were complemented with studies cited in the Cochrane Handbook or found during the search process for the rapid review. It is important to note that this was not a systematic search for evidence, though the majority of the references came from the systematic review by Ganann et al.

Table C41a identified shortcuts that could be considered to reduce the time needed to complete the reviews

Systematic review step	Possible 'shortcuts '	Potential impact on the validity of the results	Relevant AMSTAR question and potential impact of shortcut on AMSTAR score
Preparation of a Protocol	▪ Omit protocol	Unknown	Q1. Loss of one point if a protocol is not prepared and/or not mentioned in report
Question formulation	▪ Limit the number of questions and ▪ sub-questions	None expected	
Selecting relevant studies	▪ One reviewer screens titles and abstracts ▪ One reviewer screens full text	Unknown, though one reviewer could miss up to 9 % of eligible randomized controlled trials	Q2. Loss of one point if only one reviewer does screening and/or only one reviewer does data extraction
Data extraction	▪ One reviewer extracts data	Can increase the number of errors but the impact on results is not known	
	▪ One reviewer extracts data with checking by a second reviewer	Unknown	
	▪ Data extraction limited to key characteristics, results, conflicts of interest	Unknown	
Literature search	▪ Limit number of databases searched ▪ Limit or omit hand searching of references lists and relevant journals ▪ Eliminate consultation with experts to find additional studies	Limiting the number of databases searched can increase efficiency without compromising validity, especially if combined with some hand searching and contact with experts	Q3. Loss of one point if less than two databases searched and/or no supplementary strategies
Inclusion criteria			
Grey literature	▪ Limit or omit grey literature	Could introduce publication bias but the evidence is mixed	Q4. Loss of one point if grey literature omitted
Language	▪ English only	Effect can vary depending on the question	

Dates	▪ Narrow time frame, e.g., last 5 or 10 years	None expected		
Study types	▪ Restrict study types to systematic reviews (and economic evaluations) ▪ Restrict study types to randomised controlled trials or controlled clinical trials (and economic evaluations)	None expected		
Quality assessment	▪ Limit or omit quality assessment	Not recommended. Several authors suggest that, where resources are limited, priority should be given to quality assessment rather than extensive searching	Q7 and Q8. Loss of two points if not assessed, documented and used in formulation of conclusions	
	▪ Omit “a priori” specification ▪ Done by one reviewer	Unknown		
Data synthesis	▪ Narrative synthesis only (no meta-analysis)	Unknown – meta-analysis can increase power and precision but also has potential to mislead if not applied appropriately and done correctly	Q9. None if explained that meta-analysis not possible due to heterogeneity. If not, loss of one point	
Assessment of publication bias	▪ Omit	Unknown	Q10. Loss of one point if omitted	
Assessment of conflict of interest	▪ Omitted for individual studies and or for systematic review	Unknown	Q11. Loss of one point if omitted	
Report	▪ Information included limited	Unknown but can impact on AMSTAR score if insufficient detail of methods provided to enable a quality assessment. Sufficient detail of methods will help the reviewer to assess the validity of the results	Q1–11. Potential large loss of points if key AMSTAR questions not covered	

	<table><tr><td>External peer review</td><td>▪ Omit or limit</td><td>Unknown</td><td></td></tr></table>
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	While research evidence is only one of many inputs into decision-making when it comes to health policy, it is important to try to maximise its usefulness and uptake. A range of programmes and efforts already exist to promote the uptake of research evidence into policy and practice. These include efforts to conduct systematic reviews of the evidence (e.g., the Cochrane Collaboration) as well as efforts to package research evidence, including systematic reviews to inform policy and practice (e.g., Evidence-Informed Policy Network, Health Technology Assessment agencies). However, while extremely useful, these programmes are often not able to provide access to research quickly nor answer specific policy questions in a timely way. There is some evidence from a good quality randomised controlled trial with low risk of bias that rapid reviews may improve clarity and accessibility of research evidence for decision-makers when compared to a systematic review alone.[5]
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	Four case studies of rapid response models are reported: 1. Cochrane Response Rapid Reviews by Cochrane Innovations started in 2013 with potentially global reach. 'User-pays' model for funding reviews. Approximately eight weeks to complete a rapid review. Rapid reviews are reviewed externally and made publicly available with a lag period before publication. 2. McMaster Health Forum Rapid Response Programme – McMaster University started in 2012 with potentially national (Canada) reach. Funded by Ontario government, with a 'user-pays' model for funding reviews outside of Ontario. Maximum of eight weeks to complete a rapid review. Rapid reviews are reviewed externally and made publicly available with a lag period before publication. 3. Sax Institute Evidence Check programme started in 2006 with state-wide and potentially national (Australia) reach. 'User-pays' model for funding reviews. Approximately 12-16 weeks to complete a rapid review. Rapid reviews are not reviewed externally, most made publicly available (some kept confidential if requested by funder) with a lag period before publication. 4. REACH Policy Initiative, Uganda started in 2010 with national (Uganda) reach. 'User-pays' model for funding reviews. Maximum of four weeks to complete a rapid review. Rapid reviews are not reviewed externally. Whether rapid reviews are made publicly available or if there is a lag period before publication is not reported
Notes	
Reviewer notes	RQ3: innovation of rapid response programme. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: AMSTAR – A MeaSurement Tool to Assess systematic Reviews; N/A – not applicable; N/R – not reported; PICO – Population/patient, Intervention, Comparison, Outcomes; REACH – Regional East African Community Health.

Table C42 Living systematic reviews: 4. Living guideline recommendations

Publication identification	
Authors (year)	Akl et al. (2017b)
Country	International
DOI	https://doi.org/10.1016/j.jclinepi.2017.08.009
Publication description	
Design	Narrative review
Objective	To describe the concept of living practice guidelines and living recommendations, the workflows required to support them, the collaboration between living systematic reviews and living guideline teams, the thresholds for changing recommendations, and potential approaches to publication and dissemination.
Summary/Overview	This paper provides information on living systematic reviews and living guideline recommendations by the Living Systematic Review Network. It explains the concept of living guidelines and how they can provide timely and up-to-date guidance to users. The document also discusses the essential elements required for producing living recommendations and criteria for deciding when a living guideline is appropriate.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	Living practice guidelines and living recommendations ▪ **Living practice guideline**: an optimisation of the guideline development process to allow updating of individual recommendations as soon as relevant new evidence becomes available. ▪ **Living recommendation:** a recommendation that is updated as soon as relevant new evidence becomes available. One implication of the definition of the living practice guideline is that the unit of update becomes the individual recommendation and not the whole guideline, which is likely to make the process less onerous. The second implication is that the living recommendation should build on an existing, high-quality recommendation developed using the standard guideline development methods
What are the core elements of the key innovation?	**Elements necessary for producing living recommendations** **1. Living systematic review** ○ A living systematic review is a process that uses continual surveillance of the literature to allow updating of a systematic review with new evidence as it becomes available. **2. Living summary tables**

	○ Guideline panels rely on standardised summary tables to make the judgments required for developing recommendations. Updating these tables as soon as new evidence emerges is needed to swiftly relay the findings of the living systematic reviews to guideline panels in a form that allows them to reconsider the recommendation ▪ **Living Evidence Profile:** Evidence profile provides the statistical information on the effects on health benefits and harms of the alternative interventions for each outcome of interest, as well as a detailed assessment of the certainty of supporting evidence ▪ **Living EtD table:** Living EtD framework provides the information on the factors needed to judge the strength and direction of each recommendation. These factors include the health effects of interventions, the certainty of evidence, resource use, impact on equity, and acceptability of the intervention, among others. **3. Living guideline panel** ○ A living guideline process can circumvent the challenges associated with recruiting a guideline panel by recruiting ahead of time panel members committed to making themselves available within very short notice, whenever the updating process is triggered. This implies that virtual meetings would be more feasible than in-person meetings. Also, a living guideline panel would likely be engaged for more frequent periods of times, compared with a standard guideline panel. It is also likely that the membership of the panel will change over time. This raises the challenge of preserving the ''institutional memory'' of the group. **4. Living peer review process** ○ A living guideline process could recruit a larger number of reviewers than needed and ensure their commitment to a timely review. The guideline developer could minimise the involvement of the reviewers and provide them with advance notice (e.g., as soon as the updating process is triggered). Also, any internal (e.g., professional society) or external (e.g., governmental) review and approval processes or periods of public comment need to be carefully planned and weaved into the guideline timeline.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	Rationale: Given a guideline is based on a set of systematic reviews, some of its recommendations would be out of date by the time of the guideline publication. This problem is accentuated when factoring in the additional amount of time needed to move from the evidence to the recommendations and to publish the guidelines. With any frequency of update of the whole guidelines, there will be recommendations that go out of date prior to the update (decreasing validity). There will also be recommendations that are revisited unnecessarily as they are still up to date (decreasing efficiency). Conducting more frequent updates of the whole guideline will enhance validity but negatively impact efficiency. Conducting less frequent updates will enhance efficiency but negatively impact validity. The authors hypothesised that the process for living guidelines will lead to more valid recommendations, while potentially improving efficiency of guideline development. Criteria proposed to determine if a living guideline is appropriate: The prioritisation criteria for living recommendations include:

	1. The recommendation is a priority for decision making. This could be affected by a high prevalence of the condition, high rates of associated morbidity and mortality, known variation in practice, and interest in emerging interventions and diagnostic tools. 2. There is a reasonable chance that the existing recommendation changes with the emergence of new evidence. This is the case when the strength of the existing recommendation is conditional (as opposed to being strong) due to the evidence not being of high certainty. This is a scenario where the consideration of emerging evidence may increase the certainty of evidence, subsequently enhancing the strength of the recommendation. 3. There is likely to be new research evidence. Active research addressing the recommendation of interest is ongoing with a reasonable likelihood of findings being published over the period when the living guideline process is implemented. When at least one of the above listed criteria for switching a recommendation into a living status cease to apply, the authors could switch back from the living guideline process to the standard guideline development methods, for individual recommendations or entire guidelines.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	The process of developing living guidelines typically starts by developing a ‘‘base recommendation’’ using standard guideline development methods. In parallel, a literature surveillance process is put in place. The identification of new evidence would trigger the process for the living systematic review and its Evidence Profile, for the living EtD table, and subsequently for the living recommendation.
Notes	
Reviewer notes	RQ3: Innovation of living guidelines. Category of evidence: Grade D.
Associated handbook(s)	N/R

Key: EtD – evidence to decision; N/A – not applicable; N/R – not reported.

