

# Trustworthy evidence-based versus untrustworthy guidelines: detecting the difference

João Pedro Lima  , Wimonchat Tangamornsuksan, Gordon H Guyatt

**To cite:** Lima JP, Tangamornsuksan W, Guyatt GH. Trustworthy evidence-based versus untrustworthy guidelines: detecting the difference. *Fam Med Com Health* 2023;11:e002437. doi:10.1136/fmch-2023-002437

JPL and WT contributed equally.

Received 15 August 2023  
Accepted 15 September 2023

## ABSTRACT

Guidelines are essential tools in healthcare decision-making. Trustworthy guidelines inform clinicians not only on the direction (against or in favour) and strength (strong or weak/conditional) of recommendations but also on the certainty of the underlying evidence. Developing trustworthy guidelines requires panellists with clinical and methodological expertise who consider patients' values and preferences. Adherence to trustworthiness standards remains variable; clinicians should, therefore, be able to distinguish trustworthy from untrustworthy guidelines. In this paper, we offer eight domains of disparities between trustworthy evidence-based guidelines and less trustworthy guidelines.

## INTRODUCTION

In making decisions, patients and clinicians must trade off the benefits against the downsides of alternative strategies. Formal guidelines are essential tools in healthcare decision-making which, when optimally developed, provide health professionals with trustworthy evidence-based recommendations to guide their practice when confronting such trade-offs.<sup>1</sup>

In producing trustworthy guidelines, developers must consider, in addition to the best estimates of effect size for both harms and benefits, their confidence in these estimates (the certainty or quality of the evidence).<sup>2</sup> Trustworthy guidelines should, therefore, inform clinicians not only on the direction (against or in favour) and strength (strong or weak/conditional) of recommendations but also on the certainty of underlying evidence.<sup>2,3</sup> To obtain the best estimates of intervention benefits, harms and burdens and assess the credibility of the evidence, guidelines require rigorously conducted systematic reviews.

Once guideline panels have such reviews available, they must interpret results, which requires both clinical and methodological expertise, and consider patients' values and preferences.<sup>4</sup> When panels have conducted the process optimally, healthcare professionals can confidently consider

recommendations to help make decisions in complex clinical circumstances, evaluate different treatment options and through shared decision-making facilitate patients' choice of clinical management.

This usefulness of guidelines presupposes that they have been rigorously developed and, thus, meet standards of trustworthiness. Historically, very few if any guidelines adhered to these standards; we sometimes refer to this traditional approach as GOBSAT—good old boys sitting around a table. Many guidelines still fail to adhere to trustworthiness standards, still looking more like GOBSAT.<sup>5–7</sup>

It is, therefore, essential that healthcare professionals—and, for system-level guidelines, policymakers—are able to distinguish guidelines that are trustworthy from those that are not. The aim of this paper is to inform clinicians and policymakers on how to determine guideline trustworthiness. To do so, we offer eight domains (see [table 1](#)) in which a lack of rigour can compromise guideline trustworthiness.

## Distinguish between trustworthy evidence-based and untrustworthy guidelines

### Clarity

One of the important conditions for evidence-based recommendation is clarity and, related to clarity, specificity and actionability.<sup>8,9</sup> Useful recommendations are easy to follow, avoid ambiguous language and provide clear direction (eg, in favour or against an intervention) and strength (ie, the level of confidence in the benefits outweighing potential harms and burdens). Panels will develop clear recommendations using the patient/intervention/comparator/outcome (PICO) format, a well-accepted and effective structure for framing the question.<sup>10</sup> When the panel starts with the PICO question, guideline end-users can quickly understand the patient population, intervention and comparators involved in a recommendation. Clearly presented guidelines will not only make explicit the target



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada

### Correspondence to

João Pedro Lima;  
limaj1@mcmaster.ca

**Table 1** Differences between trustworthy evidence-based guidelines from untrustworthy guidelines

	Trustworthy guidelines	Untrustworthy guidelines
Clarity	Clear and actionable recommendation	Ambiguous recommendation
Panel composition	Multidisciplinary panel including all key components	Panel does not include a relevant perspective (eg, patients)
Conflicts of interest	Declared and, when appropriate, managed	Conflicts not declared and/or managed
Outcome selection and prioritisation	Using PICO format with outcomes important to patients	Failure to specify the key PICO elements and use the board panel concern than patients-important outcomes.
Summary of evidence	Recommendations based on systematic review with the best evidence available	GOBSAT; poor-quality systematic review
Values and preferences	Transparent consideration of patients' values and preferences	Ignores patients' values and preferences
Strength of recommendation	Matches trade-off of up and downsides, rating the certainty of the evidence and variability of values and preferences	Ignores trade-off of up and downsides, rating the certainty of the evidence and/or variability of values and preferences
Presentation and rationale for recommendations	Supporting evidence presented as absolute risks followed by certainty of the evidence in a user-friendly manner (SoF table)	Supporting evidence presented as relative risks with no or poorly assessed the certainty of the evidence not.