Technological innovations (not evaluated)

Table C43 Transforming clinical practice guidelines and clinical pathways into fast-and-frugal decision trees to improve clinical care strategies

Publication identification	
Authors (year)	Djulbegovic et al. (2018)
Country	USA
DOI	https://doi.org/10.1111/jep.12895
Publication description	
Design	Narrative review
Objective	To describe the process of converting clinical practice guidelines and clinical pathways into fast-and-frugal heuristics to provide a simple, transparent, and robust methodological framework to connect decision science to clinical care.
Summary/Overview	This paper describes the process of converting clinical practice guidelines/clinical pathways into fast-and-frugal decision trees (FFTs); this provides a simple and transparent, yet solid and robust, methodological framework connecting decision science to clinical care. The authors recommend that all guideline panels and clinical pathway developers express their recommendations as flow-charts or clinical algorithms, which in turn should be converted into FFTs to guide clinical care.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	**Supplementing clinical pathways with a firm theoretical decision framework such as FFTs heuristics** FFTs are highly effective, simple decision trees composed of sequentially ordered cues (tests) and binary (yes/no) decisions formulated via a series of if-then statements. FFTs represent a particularly effective class of heuristic strategies, which rely on limited information to reduce estimation error and facilitate fast decisions. The use of FFTs provides practical implementation of the satisficing principle—there must exist a point (threshold) at which obtaining more information or performing another computation becomes detrimental and costly; the application of FFT heuristics helps decision-makers stop searching before this threshold has been crossed. FFTs are very efficient for solving binary decision tasks such as making a diagnosis, prediction, or deciding whether to order tests or initiate treatment. Decision-making strategies based on FFTs have been found to be superior to other strategies, including those using complex multivariate regression models. FFTs provide a potentially fundamental link between evidence and action. FFT has an in-built theoretical structure that allows quantitative analysis of the accuracy of clinical management strategies. This is possible because the FFT heuristic strategy of decision-making can be conceptually linked to signal detection theory, evidence accumulation theory, and the threshold model to improve decision-making. Thus, metrics of signal detection theory (true and false positives, true and false negatives), the effects of sequential accumulation of

	evidence and the consequences of actions expressed via the evidence accumulation theory and threshold model, respectively, become metrics that can be used by FFTs to calculate the accuracy of clinical decisions.
What are the core elements of the key innovation?	In an FFT, clinical information ("cues") for a series of binary decisions (yes/no) are assembled. The relation among the cues is framed as a series of if-then statements (e.g., if risk for a cardiovascular event is high [cue], then administer statins [decision]). If the condition is met, the decision is made and the FFT is exited. Otherwise, the FFT sequentially considers additional cues until the exit condition of a cue is met. If the exit occurs after a positive cue ("yes"), an intervention is enacted. If the exit occurs after a negative cue ("no"), no intervention is administered. The last cue of an FFT has two exits, to avoid indefinite loops, and to ensure that a decision is ultimately made. An essential feature of FFT is that the structure of the exits from the cues (yes/no) determines the ratio between false negatives vs. false positives. Thus, FFTs where all cues in a fixed order have all their exits on the "yes" side of each cue has a high true positive rate at the expense of a large number of false alarms. In contrast, FFTs where all cues in a fixed order have all their exits on the 'no' side of each cue reduces false alarms at the expense of large false negative rates.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	Clinical recommendations derived from clinical pathways are typically ad hoc and developed by experts in a theory-free environment. As any recommendation can be right (true positive or negative) or wrong (false positive or negative), the lack of theoretical structure precludes the quantitative assessment of the management strategies recommended by clinical practice guidelines/clinical pathways. To realise the full potential of clinical practice guidelines/clinical pathways, they need to be placed on more solid theoretical grounds. This can be best realised by converting clinical practice guidelines/clinical pathways within the heuristic theory of decision-making, often implemented as fast-and-frugal decision trees.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	Fast-and-frugal decision trees allow an assessment of the accuracy for both the individual patients who meet the given decision criteria, and the evaluation of the entire management strategies, at the population level. For example, the conversion of the American College of Cardiology and American Heart Association statins guidelines into an FFT with five cues shows the performance of the FFT at individual patient levels with the answer "yes" to all cues. The overall accuracy of this FFT expressed as a positive predictive value = 11.4%, indicating that 11.4% of treated patients will be appropriately given statins, but 88.6% will not, suggesting there is a better FFT sequence. One of the powerful features of FFT methodology is that it allows evaluation of the performance of all possible clinical strategies by changing the orders in which available clinical information (cues) are collected (and acted upon). The FFT can demonstrate performance characteristics for all 1,920 combinations that can be generated with five cues. The analysis can help identify the most sensitive and specific clinical management strategies; in addition, if data on the benefits and harms of statins are used, the various consequences of actions can be incorporated into the analysis to generate the FFT with the most optimal trade-offs between true positives and false positives.
Notes	
Reviewer notes	RQ3: innovation of fast and frugal heuristics that employ minimum of time, knowledge, and computation to make adaptive choices in clinical care. Category of evidence: Grade D.
Associated handbook(s)	N/R

Key: FFT – fast-and-frugal tree; N/A – not applicable; N/R – not reported.

Table C44 Development of health pathways to standardise cancer care pathways informed by patient-reported outcomes and clinical practice guidelines

Publication identification	
Authors (year)	Girgis et al. (2018)
Country	Australia and UK
DOI	https://doi.org/10.1200/CCI.18.00024
Publication description	
Design	Mixed methods: Algorithm development informed by a literature review and a consensus decision-making approach.
Objective	To describe the development of an evidence-based automated decisional algorithm for patients with cancer that had specific, actionable, clinical, evidence-based recommendations to improve patient care, communication and management.
Summary/Overview	This study describes the development of an evidence-based automated decisional algorithm for patients with cancer. Using automated algorithms and clinical recommendations provides a platform for streamlining and systematising the use of PROs to inform risk-stratified guideline-informed care.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	Use of PROs to inform clinical pathways through use of an automated decisional algorithm.
What are the core elements of the key innovation?	The algorithm comprised three steps: 1. Evaluating/identifying issues ○ Patient distress, symptoms, unmet needs, and quality of life were the areas of clinical care that were initially targeted for the development of the algorithms to enhance patient care. 2. Mapping to domains of care ○ On the basis of data extracted from the literature and existing clinical practice guidelines, items were mapped and grouped into five domains of patient well-being: physical, emotional, social and family, practical support, and maintaining well-being. 3. Clinical recommendations ○ Recommendations were mapped to the following four main categories: (1) Consider reasons for concern and, if required, refer to (types of specialties indicated here, depending on issue) for further assessment and care; (2) Clinically address as appropriate or refer to (types of specialties indicated here, depending on issue) for further assessment and care; (3) Address (type of) needs and identify appropriate sources of support and information; and (4) No action required.

What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	Collecting PROs assessments can provide health service professionals with tailored actionable information. However, for them to be effective, they need to be integrated into the work flow and available at the point of care, and the pathways developed need to be evidence-based. The use of computing devices and automated data collection in clinical settings enables rapid execution of automated algorithms. The algorithms can incorporate a variety of patient factors into patient-specific recommendations, thereby standardising care in an effort to optimise patient care.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	▪ eHealth systems, such as PROMPT-Care, provide a platform for streamlining and systematising the use of PROs to inform risk-stratified guidelines for healthcare within time-limited clinical consultations. The series of algorithms that set out systematised care pathways for the clinical care of patients with cancer could potentially inform patient self-management recommendations. ▪ The process of computing algorithms and generating the clinical recommendations is fully automated and embedded with the PROMPT-Care eHealth system. The PROMPT-Care system helps patients to complete PRO assessments online by using an electronic device (e.g., tablet, smartphone, or computer) from any location (at home or in the clinic) and automatically transfers data into the patient's electronic medical record within the point-of-care oncology information system in real time. ▪ By using the algorithms, the patient PROs and clinical recommendations are automatically mapped and summarised in clinical feedback reports. These reports remain within the patient electronic medical record and are available in real time for clinical staff to review in clinic and take action upon as necessary. ▪ Clinical staff will also receive e-mail alerts in which the algorithms identify patients who breach the same items on two consecutive assessments, prompting the clinical care team to review the clinical feedback reports outside the clinic consultations. ▪ The algorithms can also be used to facilitate patient-self management, with patients receiving targeted links to evidence-based information resources related to the issues identified in the PRO assessment.
Notes	
Reviewer notes	RQ3: Development of an automated decisional algorithm that had specific, actionable, clinical, evidence-based recommendations to improve patient care, communication, and management. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: PRO – patient reported outcome; PROMPT-care – Patient Reported Outcome Measures for Personalised Treatment and Care; N/A – not applicable; N/R – not reported.

Table C45 Design and execution of integrated clinical pathway: A simplified meta-model and associated methodology

Publication identification	
Authors (year)	Ardito et al. (2020)
Country	Italy
DOI	10.3390/info11070362
Publication description	
Design	Descriptive cross sectional study: Description of technology/framework
Objective	To provide a method for obtaining machine executable ICP.
Summary/Overview	The study describes an innovation for integrated clinical pathway (ICP) that utilises a chatbot engine that acts as a virtual assistant and gathers patients' health data. The chatbot follows the logic of the process by asking the patient to select one or more items among the set of closed answers encoded in the meta-model.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	A simplified meta-model for process information management that is compliant with most general purpose description languages and ICP process execution assisted by a chatbot.
What are the core elements of the key innovation?	▪ The chatbot's role is to lead the patient to provide information useful for determining a change in state in the clinical process instance. To conduct the dialogue with the patient, the chatbot uses some sentences contained in the process template. ▪ The proposed process execution strategy is to allow a chatbot to access the topological structure of the template process, stored into the meta-model, and use it to start a conversation with the user that aims to advance the status of the process instance associated with the user. ▪ The chatbot uses a PROCESS_NODE attribute called patient description, which contains a description of the task to be performed. In this case, the task can be one of the following: o Informative task: A task that aims to provide information and does not require the production of specific results. The chatbot sends the information message stored in patient description and then moves the execution token to the next task; o Processing task: In this case, the task requires the production of specific outputs. The chatbot guides the user to produce the expected results. When all the results have been produced, the chatbot independently moves the execution token to the next task. ▪ The chatbot uses the PROCESS_NODE patient description attribute in order to express the conditional situation to the user (e.g., ask a question); at the same time, it proposes the possible answers in exclusive mode using the patient description attribute inserted in its two child nodes and leaving the user the possibility to indicate the

	right branch. After the choice, the chatbot will move the execution token to the next node of the one representing the branch related to the choice made. ▪ This is the same behaviour as in the condition case, but the user can activate multiple child branches of the PROCESS_NODE that expresses the condition. The chatbot will then move the execution token to the child nodes of all the selected branches. ▪ The chatbot will automatically move the execution token to all the child nodes of the parallel PROCESS_NODE.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	To date, research has focused more on finding solutions to the machine-executability issue of the ICP process, acting on the design language. This approach has resulted in an increase in the complexity of the process graphical representation, which, in turn, has become an obstacle to the adoption of the proposed solutions. In this paper, the authors show that it is possible to explore new possibilities by acting on the representation of the process knowledge, independently from the language adopted for the process design.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	**Methodology for executable clinical pathway generation** 1. **ICP Formalisation:** This phase is typically carried out by an interdisciplinary medical team involving case managers. The aim is to design a clinical pathway process in which medical guidelines and evidence converge. 2. **ICP Normalisation:** At this stage, the clinical team is not constrained to using a specific formalism for the process descriptive diagram. If the clinical team does not used a standard formalism, some flow ambiguities might occur. In this case, a disambiguation activity is carried out. 3. **ICP4Patient generation:** The third step addresses the issue of the centrality of the patient with respect to the treatment pathways. A clinical process designed to be read and interpreted exclusively by doctors is not understandable by the patients to whom it is addressed because the medical language is too technical. It is therefore necessary to proceed to a transformation of the text that describes each task in a language that is comprehensible by patients. This transformation could be performed by the Case Manager or the most suitable professional figure in this critical phase. 4. **ICP-Graph Generation:** At this point, it is possible to generate a graph related to the template process by feeding the Process Design Data that will be the container of all the clinical pathways templates managed by the ICP-Graph Generation platform. 5. **User-ICP linking:** Once the template process is inserted into the meta-model, it will be possible to link patients to the template process by feeding the data of the process instance. 6. From this moment on, the General Practitioner invites the patient to interact with the chatbot. The status of the process is then updated by the chatbot by triggering a continuous conversation with the patient using pre-coded sentences in the template process. At the same time, the medical staff can monitor the status of the process, acquire knowledge for each patient, and take actions in the case of need or upon chatbot request.
Notes	
Reviewer notes	RQ3: Innovation of integrated clinical pathway that utilises a chatbot engine that acts as a virtual assistant and gathers patients' health data. Category of evidence: Grade D.

Associated handbook(s)	N/R

Key: ICP – integrated clinical pathway; N/A – not applicable; N/R – not reported.

Table C46 Comprehensive mitigation framework for concurrent application of multiple clinical practice guidelines

Publication identification	
Authors (year)	Wilk et al. (2017)
Country	Poland, Canada, USA
DOI	https://doi.org/10.1016/j.jbi.2016.12.002
Publication description	
Design	Descriptive cross sectional study: Framework description
Objective	To describe a comprehensive framework based on first-order logic (FOL) for mitigating (identifying and addressing) interactions between multiple clinical practice guidelines applied to a multi-morbid patient. FOL is a widely used formal system for representing and reasoning about knowledge, expressed in domain-specific FOL language.
Summary/Overview	The paper describes a mitigation framework that provides a formal representation for the clinical practice guideline (considered as primary medical knowledge) and the secondary medical knowledge related to adverse interactions and patient preferences, and a mitigation algorithm to identify and address adverse interactions for a multi-morbid patient. The framework handles discordant morbidities and incorporates patient preferences. Specifically, it customises clinical practice guidelines represented as actionable graphs to a specific patient by revising them using revision operators to address adverse interactions and unmet patient preferences. Finally, it creates a management scenario that includes activities that are clinically appropriate and preferred by the patient.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	The framework mitigates adverse interactions between multiple clinical practice guidelines applied to a multi-morbid patient. It introduces a mitigation-oriented clinical practice guideline representation that on one hand captures the complex characteristics of clinical practice guidelines (e.g., parallel activities and their temporal attributes), and on the other hand facilitates their processing by the mitigation algorithm. Moreover, the framework allows for the specification of patient preferences and considers them when addressing encountered interactions and revising the applied clinical practice guidelines.
What are the core elements of the key innovation?	The main components of the mitigation framework include: ▪ Actionable graphs ○ An actionable graph represents a clinical practice guideline for managing disease (here the notion of a disease is extended to include conditions such as trauma). ▪ Mitigation-specific FOL language ○ To develop a FOL-based description of a mitigation problem, a mitigation-specific FOL language was introduced.

	▪ Combined mitigation theory o A combined mitigation theory brings together the core components of a mitigation problem. ▪ Translating actionable graphs to FOL ▪ Revision operators o Revision operators encode two types of secondary medical knowledge: ▪ Knowledge required to mitigate adverse interactions resulting from the concurrent application of clinical practice guidelines ▪ Knowledge about patient preferences. o Therefore, two classes of revision are considered — interaction-related and preference related. ▪ Interaction-related revision operators define adverse interactions between two or more actionable graphs due to discordant morbidities and specify revisions necessary to mitigate these interactions. ▪ Preference-related revision operators describe clinical circumstances (e.g., a sequence of actions) that are not consistent with patient preferences and describe revisions that modify actionable graphs according to these preferences. ▪ Management scenario o A management scenario is a theory expressed in FOL language that represents a safe (free of adverse interactions) and preferred (consistent with patient preferences, whenever possible) course of actions and decisions for a specific patient. The management scenario highlights the current and future clinical actions and their characteristics (action and dosage predicates) and future clinical decisions (decision predicates) along with their timing (start time, and duration predicates), and includes the assumptions made about the patient's future state (result predicates). ▪ Mitigation algorithm ▪ It consists of three procedures: customise, mitigate, and revise. o Customise procedure: The purpose of the customise procedure is to revise a combined theory according to patient preferences (if any) while ensuring that all adverse interactions have been mitigated, and to return a management scenario. o Mitigate procedure: The goal of the mitigate procedure is to identify those adverse interactions encountered in a combined mitigation theory that are described by interaction-related revision operators, to revise the combination theory by applying operations described by appropriate operators, and to create a management scenario. In summary, this paper proposed a comprehensive FOL-based mitigation framework for the concurrent application of multiple clinical practice guidelines to a multi-morbid patient. The framework handles discordant morbidities and incorporates patient preferences. Specifically, it customises clinical practice guidelines represented as actionable graphs to a specific patient by revising them using revision operators to address adverse interactions and unmet patient preferences. Finally, it creates a management scenario that includes activities that are clinically appropriate and preferred by the patient. As such, the framework can be used as a simulation and verification tool supporting a physician in planning a complex therapy for a multi-morbid patient.

What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	Almost all of the research on guidelines is focused on developing and executing clinical practice guidelines designed for the management of a single clinical condition, while many patients do not have a single clinical condition. Developing clinical practice guidelines for patients with multi-morbidity is difficult because most of the clinical studies used for guideline development exclude such patients from clinical trials. Hence, the authors proposed a mitigation framework addressing adverse interactions resulting from conflicting therapies, while also considering patient preferences related to the prescribed treatment.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	The high-level implementation involves two architectural components — the mitigation CDSS and the FOL-based knowledge repository. The mitigation CDSS is responsible for computations, i.e., processing of combined mitigation theories, and for interfacing clinical users (physicians) and an electronic patient record. Its mitigation subsystem executes the mitigation algorithm and invokes an external FOL solver for theorem proving and model finding. Moreover, the run-time interface subsystem has a data capture facility for entering patient information and for importing it from the EPR through standardised protocols. The run-time interface subsystem reports detailed mitigation results (management scenario, applied revision operators). For better comprehensibility, these results would be translated into a narrative, textual format, so the clinical user is fully insulated from the underlying FOL-based representation. The FOL-based knowledge repository stores the knowledge used by the mitigation framework. Specifically, it is divided into three sub-repositories with actionable graphs, revision operators and patient information. The first two sub-repositories need to be prepared in advance and preloaded with the encoded primary and secondary medical knowledge. This process is conducted by a knowledge engineer who works with a clinical expert to elicit knowledge from textual clinical practice guidelines, medical literature and the expert herself, and to encode it in FOL. While knowledge elicitation and encoding requires significant effort and time, it needs to be conducted before system deployment and repeated only when significant updates are necessary (e.g., new clinical practice guidelines are published). The sub-repository with patient information is accessed by the mitigation CDSS during regular execution that automatically retrieves and updates its content. The knowledge engineer may also be consulted by the physician during regular operation in order to encode patient preferences into preference-related revision operators. Again, this should happen infrequently, for example, when the management of a specific patient starts (or when her preferences have changed), and then the mitigation CDSS will use the defined operators. Thus, although the interventions by the knowledge engineer are crucial for the proper operation of the mitigation CDSS, they should be relatively infrequent and for the most part the CDSS will be interacting with physicians and automatically performing computations based on provided information.
Notes	
Reviewer notes	RQ3: Innovation of framework for the concurrent application of multiple clinical practice guidelines to a multi-morbid patient. Category of evidence: Grade D.
Associated handbook(s)	

Key: CDSS – clinical decision support system; EPR – electronic patient record; FOL – first-order logic; N/A – not applicable; N/R – not reported.