GOBSAT, good old boys sitting around a table; PICO, patient/intervention/comparator/outcome; SoF, summary of findings.

patient population but also whether there are subgroups within the population that may benefit more or less from the interventions under consideration.<sup>4</sup> When guidelines provide clear recommendations, healthcare professionals can easily understand and interpret them. Clear recommendations, therefore, facilitate increased adherence and consistent application of evidence-based practices.

### Panel composition

The process of developing guidelines often involves conflicting opinions, and, thus, negotiation among panel members. The composition of the guideline panel can, therefore, influence the recommendations, and an optimal outcome requires incorporating perspectives from the right mix of stakeholders.

The composition across guideline panels may vary, but panels typically consist of 10–25 members who, for guidelines to be trustworthy, must constitute a multidisciplinary group drawn from a minimum of four constituencies (often referred to as stakeholders).<sup>10–12</sup> These constituencies include two groups of clinicians: clinical experts, often research leaders, and clinicians who spend most of their time involved in patient care. If guidelines are directed both at primary care and specialist physicians, the panel should include both generalists and subspecialists.<sup>10–12</sup> Clinicians involved in the care of patients often come from allied health professions, including nurses, physiotherapists, occupational therapists, social workers, speech and language pathologists and others.

To ensure optimal understanding of the relevant evidence, trustworthy guidelines also require participation from one or more methodologists. Such individuals will ideally have expertise in aspects of guideline development beyond interpreting the evidence, including group

process and the ultimate framing of recommendations. Depending on the scope and topic, other non-clinical disciplines may include health economists (for guidelines that address resource use or costs), public health experts or other decision-makers (for guidelines that address public health or systems issues) and ethicists (for guidelines in which issues such as equity are highly relevant).<sup>10–11</sup>

The presence of people with lived experience of the condition under consideration will bring another perspective to the panel. Such individuals are often mistakenly seen as representing patient views: it is impossible for a small number of individuals to provide such representation, and impossible for a guideline development group to ensure that the individuals they recruit are in any way representative of population values and preferences. Patient partners are, thus, recruited not necessarily to have more insight into typical patient values and preferences than other panel members (they may have more insight but may often have less) but rather to bring in patient perspectives that may otherwise be neglected.

Finally, within those constituencies, it is also important to ensure geographical representativeness. In other words, the panel should include individuals from the geographical regions where recommendations are applicable. International guidelines, for instance, will include panellists not only from European and North American countries but also from developing countries. Gender balance is another aspect that should be considered, especially in topics where the perspectives of some subgroup populations are essential (eg, PrEP adherence, gynaecological cancer, etc). In summary, it is recommended to consider geographical representation and gender

balance because expert groups should not only encompass professional diversity but also reflect geographical and gender equality.

### Conflicts of interest

Both financial and non-financial conflict of interest (COI) may bias clinical practice guideline recommendations.<sup>10</sup> COI may detrimentally influence multiple steps of the guideline development process, including scoping and framing the key questions, selecting choice of comparisons, interpreting the evidence and developing and presenting recommendations.<sup>13</sup>

Ideally, guidelines should exclude panellists with appreciable COI. However, in some circumstances, the panel may not be able to perform its work without members who have COIs. Under such circumstances, members with COIs should represent only a minority of the guideline development panel (GDP) and those with the most serious COI should recuse themselves from the discussion and voting on recommendations on which they are conflicted.<sup>10 11 13 14</sup>

In trustworthy guidelines, all panel members will complete and agree to publish the summary of their disclosed interests in both financial and non-financial conflicts. Moreover, guidelines must have a plan for managing conflicts that do exist (such as selective exclusion or labelling of conflicts at the time of discussion).<sup>10 15</sup> Finally, due to the timeframe in which a guideline is often developed, sometimes as long as 2 years, COIs must be continuously monitored and updated.

### Outcome selection and prioritisation

The objective of any recommendation is to improve patient-important outcomes. Thus, selecting and prioritising the outcomes are critical to produce a trustworthy guideline. The panel must prespecify the important outcomes in both benefits (desirable outcomes) and harms and burdens (undesirable outcomes) that they will consider.<sup>10 16 17</sup> The importance of these outcomes may vary according to the perspective of patients, clinicians or policymakers.<sup>16</sup> When the target audiences for a guideline are clinicians and patients, the guideline should consider the patient-important outcomes, as patients are the primary group that will benefit—or not—from incorporation of guideline recommendations into decision-making.<sup>16 17</sup>

In some cases, when patient-important outcome events are rare or occur over long periods of time, guideline panels turn to surrogate outcomes as a substitute for patient-important outcomes.<sup>16 17</sup> Such outcomes include glucose for macrovascular and microvascular complications of diabetes, bone density for fractures and cognitive function for behaviour in patients with dementia. Unfortunately, there are a myriad of examples in which an intervention has modified a surrogate outcome in what should be a positive way only to find no improvement—and in quite a few instances, a deterioration—in the patient-important outcome for which the surrogate

is standing in.<sup>17–19</sup> Thus, the panel should consider surrogate outcomes only when data on patient-important outcomes are lacking<sup>16</sup> and recognise that when they do rely on surrogates, the quality or certainty of the resulting evidence will be lower, and sometimes very much lower, than if studies had measured patient-important outcomes directly.<sup>20</sup>