Table C47 STANDING Collaboration: a study protocol for developing clinical standards

Publication identification	
Authors (year)	Wiles et al. (2016)
Country	Australia
DOI	http://dx.doi.org/10.1136/bmjopen-2016-014048
Publication description	
Design	Protocol: description of modified e-Delphi process
Objective	To provide proof of concept for an alternative method for creating sets of nationally-agreed evidence-based standards and clinical indicators and obtain consensus on 'appropriate care' for a range of common medical conditions.
Summary/Overview	This study describes the STANDING Collaboration — a method to develop an inclusive, transparent, collaborative process, which allows HCPs and patients or consumers to develop and keep up-to-date clinical standards comprising indicators with defined attributes, using an online curated wiki-based platform to facilitate ongoing review and updating of the standard, or individual indicators, as soon as new evidence emerges. It uses a three-phase approach: engage relevant stakeholders, develop clinical indicators representative of 'appropriate care' (which constitute the standard) for a range of common conditions, and evaluate the processes, products and feasibility.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	STANDING Collaboration: A transparent and inclusive online curated (purpose-designed, custom-built, wiki-type) system that uses an ongoing and iterative documentation process to facilitate synthesis of up-to-date information and make available its provenance.
What are the core elements of the key innovation?	The STANDING Collaboration methodology comprises of three approaches: ▪ Stakeholder analysis ▪ Development and test of a process for creating clinical indicators representative of 'appropriate care' for a range of common conditions ▪ Evaluation of processes, products and feasibility. Through an online curated wiki-based platform to facilitate ongoing review and updating of the standard, or individual indicators, as soon as new evidence emerges. In this study, the term 'wiki' refers to an interactive information management system which will allow users (e.g. healthcare professionals and patients) to collaborate directly in formulating and refining indicators that are relevant to their clinical practice and lived experience. The source and provenance of each indicator, including all suggestions, will be posted online and updated as necessary.
What is the rationale behind the methodology? **OR**	Emerging schools of thought suggest that 'appropriate care' may be enhanced through greater patient (health consumer) engagement. This could be facilitated by involving patients and interested laypeople as well as HCPs in clinical practice

What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?	guideline development, and using online technologies to enhance transparency, accessibility and currency of both content and development processes. The strategies employed in this protocol aim to mitigate problems with existing clinical practice guideline development processes by adopting a single approach to avoid duplication, using an ongoing and iterative documentation process to facilitate transparent synthesis of up-to-date information and make its provenance accessible, and requiring all participants to declare their conflict of interest.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	**Development and test of a process for creating clinical indicators representative of 'appropriate care' for a range of common conditions** **Clinical indicators will be developed for individual conditions using a four-stage process:** ▪ **Source, select and search relevant clinical practice guidelines:** o Interview data from the stakeholder analysis (phase 1) will be used in conjunction with national health priority areas, burden of disease and prevalence data to identify candidate conditions for clinical standard and indicator development. Clinical indicators will be drawn initially from the latest clinical practice guidelines. A systematic search will be undertaken of national-level Australian clinical practice guidelines endorsed by the NHMRC, and international-level guidelines from NICE, SIGN, and the AHRQ National Guideline Clearinghouse in the USA, and GIN. In the absence of Australian national or international clinical practice guidelines, relevant professional medical college and association clinical practice guidelines may also be searched, as well as those published at state or professional level and in international journals. ▪ **Extract all concepts from each clinical practice guideline together with the relevant text in which they appear (original recommendation), and tabulate common concepts to select, draft and format the proposed clinical indicators based on identified concepts:** o Recommendations (and their key underlying concepts) from each clinical practice guideline will be collated and used to inform the content of the proposed clinical indicators. Not all recommendations published in clinical practice guidelines will become indicators. Recommendations will be flagged for potential exclusion based on the following criteria: ▪ strength of the wording of the recommendation (i.e., 'may' and 'could' statements would be excluded; 'should' and 'must' statements would be included) ▪ vague guiding or aspirational statements and those without recommended actions ▪ conflicting recommendations from less recent clinical practice guidelines and those with lower AGREE-II scores. o All clinical indicators will be written in plain English, one concept at a time, using a structured and standardised format (e.g., commencing with the inclusion criteria followed by the compliance action). For each condition, indicators will be arranged according to phases of care (i.e., screening, diagnosis, assessment, acute care, ongoing care) so that together, they constitute a clinical standard amendable to inclusion as a clinical tool over the patient journey.

	▪ **Review the indicators internally:** o Internal reviews will first be conducted within the research team, and subsequently by Clinical Champion and Curator Group members who will comprise a mix of at least two members of the following: clinicians (e.g., general practitioners, medical specialists, allied health professionals, nurses), researchers, policymakers or public health specialists or healthcare quality improvement experts, and consumers. o In the first round, drafts of proposed clinical indicators, and the recommendations on which they are based, will be sent via email to the Curator Group members. The review criteria to be used are based on the methods from previous studies for developing and measuring indicators of appropriate care. Curator Group members will be asked to: recommend indicators for inclusion (with or without amendments) or exclusion, provide comments in relation to three key criteria: evidence, feasibility and importance, and make additional suggestions (with supporting material). In addition, research team members will pose specific questions to the Curator Group members about individual indicators to highlight inconclusive or conflicting clinical practice guideline recommendations, or to clarify definitions for inclusion criteria and compliance actions. o In particular, consumer members of the Curator Group will be asked to vet the plain English wording of clinical indicators and a linked glossary of terms to ensure that content is appropriately targeted to the consumer audience. o In this round, Curator Group members will complete their assignments independently to minimise 'group-think.' Research team members will collate the feedback and revise the content, structure and format of each indicator. The refined set of indicators (including the original indicators and any feedback and suggestions) will be sent to the same Curator Group members for a second round of scoring. The same approach will be used in the second round, with a request for further refinement and identification of indicators to be included or excluded. If necessary, Curator Group members will discuss the proposed set of indicators via a third round teleconference, with a view to achieving consensus and approving the indicators for the external online wiki-based review process. ▪ **Review the indicators externally** o External reviews will be conducted by healthcare professionals and consumers who have registered to the wiki as reviewers (Wiki Registrants). Relevant medical colleges, professional and consumer associations and networks will be contacted to request assistance with the identification of potential clinical indicator reviewers. Invitations will be by email, media releases and articles within newsletters. Healthcare professionals and consumers will self-nominate as reviewers for one or more of the STANDING Collaboration conditions based on their interests, scope of practice and experience. Wiki Registrants for this process will be required to declare their conflicts of interest, which will be taken into account by the Clinical Champion and Curator Group when considering reviewers' feedback on the indicators. The external review will involve an interactive wiki-based process where indicators for each condition from round 3 of the internal review will be posted on an online wiki site. A software development company will be engaged to purpose design and custom build the wiki for this project. Reviewers will provide comments on indicators in relation to the three key criteria: evidence, feasibility and importance, make recommendations (i.e., inclusion, inclusion with

	amendments, exclusion, hold) and be able to suggest edits in real time. The Clinical Champion and Curator Group for each condition will follow up and manage external reviewers' responses, and make final recommendations for that version regarding the inclusion, content, structure and format of indicators. ○ The Clinical Champion and Curator Group will use supporting references when considering and responding to each suggestion related to whether and why they have been included or rejected. In addition, all external reviewers' comments and recommendations will be logged, classified and presented in subsequent rounds according to whether and why they have or have not been incorporated into the next iteration. This will allow tracking of the evolution of the standards and indicators from the original recommendations on which they were based to their final iteration, as well as the nature and influence of review feedback in shaping the standard. ○ Once the indicators are 'stable' with no further significant changes being suggested, that version of the standard will be published as comprising a set of clinical indicators that represents 'appropriate care' for Australians with the candidate conditions at that time. Endorsement will then be sought by relevant professional bodies and consumer organisations. For each medical condition that has undergone indicator development via the STANDING Collaboration process, it will be possible for evidence to be monitored by the Curator Group (or a subgroup comprised of key members of the Curator Group) in order to update standards and indicators as necessary. For each condition, the initial monitoring plan involves using information from automated database searches and feedback from the wiki to initially update indicators every three months and, once stable, at a minimum of every six months.
Notes	
Reviewer notes	RQ3: Innovation of the use of online-platform for development of standards and quality indicators. Category of evidence: Grade D.
Associated handbook(s)	N/R

Key: AGREE – Appraisal of Guidelines for REsearch & Evaluation; AHRQ – Agency for Healthcare Research and Quality; GIN – Guidelines International Network; HCP – healthcare professionals; N/A – not applicable; NHMRC – National Health and Medical Research Council; NICE – National Institute for Health and Care Excellence; N/R – not reported; SIGN - Scottish Intercollegiate Guidelines Network.

Table C48 A framework for practical issues was developed to inform shared decision-making tools and clinical guidelines

Publication identification	
Authors (year)	Heen et al. (2021)
Country	Norway, Canada, USA
DOI	https://doi.org/10.1016/j.jclinepi.2020.10.002
Publication description	
Design	Iterative mixed methods study design.
Objective	▪ To develop a generic framework of patient-important practical issues from patient experience databases. ▪ To integrate this framework in the MAGIC authoring and publication platform (MAGICapp) for the digital structuring of summaries of practical issues and inclusion in guidelines and shared decision making tools. ▪ To test the feasibility of using this framework and authoring tool in the production of international recommendations and linked SDM tools in The British Medical Journal Rapid Recommendations.
Summary/Overview	Practical issues are central in decision-making and need to be considered as a part of evidence summaries informing guidelines and SDM tools. The authors developed a generic framework that allows for inclusion of practical issues that proved feasible and, when implemented in guidelines and SDM, very useful.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	Incorporation of a generic framework of patient-important practical issues into an online authoring and publication platform (MAGICapp) to facilitate shared decision making.
What are the core elements of the key innovation?	The framework of patient-important practical issues included the following 15 categories: medication routine, tests and visits, procedure and device, recovery and adaptation, coordination of care, adverse effects, interactions and antidote, physical well-being, emotional well-being, pregnancy and nursing, costs and access, food and drinks, exercise and activities, social life and relationships, work and education, travel and driving. Incorporation of a generic framework of patient-important practical issues from patient experience databases into MAGICapp, an online authoring and publication platform for guidelines and evidence summaries. Here, evidence summaries are digitally structured, which allows for translation of data into various formats including multi-layered guidelines and SDM tools. After defining the final set of generic categories, the authors integrated the framework in the data structure of the MAGICapp structuring of data so that practical issues were included in guidelines and supporting evidence summaries created through the online platform, in complement to traditional outcomes for benefits and harms.

What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	The goal of SDM tools is to provide individual care that makes intellectual, emotional, and practical sense to each patient. However, many current decision support tools typically lack or have minimal information addressing the practical issues patients face when implementing treatment options or tests, and how implementation affects their daily life. Such omission is problematic: Addressing practical issues often contributes to what has been described as "the work of being a patient" and, when excessive, constitutes an onerous burden of treatment. This work was based on insights from the research and innovations from the MAGIC Evidence Ecosystem Foundation, a non-profit initiative set up to facilitate the digital creation, publication, and updating of guidelines, evidence summaries, and SDM tools.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	**Integration of practical issues in SDM tools** By clicking on each icon, users can access the content as a superimposed box displaying narrative information on the relevant practical issues, either as key words or short sentences. Categories that do not include any relevant content are still displayed on the top layer grid, as they may trigger meaningful questions from patients.
Notes	
Reviewer notes	RQ3: innovation of integration of patient-important practical issues in an online authoring tool and publication platform. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: MAGIC – MAking Grade an Irresistible Choice; N/A – not applicable; N/R – not reported; SDM – shared decision making.

Table C49 Towards the semantic enrichment of Computer Interpretable Guidelines: a method for the identification of relevant ontological terms

Publication identification	
Authors (year)	Quesada-Martínez et al. (2018)
Country	Spain
DOI	PMID:30815135 / https://pubmed.ncbi.nlm.nih.gov/30815135/
Publication description	
Design	Descriptive cross-sectional study
Objective	To describe a method to enrich a Computer Interpretable Guideline (CIG) by means of an OWL ontology that describes the clinical domain of the CIG, which could be exploited (e.g., for interoperability with an EHR).
Summary/Overview	The study describes a method to support the development of an ontology beginning with a CIG. CIGs are formalised versions of clinical practice guidelines for use as decision-support systems.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	CIGs: defined as formalised versions of clinical practice guidelines for use as decision-support systems.
What are the core elements of the key innovation?	The use of ontologies for bridging the gap between CIGs and other resources/systems, such as databases of clinical trials and electronic health records. Ontologies can be defined as controlled vocabularies that allow the description of the meaning of data (its semantics) in a human and machine-readable way. They are used increasingly often to aid processing of information in biomedical research and in healthcare systems.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	The emergence of CIGs was motivated by an interest in making clinical practice guidelines recommendations available to clinicians in an easier and immediate way, compared to clinical practice guidelines in text form. The authors suggest that the benefits of the use of CIGs in clinical settings include improved clinical practice guidelines compliance and increased efficiency.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	N/R
Notes	
Reviewer notes	RQ3: Innovation of CIGs. Category of evidence: Grade D.
Associated handbook(s)	N/R

Key: CIG – computer-interpretable guideline; EHR – electronic health record; OWL – web ontology language; N/A – not applicable; N/R – not reported.

Evidence and or guidance translation (evaluated)

Table C50 Improving the user experience of patient versions of clinical guidelines: user testing of a Scottish Intercollegiate Guideline Network (SIGN) patient version

Publication identification	
Authors (year)	Fearns et al. (2016)
Country	UK
DOI	https://doi.org/10.1186/s12913-016-1287-8
Publication description	
Design	User study using a think-aloud protocol method.
Objective	To user test a patient version of a SIGN clinical guideline that was designed based on preliminary work for the DECIDE project.
Summary/Overview	The study describes the testing of a patient version of a SIGN clinical guideline. Patient versions of clinical guidelines help patients to understand what to expect in an intervention and helps them participate more actively in the decision making. The patient versions of guidelines should be colourful, have simple language, simple diagrams, and icons/headings to indicate clear recommendations.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	Patient versions of guidelines.
What are the core elements of the key innovation?	**Themes emerging from user testing** ▪ Usefulness/Value ○ Information about risks is most useful if directly associated with information about self- management or any form of action that can be taken to counter the risks. ○ Simple diagrams and charts can communicate information clearly. ○ It is helpful to flag clearly any important areas not covered by the guideline. ○ Signposting to organisations that can provide help and further information is valued. ▪ Usability ○ Language should be kept as simple as possible. ○ Small font size, use of light/pale colours, and too much material on a page were major barriers to use of the guideline by the patient group. ○ Clear flagging of recommendations using headings/icons works well. ○ A risk of 2 in 100 was interpreted by some as very high and others as very low.

	○ Icons for levels of recommendation worked best when kept recognisable, with a clear link to the intended message. ○ Vague or generic icons can cause confusion and be misinterpreted e.g. a blue circle can be interpreted as a zero. ○ Uncertainty was effectively communicated by the "?" icon, but people may not know how to respond to this information. ▪ Credibility ○ Credibility arose from information on the guideline production process, and the involvement of qualified professionals. ○ The status of the guideline is important (do health services recognise the recommendations). ○ Credibility may be threatened by pathways or recommendations that do not fit with the patient's own experiences. ▪ Desirability ○ Aspects that increased desirability included a friendly tone, simple language, chunking of text, the use of colour, glossy "high quality" look, and use of icons/images. ○ A friendly feel was achieved by informal language, use of colour, and the inclusion of quotes and images/icons. ○ Negative language or images, and a bureaucratic/dogmatic tone were disliked. ○ Quotes can personalise the material, giving it an engaging and friendly tone, and emphasising a particular message. ▪ Accessibility/Findability ○ The brief contents page, with simple question based headings was clear and facilitated flicking to relevant sections. ○ The participants were very concerned about the apparent lack of dissemination of patient versions of guidelines. ○ It is important for printed copies of the guideline to be available. ○ The patient version must be tailored to the intended audience's needs (e.g., font size, language/numerical information). ○ Information on how to access the services/interventions recommended is important. ○ Clear branding as a patient version is required. ○ Clear information on "who this booklet is for" encouraged people to read and share the guideline. ○ It is important to provide telephone numbers and addresses as well as websites for signposted organisations.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	▪ Despite many guideline bodies producing patient versions, there is low awareness of the public towards clinical practice guidelines. The authors suggest that the public may not perceive this format of health information any more positively than alternative sources. ▪ There is a need for guideline producers to make clear how the information contained in the guideline is relevant to the patient and how it can be used in their healthcare. ▪ Patient versions of guidelines can inform and empower people to ask questions.

	▪ Patient versions can help people to anticipate what to expect when seeing a healthcare professional or having an intervention. They may be most useful to patients around the time of their diagnosis. ▪ Developing and testing patient versions of SIGN guideline could provide insights on these issues.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	N/R
Notes	
Reviewer notes	RQ3: Innovation of patient version of guidelines. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: DECIDE – Developing and Evaluating Communication Strategies to support Informed Decision and practice based on Evidence; N/A – not applicable; N/R – not reported; SIGN – Scottish Intercollegiate Guidelines Network.

Technological innovation (evaluated)

Table C51 A knowledge-modelling approach to integrate multiple clinical practice guidelines to provide evidence-based clinical decision support for managing comorbid conditions

Publication identification	
Authors (year)	Abidi (2017)
Country	Canada
DOI	https://doi.org/10.1007/s10916-017-0841-1
Publication description	
Design	Descriptive evaluation study .
Objective	To describe the integration of multiple disease-specific clinical practice guidelines in order to manage comorbidities within a computerised CDSS.
Summary/Overview	This paper describes the innovative manifestation of clinical practice guidelines integration with a CDSS that can provide clinical practice guideline-mediated recommendations to manage patients with comorbid conditions—a step up from existing CDSSs that manage single chronic diseases.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	Systematic integration of multiple disease-specific clinical practice guidelines in order to manage co-morbidities within a computerised CDSS. Clinical practice guidelines integration framework—known as COMET—that manifests a knowledge management approach to model, computerise and integrate multiple clinical practice guidelines to yield a comorbid clinical practice guidelines knowledge model that can provide evidence-based recommendations for handling comorbid patients. COMET exploits semantic web technologies to achieve: ▪ Clinical practice guidelines knowledge synthesis to translate paper-based clinical practice guidelines to disease-specific CPs that include specialised comorbidity management procedures based on input from domain experts. ▪ Clinical practice guidelines knowledge modelling to computerise the disease-specific CP using a comorbidity clinical practice guidelines ontology. ▪ Clinical practice guidelines knowledge integration by aligning multiple ontologically-modelled CPs to develop a unified comorbid clinical practice guidelines knowledge model. ▪ Clinical practice guidelines knowledge execution using reasoning engines to derive clinical practice guidelines-mediated recommendations for managing patients with comorbidities.

What are the core elements of the key innovation?	The proposed clinical practice guidelines integration approach exploits semantic web technologies to: (1) develop a high-level formal knowledge model — i.e., a comorbidity clinical practice guidelines ontology that represents the underlying concepts about clinical practice guidelines tasks, relationships, workflows, constraints and decision rules; (2) computerise the clinical practice guidelines as an instantiation of the comorbidity clinical practice guidelines ontology; (3) integrate the different clinical practice guidelines by establishing semantic alignments between their ontological representations, thus realising a comorbid clinical practice guidelines knowledge model; (4) execute the comorbid knowledge model, using reasoning engines, based on patient data and physician input to derive clinical practice guidelines-mediated recommendations for managing patients with comorbidities. **The COMET Framework: Clinical practice guidelines integration to handle comorbidities at the knowledge modelling level** The proposed clinical practice guidelines integration approach stipulates that: ▪ Multiple disease-specific clinical practice guidelines are modelled in terms of an illustrative CP model that articulates the clinical actions, decisions, sequence of the actions, associated constraints and local considerations. The localised CP outlines an operational schema of a disease specific clinical practice guideline. ▪ Experts review the multiple clinical practice guidelines and their operational CP to identify potential comorbidity considerations that pertain to specific clinical decisions and tasks within a clinical practice guideline, and how it influences the other concurrently applied clinical practice guidelines in a comorbid situation. ▪ Experts augment the disease-specific CP with comorbid compliant therapeutic decisions, actions and recommendations that correspond to the identified comorbidity considerations. ▪ The comorbid compliant clinical actions recommended by experts (in response to comorbidity considerations) are augmented to the original disease specific CP model so that it can now handle comorbid situations in a safe and efficient manner. The update to the CP model includes: (1) integrating the clinical actions of one CP with related actions in another CP; (2) adding specific comorbid actions to a CP; and (3) altering the clinical workflow of a CP to align with a concurrent CP in order to avoid duplicate tasks or to minimise time and resources. This achieves clinical practice guidelines integration for managing comorbid conditions at the knowledge level. ▪ The comorbid compliant CP model—manifesting comorbidity management knowledge—is then computerised in terms of an executable ontological CP knowledge model that integrates multiple clinical practice guidelines to realise a comorbid clinical practice guidelines knowledge model. In essence, the knowledge model comprises the independent disease-specific CP (each CP is modelled as a separate instantiation of the model) and the integrated comorbid CP model. This achieves clinical practice guidelines integration for managing comorbid conditions at the knowledge modelling level. The clinical practice guidelines integration approach takes as input multiple clinical practice guidelines corresponding to comorbid diseases, and in return yields a reconfigured comorbidity compliant CP that is systematically formulated by reviewing, revising and resolving common and conflicting tasks across the comorbid clinical practice guidelines. The comorbid CP coordinates the overall comorbid care process by recommending clinical actions that are comorbid compliant in terms of addressing the overall patient condition (as opposed to disease-based recommendations), while avoiding adverse interactions across tasks stipulated by concurrent independent clinical practice guidelines. The authors

	formulated a generic multi-step clinical practice guidelines integration framework — i.e., COMET — that covers the entire lifecycle of developing a CDSS for handling comorbidities. COMET leverages Semantic Web methodologies and tools for clinical practice guideline representation, integration and execution. The development of the comorbid clinical practice guidelines knowledge model is an extension of the traditional ontology development lifecycle, where clinical practice guidelines integration methods that capture the functional and temporal relationships between overlapping care processes across multiple clinical practice guidelines were augmented.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	The rationale behind systematically integrating multiple disease-specific clinical practice guidelines within a computerised CDSS to formulate a safe and clinically pragmatic comorbid clinical practice guidelines knowledge model include: ▪ Reducing potential for patient harm when managing comorbidities. The concurrent application of multiple clinical practice guidelines is not a safe approach because the recommendations stipulated by one disease-specific clinical practice guideline may be clinically incompatible and in conflict with recommendations from other disease-specific clinical practice guidelines. ▪ Resolving conflicts between the disease-specific diagnostic and therapeutic decisions/action as recommended by multiple clinical practice guidelines. ▪ Outlining the benefits and risks of clinical actions from a patient perspective (and not an individual disease perspective). ▪ Synergising the underlying clinical workflows of multiple clinical practice guidelines to avoid duplication of tasks in order to optimise resource utilisation.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	Clinical practice guidelines integration at the knowledge modelling level is clinically pragmatic, as this approach offers the following advantages: (1) The integrated clinical practice guidelines model is guided and validated by domain experts; (2) The integrated clinical practice guidelines model is scalable — i.e., more than two clinical practice guidelines can be integrated; (3) The integrated clinical practice guidelines model renders an executable clinical workflow that is comorbid compliant; and (4) The integrated clinical practice guidelines model can be readily reconfigured to incorporate clinical practice guidelines updates. Using a knowledge modelling-level approach, a therapy plan for comorbid patients can be formulated by re-using and re-assembling specific sections of multiple clinical practice guidelines.
Notes	
Reviewer notes	RQ3: Innovation of integration of multiple disease-specific clinical practice guidelines in order to manage co-morbidities within a computerised CDSS. Category of evidence: Grade D.
Associated handbook(s)	N/R

Key: CDSS – clinical decision support system; COMET – Comorbidity Ontological Modeling & ExecuTion; CP – clinical pathway; N/A – not applicable; N/R – not reported.

Table C52 Automatically finding relevant citations for clinical guideline development

Publication identification	
Authors (year)	Bui et al. (2015)
Country	USA
DOI	https://doi.org/10.1016/j.jbi.2015.09.003
Publication description	
Design	Mixed methods: The study design consisted of three main parts: (1) development of a gold standard composed of studies used in the development of cardiovascular guidelines; (2) iterative development of a citation finding system composed of two main components: query expansion and citation ranking; and (3) evaluation of each system component using standard information retrieval metrics and comparison with baseline approaches.
Objective	To improve the traditional literature database search approach using innovative query expansion and citation ranking.
Summary/Overview	The study describes an automated approach to retrieve relevant and high-quality citations from PubMed. The approach can be used to assist the development of clinical guidelines and systematic reviews. The results showed that the proposed method outperformed the default PubMed query expansion in terms of recall (80.2% vs. 51.5%) and seeding recall (90% vs. 63.5%), with a non-significant loss in precision (0.6% vs. 0.4%; p = 0.09).
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	Citation retrieval system composed of query expansion and citation ranking methods.
What are the core elements of the key innovation?	**Gold standard** ▪ The gold standard consisted of citations that have been used to support guideline recommendations. For those guidelines discussing the comprehensive management of the conditions of interest, the following steps were performed to build the gold standard: (1) Extracted all the citations listed in the 'References' section of the guideline; (2) extracted the guideline recommendations whose evidence sources were provided in the guideline and the citations that were used as evidence sources to support each recommendation; and (3) automatically mapped those citations in free-text to PMIDs using the NCBI Batch Citation Matcher tool. Manual mapping was performed to supplement the citation IDs that could not be matched by the NCBI tool. **System overview** ▪ The system is an extension of PubMed's search engine to enhance the ability to retrieve citations for clinical guideline development. The system has a pre-processing stage and two other main stages: query expansion and document ranking. The query expansion stage aims to improve recall (that is, the identification of possibly

relevant studies from electronic literature databases) while the document ranking aims to improve precision on top-ranked documents.

Pre-processing

- This step takes the title of the guideline as input and extracts the conditions of interest. This step also detects whether a particular guideline focuses on one or more conditions.

Query expansion

- Based on the extracted condition terms, a search using PubMed's default search behaviour was conducted. When entering a query on the PubMed search interface, PubMed automatically expands the query to maximise recall. The results of PubMed expansion were used as the baseline to compare with the proposed expansion approach. The proposed approach aimed to find relevant and meaningful MeSH terms of the condition topics. Additional MeSH terms were injected to the original query using the Boolean OR operator.
    - **Common filter:** A set of filters (i.e., publication date, human study and English language) were consistently applied for all queries generated.
    - **MeSH expansion:** An algorithm was developed to expand the seed query using MeSH resources (MeSH descriptors, MeSH Tree), and a natural language processing application (Metamap). The algorithm takes input as a single search query and outputs the expanded query. If there were multiple queries (multiple conditions), they were joined by the Boolean OR operator. Eventually, the query was adjusted by the common filter and applied the PubMed sorting mechanisms. To conduct a PubMed query, the PubMed query was formulated into the URL syntax and used the Entrez Programming Utilities (E-utilities) to submit and retrieve results from the NCBI servers. The algorithm uses the following methods to find relevant MeSH concepts:
        - Statistical expansion
        - Body-part expansion
        - Parent expansion
        - MeSH stop list.

Document ranking

- There are three ways searchers can obtain a ranked list of citations:
    - **PubMed sorting functionalities**: PubMed offers seven ways to sort order for search results: Most Recent, Relevance, Publication Date, First Author, Last Author, Journal, and Title. Most Recent is PubMed's default sorting, which ranks citations by the time they were added to the MEDLINE database. The Relevance sort uses PubMed's internal algorithm to assign weight to citations depending on the frequency with which search terms were found and the fields in which they were found. The authors used and evaluated the Most Recent and Relevance sorts to compare with the proposed ranking approach. The other sorts based on publication time and alphabetical orders were less likely to identify relevant citations.
    - **A machine learning approach**: Using a general purpose machine learning classifier to identify scientifically rigorous clinical studies. In 2009, Kilicoglu et al.[6] implemented an ensemble approach

	combining several machine learning classifiers (Naïve Bayes, support vector machine, and boosting) to identify scientifically rigorous studies. The classifier was built on five basic features: words, MEDLINE metadata, sematic predications, relations, and UMLS concepts. In the original study, the classifier trained on 10,000 citations could achieve 82.5% precision and 84.3% recall on an unseen test set of 2,000 citations. The classifier outputs the probability that a citation is scientifically rigorous. These classifiers were used as the baseline ranking approach. ○ **Clinical research scoring approach**: The authors proposed an alternative method for ranking MEDLINE citations using three dimensions: MeSH majority, study design, and journal ranking. These dimensions attempt to capture three characteristics that are desirable for retrieved studies: relevancy, study quality, and study impact. ▪ **MeSH majority**: a PubMed document can be indexed with multiple MeSH concepts, but only a small subset are indexed as 'major topic'. Using the expanded MeSH concepts from the query expansion stage, the authors assigned a MeSH score of 2.0 if one of the MeSH concepts or any of its children was tagged as a major topic. Otherwise, a MeSH score of 1.0 was applied. ▪ **Study design**: The authors assigned a Study Design score to a study based on the publication type of the retrieved document (score 4.0: Practice Guideline, Guideline, Review with Meta-Analysis; score 3.0: Randomised Controlled Trial; score 2.0: Clinical Trial, Controlled Clinical Trial, Case-Control Studies, Cohort Studies, Longitudinal Studies, Cross-Sectional Studies, Cross-Over Studies, Observational Study, Evaluation Studies, Validation Studies, Comparative Study; and score 1.0: any other types). The rationale for the Study Design scoring was adapted from the GRADE system. If a study had multiple publication types, the maximum SD score found on the matrix was chosen. The Study Design score was increased with the presence of blinding methods (single-blinded method +0.1, double-blinded method +0.2) and setting (multicentre study +0.1). ▪ **Journal ranking**: Journal ranking is an estimation of scientific quality and clinical impact of the study based on the popularity of the publishing source. The authors used the open-access SJR, an impact factor metric, published by Scopus in 2012. The National Library of Medicine's journal records were mapped to Scopus' records using the journal's ISSN number, from which the SJR metric was retrieved. ▪ **Ranking score:** Finally, the ranking score was calculated by multiplying all three metrics (ranking score = MeSH Major Score x SD score x SJR). Since those metrics are independent, multiplication was considered to be the most appropriate method to aggregate the three metrics.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	Literature database search is a crucial step in the development of clinical practice guidelines and systematic reviews. In the age of information technology, the process of literature search is still conducted manually, therefore it is costly, slow and subject to human errors. Hence, the authors investigated the literature search stage and aimed to maximise recall while controlling the impact on precision. They developed and assessed query expansion and ranking methods to enhance information retrieval performance in the context of clinical guideline development. The solution was based on an extension of PubMed's search engine, optimised to retrieve and rank relevant studies for cardiovascular guidelines.

What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	**Evaluation of the innovation** From 2010 to 2014, the American College of Cardiology published 17 guidelines about cardiovascular topics. Focus Update releases (without systematic searches) and guidelines without comprehensive management of a condition were excluded. Eight guidelines met the inclusion criteria for the evaluation study. The authors extracted 653 practice recommendations, which cited 1,863 citations. Of those, the authors found PMIDs in 1,848 citations (99.2%). A small portion of citations such as book chapters, online resources (e.g., FDA site), and studies not indexed in MEDLINE did not have PMIDs. Overall, the query expansion algorithm achieved recall of 80.2% and seeding recall of 90.1%. In comparison with the default PubMed expansion, the algorithm improved recall by 28.7% and seeding recall by 26.5%, with a 0.2% drop in precision. The ability to find seed studies (seeding recall) improved by 26.6%. The query expansion algorithm could find all citations for more guideline recommendations than the default PubMed expansion (64.5% vs. 37.2%, $p<0.0001$). For citation ranking, the clinical research scoring approach had the best average precision of 7% compared to 2.1% machine-learning classifier, 0.9% PubMed's sort by Relevance, 0.5% PubMed's sort by Most Recent. Similarly, the scoring approach had the highest average recall, improved 4.2% over the machine-learning classifier (66.2% vs. 62%, $p<0.001$), 14.8% over PubMed's sort by Relevance (66.2% vs. 51.4%, $p<0.001$), and 21.1% over PubMed's sort by Most Recent (66.2% vs. 45.1%, $p<0.001$).
Notes	
Reviewer notes	RQ3: Innovation of an automated approach to retrieve relevant and high-quality citations from PubMed to assist in the development of clinical guidelines and systematic reviews. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: FDA – Food and Drug Administration; GRADE - Grading of Recommendations Assessment, Development and Evaluation; ID – identifier; MeSH – Medical Subject Headings; ISSN – International Standard Serial Number; N/A – not applicable; NCBI – National Center for Biotechnology Information; N/R – not reported; PMID – PubMed identifier; SJR – SCImago Journal Rank; UMLS – Unified Medical Language System.

Table C53 Adoption of an electronic template to promote evidence-based practice for policies, procedures, guidelines and directives documents

Publication identification	
Authors (year)	Corey et al. (2018)
Country	USA
DOI	https://doi.org/10.1097/DCC.0000000000000305
Publication description	
Design	Mixed methods: survey and consensus
Objective	To develop an online template format to introduce an evidence-based practice system that supports and guides the clinicians' clinical practice.
Summary/Overview	This study describes the process of developing a standardised electronic template for clinical practice guidance documents, such as protocols, policies, guidelines and directives, and describes the inclusion of clinical staff in the development and implementation of evidence-based policies and the need for clinical staff to be literate with technology to assist implementation.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	Electronic template for PPGD.
What are the core elements of the key innovation?	**Topics/headings included in the electronic template for PPGD** ▪ Standardised purpose statement o The purpose of this policy is to align clinical practice with the best evidence or consensus of expert opinion available in the literature. ▪ Table of contents o Navigation enabled. ▪ Policy statement(s) o In bullet points, write statement defining what must be done or followed and describe the legal responsibility of the clinician performing the task or operating the equipment, including who can administer/perform therapy/procedure, if applicable. ▪ Clinical indications/contraindication o In bullet points, state the conditions or situations that a clinician must consider when deciding to include or exclude the use of this therapy or protocol. ▪ Equipment o In bullet points, list specific equipment needed for all procedural steps.

	▪ Room entry procedure o All patients are provided care with respect and dignity. Always adhere to room entry procedures: identify yourself to the patient and family, perform hand hygiene, use two patient identifiers, and educate the patient and family before procedures. o All patients are provided care using the principles of standard precautions. Special precautions are managed according to the hospital infection control policies/guidelines: Infection Control Manual. ▪ Procedure title o Followed by table with: Procedure Steps and Rationale. ▪ Patient Monitoring and Care with rationale and consideration o List element(s) necessary to the care and monitoring of patient. Write none if not applicable. ▪ Nursing documentation o List data points, required frequency, and location of documentation(s) e.g., vital signs hourly on electronic flow sheet. Write none if not applicable. ▪ Considerations/additional education o In bullet points, state important clinical information that would support the clinician when utilising this policy/procedure. ▪ Appendix o Add flow charts, graphs, pictures, tables, flow sheets, etc. as necessary. ▪ Document history o Related policies or guidelines. ▪ Key words o Add likely search terms that end users would type to search for this policy. ▪ References o American Psychological Association referencing format includes levels of the evidence.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	PPGDs were underused by clinicians within the CCPM by the intensive care unit. The reasons were: ▪ Inconsistencies in the formatting of PPGDs ▪ Updated PPGDs provided new references but did not always review and remove outdated references when applicable ▪ Citations were absent within the text of PPGDs, so identifying where the evidence came from for best practice was difficult ▪ Some authors of critical care policies, procedures, and guidelines were unclear as to what constituted a policy statement, leading to some guidelines being inaccurately identified as policies ▪ The end users, direct care critical care nurses, consistently identified two barriers to utilisation of the CCPM: difficulty finding PPGDs and lengthy policies that were cumbersome.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	The electronic template for PPGD developed using Microsoft Word 2010, contained the elements required by hospital policy regarding PPGDs but leveraged technology to enhance navigation of the PPGDs. The new policy template provided information in each section to guide the authors. The hospital-wide policy and procedure manual contained a document with the hospital's definition of these terms, which were used in a condensed version.

	Evaluation of the innovation A survey was conducted with 106 critical care nurses to assess the effectiveness of the new PPGD format, as well as process changes identified as problems that needed to be addressed by an audit of the PPGDs conducted in 2013. ▪ The survey included a question on how often the critical care nurses accessed the PPGDs in a month, as a means of assessing their utilisation of the contained policies and procedures. One hundred percent of the respondents (n=106) accessed the PPGDs at least once a month. The nurses' ability to find a desired PPGD was also assessed, because it was identified as a key barrier to utilisation of PPGDs by an audit in 2013. Only 24% of the respondents (n=15) still felt that PPGDs were difficult to find. 75% of the respondents were looking for PPGDs in the correct place on the hospital portal. ▪ A Likert scale was used to determine the nurses' perception of the effectiveness of the PPGDs by asking specific questions to elicit feedback as to their opinions regarding the quality and value of the PPGDs found on the hospital portal. If the critical care nurses did not perceive PPGDs to contain valuable information, then it would be expected they would not be used. The overall results indicated that the respondents agreed or strongly agreed that the PPGDs provided guidance (85%), reflected current practice (76%), were clear and concise (75%), and were evidence-based (73%). ▪ To assess the value the nurses placed on the body of evidence from which the PPGDs were derived, a number of questions were asked regarding the references (that is, references to the evidence base). All 106 nurses looked at the PPGDs' references sometimes (61%), most of the time (32%), or always (6%). The nurse respondents indicated that they believed it was essential to know the source (44%) and strength (68%) of the evidence behind a PPGD. Whereas only 22 nurses commented on the value of citing the reference within the body of the PPGD, most of the respondents (59%) were neutral as to the citing of evidence within the PPGD. ▪ Questions were designed to clarify the utilisation of the navigation panel embedded in the table of contents. Forty-four percent of the nurses found this function helpful, whereas 13% did not. A large number of nurses (41%) were neutral on the subject, which suggests that they may not have used this function, leading the authors to wonder whether this was due to a lack of knowledge regarding the use of technology within the PPGD.
Notes	
Reviewer notes	RQ3: innovation of standardised electronic template for clinical practice guidance including standard practice elements (quality indicators). Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: CCPM – Critical Care Practice Manual; N/A – not applicable; N/R – not reported; PPGD – protocols, policies, guidelines, and directives.

Table C54 Efficiency of pragmatic search strategies to update clinical guidelines recommendations

Publication identification	
Authors (year)	Martínez García et al. (2015)
Country	International
DOI	https://doi.org/10.1186/s12874-015-0058-2
Publication description	
Design	Evaluation: Descriptive study of search strategies
Objective	To evaluate the efficiency and feasibility of two new approaches to identify the need to update clinical guidelines recommendations: the development of search strategies using PubMed Clinical Queries for MEDLINE and the use of the PLUS (McMaster Premium Literature Service) database.
Summary/Overview	The study evaluated two search strategies to identify signals for updating recommendations and compared them to an exhaustive search strategy using a random sample of recommendations from a cohort of clinical guidelines from a national guideline development programme. The proposed method of developing restrictive search strategies, using PubMed Clinical Queries filters in the MEDLINE database, provides a feasible and efficient method for guideline developers to identify significant new studies that are likely to trigger a recommendation update. Searching only in the PLUS database was a suboptimal approach that needs topic specific tailoring.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	▪ **Restrictive approach** Guideline methodologists, trained by researchers with experience designing search strategies, developed restrictive search strategies for each clinical question using the PubMed Clinical Queries search filters for the MEDLINE database. Clinical questions that had at least two PICO components were considered. The restrictive search strategies were developed considering the minimum number of MeSH terms and text words required from the original exhaustive search strategies. The search strategies were designed in four stages: 1) Development: selection of keywords from the clinical questions and identification of MeSH terms and text words in titles; 2) Validation: Evaluation of whether each search retrieved all the original references for its corresponding recommendation; 3) Refinement: If a search did not retrieve all the original references, selection and search of less specific MeSH and/or text words in the title or abstract; and 4) Application of each of a broad and a narrow treatment Clinical Queries filter (www.ncbi.nlm.nih.- gov/pubmed/clinical), and a systematic review filter. ▪ **PLUS approach**

	An information specialist from the Health Information Research Unit developed a PLUS search strategy for each guideline topic. MeSH and SNOMED indexing terms in the PLUS database with clinical guideline topics was done. Both primary and review papers were included. To take into account the time delay associated with the critical appraisal process of the articles, the PLUS search strategies were ran from the beginning of the year in which the original exhaustive searches were run, until approximately three months beyond the latest date of the exhaustive searches.
What are the core elements of the key innovation?	The approach aims to identify the need to update clinical guidelines recommendations through: ▪ Restrictive search strategies using PubMed Clinical Queries search filters for MEDLINE ▪ The use of the PLUS database.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	It is challenging for guideline developers to screen for new, relevant evidence that justifies a clinical guideline update. The current practice to update clinical guidelines is based on new evidence identified using original exhaustive search strategies. So far, little empirical work has been conducted to test the effectiveness and efficiency of alternative searching processes. While researchers have tested restrictive approaches and alternative strategies to update systematic reviews, more information is needed about the timing and type of search. This study aimed to evaluate the efficiency and feasibility of two approaches to identify the need to update clinical guidelines recommendations.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R
How is the innovation used in practice?	**Comparison between the exhaustive approach, restrictive approach and PLUS approach** The restrictive approach (using a narrow PubMed Clinical Queries filter, clustering results per clinical guideline and imputing exhaustive search results for clinical questions with less than two of the four PICO components, or clinical questions pertaining to prognosis or diagnosis) retrieved 68.1% fewer references than the exhaustive approach, and identified most of the key references (62/69, 89.9 %) and recommendations updates (22/25, 88.0%). The PLUS approach retrieved 88.5% fewer references than the exhaustive approach and identified a substantially lower number of key references (18/69, 26.1%) and potential updates (10/25, 40%) than the restrictive approach.
Notes	
Reviewer notes	RQ3: innovation of pragmatic search strategies to update clinical guideline recommendations. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: MeSH – Medical Subject Heading; N/A – not applicable; N/R – not reported; PICO – Population/patient, Intervention, Comparison, Outcome; PLUS – McMaster Premium Literature Service; SNOMED – Systematised Nomenclature of Medicine.

Table C55 Deep neural network for reducing the screening workload in systematic reviews for clinical guidelines: Algorithm validation study

Publication identification	
Authors (year)	Yamada et al. (2020)
Country	UK, Japan, USA
DOI	https://doi.org/10.2196/22422
Publication description	
Design	Evaluation study
Objective	To investigate whether a machine learning system could perform the citation screening aspect of systematic reviews more efficiently.
Summary/Overview	The study investigates the efficiency of a machine learning system to perform systematic reviews. The study found that an active machine learning system could improve the precision of the systematic review process as well as reduce the time required, thus assisting with the development of clinical guidelines.
RQ1: Description of core components of clinical practice guidance	
What core components have been stated in the document?	N/A
RQ2: Description of quality measures/criteria for clinical practice guidance development	
What quality measure tools are there to examine the robustness of methodological process used to develop the various types of clinical practice guidance?	N/A
What criteria does the tool use to assess quality?	N/A
What are the strengths and limitations of the tool?	N/A
RQ3: Description of key innovations in the development and implementation of clinical practice guidance	
What innovative methodologies have been used to develop and or implement clinical practice guidance?	Active machine learning system (Concept Encoder, Fronteo Inc.) to reduce the workload in systematic reviews.
What are the core elements of the key innovation?	**Systematic review workload reduction** Reference lists of five clinical guidelines released by the American Diabetes Association, American College of Cardiology, American Heart Association (two guidelines) and American Stroke Association were reviewed. Search strategies of systematic reviews and meta-analyses of interventional randomised controlled trials that informed these guidelines were reproduced. The resulting primary screening dataset included all of the correct articles (i.e., those cited in the original systematic review) when it was reproduced according to the published search strategies. Correct articles (those actually reviewed/cited) and incorrect articles (those not reviewed/cited) were then extracted from the dataset. Using two randomly selected correct articles (that is, selected by Concept Encoder from those cited in the original systematic review), the following steps were performed to calculate how much workload reduction could be achieved using Concept Encoder. 1. Concept Encoder was used to convert sentences into vectors, extract and learn each vector component as a feature value, identify similar vectors as indicators of the similarity of sentence content, and perform a rapid search for similar sentences. Vectorisation facilitated text analysis by providing numerical data that allowed various calculations to be performed (e.g., to assess clustering of results). In addition, vectorisation allowed searches to be based on the sums and differences of sentences, facilitating comparison of content between two sentences and resulting in a sentence retrieval engine that could be adapted to research targets.

	2. Concept Encoder read the two articles and calculated the mean value of the sentence-word vectors corresponding to the two articles. Next, this mean value was used to assign scores to the other articles by determining the cosine distance between the mean value and the vectors corresponding to each of the remaining articles. 3. A researcher reviewed the article with the higher score. If this was a correct article, Concept Encoder learned it as a correct article based on the mean value of all chosen sentence-word vectors. If it was an incorrect article, the sentence-word vector is subtracted from the mean vector of the corrected (reviewed/cited) articles. 4. Concept Encoder learned the correct and incorrect article, and thus identified and rescored the remaining articles, which had not been checked by the researcher. 5. The researcher again reviewed the article with the highest score. If this was a correct article, Concept Encoder learned it as a correct article. If it was incorrect, Concept Encoder learned it as an incorrect article. 6. After learning all of the correct and incorrect articles identified up to this point, Concept Encoder scored the remaining articles again. The mean of sentence-word vectors for all corrected articles minus the mean of sentence-word vectors for all incorrect articles was used to score the remaining articles. 7. Steps 2 to 5 were repeated until all of the correct articles had been identified. Following this, the final reading ratio was calculated as the number of articles read by Concept Encoder relative to the total number of articles. For example, if the total dataset comprised 1,000 articles, and Concept Encoder found all of the correct articles after reading 200 articles, the final reading ratio would be 20%, and the work involved in screening the literature would have been reduced by 80% (avoiding the need to read 800 out of 1,000 articles). Work saved over sampling (WSS) at targeted extraction of 95% of correct articles (R%) is an index to measure how much work is saved compared to manual screening to achieve identification of R% of correct papers. 8. Next, the first correct article (step 2) was changed, and the same process was repeated until all of the correct articles were identified. 9. The maximum reduction of the literature screening workload achieved by teaching Concept Encoder two correct articles (i.e., two articles that were actually reviewed) was determined.
What is the rationale behind the methodology? **OR** **What criteria were used to determine if an innovation was necessary and if it was necessary, the type of innovation indicated?**	Performing systematic reviews is a time-consuming and resource-intensive process. To reduce the time and cost of screening literature when performing systematic reviews, researchers have explored the use of active learning text classification systems to achieve semi-automated exclusion of irrelevant studies while retaining a high proportion of eligible studies for subsequent manual review. However, little progress has been made for the following reasons. First, previous studies did not investigate well-characterised and high-quality datasets, so the type of systematic review used as the data source was unclear, and the method of applying machine learning to the clinical studies was obscure. Second, previous reports did not specify how active machine learning was used. Third, only an approximate 30-50% reduction of the workload was achieved. Fourth, a method that extracts 100% of the correct articles from the literature has not been developed because most studies use a targeted extraction of 95% as the primary outcome; despite the importance of not missing any eligible studies when performing systematic reviews (i.e., the objective is to identify all relevant articles). To overcome some of these issues, the authors studied systematic reviews of randomised controlled trials cited in several recent international clinical guidelines to investigate whether an active machine learning system (Concept Encoder, Fronteo Inc.) could reduce the workload and accelerate the review process while improving its precision.
What changes have been made in governance procedures for tracking of guidance as it becomes available for updating?	N/R

How is the innovation used in practice?	▪ **Reproduction of the Review (Evaluation Study)** The similarity of any two articles is defined as the cosine distance of the two sentence-word vectors associated with these articles. After a correct (reviewed) or incorrect (not reviewed) article is identified, the associated row vector is defined as correct or incorrect and used as the feature vector representing a correct or incorrect article. The cosine distances for all other articles (m − 1 articles) are calculated and arranged in descending order. For the next article from the top of the list, if the article is a correct one, the mean of the vectors for the correct articles is used to train Concept Encoder in the next step of active learning. If the article is an incorrect one, the vector is subtracted to train Concept Encoder in the next step of active learning, that is, it is used as the feature vector. Cosine distances between the updated vector and all other articles are calculated and ordered again, and this process is repeated until all of the correct articles have been identified. Here, the mean vector is simply used as the feature vector for the correct articles. Classification models could be built upon using these vectors as features to arrange the remaining articles in a descending manner by active learning; however, similarity of articles seemed to be embedded in the vectors, and using the vectors directly as the features was effective. Therefore, the process was kept simple, and no further machine learning was conducted in the active learning process. The deep neural network–based active machine learning system was found to significantly reduce the workload by at least 0.867 compared with manual screening (the lowest mean work saved over sampling at 95% (WSS@95%)). The average reduction of the workload compared with manual screening was >90% or 10-fold (WSS@95%: mean 0.904), and Concept Encoder showed a high ability to discriminate between correct and incorrect studies.
Notes	
Reviewer notes	RQ3: Innovation of use of machine learning for article screening in systematic reviews. Category of evidence: Grade C.
Associated handbook(s)	N/R

Key: N/A – not applicable; N/R – not reported; R – targeted extraction of 95% of correct articles; WSS – Work saved over sampling.

References

1. Schünemann HJ, Wiercioch W, Etxeandia I, Falavigna M, Santesso N, Mustafa R, Ventresca M, Brignardello-Petersen R, Laisaar KT, Kowalski S, Baldeh T. Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise. Cmaj. 2014 Feb 18;186(3):E123-42.

2. Thayer KA, Schünemann HJ. Using GRADE to respond to health questions with different levels of urgency. Environment international. 2016 Jul 1;92:585-9.

3. Ganann R, Ciliska D, Thomas H. Expediting systematic reviews: methods and implications of rapid reviews. Implement Sci. 2010;5:56.

4. Cameron A, Watt A, Lathlean T, Sturm L. Rapid versus full systematic reviews: an inventory of current methods and practice in Health Technology Assessment. ASERNIP-S Report No. 60. Adelaide: ASERNIP-S, Royal Australasian College of Surgeons; 2007.

5. Opiyo N, Shepperd S, Musila N, Allen E, Nyamai R, Fretheim A, et al. Comparison of alternative evidence summary and presentation formats in clinical guideline development: a mixed-method study. PLoS One. 2013;8:e55067.

6. Kilicoglu H, Demner-Fushman D, Rindflesch TC, Wilczynski NL, Haynes RB. Towards automatic recognition of scientifically rigorous clinical research evidence. Journal of the American Medical Informatics Association. 2009 Jan 1;16(1):25-31.