### Summary of evidence

Trading off benefits versus harms and burdens is the core of a guideline panel's job. Doing so requires best evidence summaries of the magnitude of effects on all important outcomes and an assessment of the certainty of the evidence. In trustworthy guidelines, this summary is informed by rigorously conducted systematic reviews, often with meta-analyses providing single best estimates of effect.<sup>4</sup>

Rigorous systematic reviews include explicit eligibility criteria, a comprehensive search for eligible studies, assessment of risk of bias of individual studies, with judgement of eligibility and risk bias conducted in duplicate.<sup>21</sup> Rigorous reviews also involve judgements regarding the certainty or quality of the evidence, from high to very low certainty. Ideally, that judgement will use the rigorously developed and widely adopted the grading of recommendations assessment, development, and evaluation (GRADE) approach in which randomised trials begin as high-certainty evidence and observational studies as low-certainty evidence in a four-category system of certainty of evidence (high, moderate, low and very low).<sup>22</sup>

Beyond study design, GRADE has identified five domains that may lead to rating down the certainty of evidence: risk of bias, inconsistency, indirectness, imprecision and publication bias. Reviewers may rate up the certainty of evidence from observational studies, primarily for large or very large magnitude of effect. If a panel is fortunate, rigorous reviews will already be available; if not, they will have to commission or conduct their own.

Without systematic reviews, the evidence summaries become untrustworthy, and guidelines regress to the GOBSAT approach. Systematic reviews following methodological standards are, therefore, essential for guideline trustworthiness.

### Values and preferences

A key principle of evidence-based medicine posits that clinical decision-making must consider patients' values and preferences.<sup>23</sup> Values and preferences are also a key determinant of the direction and strength of a recommendation—guideline panels make strong recommendations in favour of an intervention when patients place much higher value on the benefits than the associated harms and burdens, strong recommendations against an intervention when the opposite is true.

Trustworthy evidence-based guidelines will make explicit the values and preferences underlying their recommendations. Ideally, to inform the values and preferences of the average patient on the outcomes of



interest, the panel will seek and find systematic reviews summarising studies of patient values and preferences.<sup>24</sup> In the absence of formal studies of values and preferences, clinical experts can consider their experience in shared decision-making. Other panel members can consider their experience with colleagues, friends and other patients. A panel may also conduct a focus group to generate their own relevant evidence.<sup>25</sup> Methods are now available to help guideline panels clarify their views of patient values and preferences.<sup>26</sup>

The greater variability among patients in their values and preferences, and the less confident panels are regarding typical values and preferences, the greater the uncertainty regarding the balance between desirable and undesirable outcomes of an intervention, and the more likely a panel will issue a weak or conditional rather than a strong recommendation.<sup>24</sup> Thus, evidence regarding benefits and harms, in combination with evidence regarding patients' values and preferences, will eventually inform panellists when formulating a recommendation.

### Strength of recommendation

According to the GRADE system, recommendations can be categorised as strong or weak/conditional recommendations against or in favour of an intervention.<sup>22 24</sup> Following GRADE guidance, guideline panels will present their strong recommendations as 'We recommend (...)' and conditional recommendations with 'We suggest (...)'. In deciding on the strength of recommendation, panels will consider the magnitude of benefits, harms and

burdens, the certainty of evidence and patient values and preferences.

While a large gradient between benefits versus harms and burdens warrants a strong recommendation, a close balance is likely to be associated with a conditional recommendation. Regarding certainty of evidence, with some few exceptions, low or very low certainty of evidence warrants a conditional recommendation: thus, the higher the certainty of evidence, the more likely is a strong recommendation. Finally, both large variability and uncertainty regarding patient values and preferences will influence the strength of recommendations: the larger the uncertainty or variability, the more likely a weak or conditional recommendations on those values and preferences. Additional considerations—addressed in the next section—may sometimes influence the strength of recommendations.

### Presentation and rationale for recommendations

The way a guideline presents its recommendations is crucial to obtain optimal transparency. Patients, clinicians and other stakeholders require not only a transparent and simple presentation of the evidence but also an explanation of how this evidence led to the recommendation formulated by the panel. GRADE offers two tools to achieve this transparency—the summary of findings (SoF) tables and the Evidence to Decision (EtD) framework.<sup>22 27</sup>

GRADE's SoF table organises in a succinct manner the relative effect of interventions, their absolute effect,

Remdesivir vs No remdesivir Patients with severe covid-19					
Outcome	Study results and measurements	Absolute effect estimates		Certainty of the Evidence (Quality of evidence)	Plain language summary
		No remdesivir	Remdesivir		
Mortality	Odds ratio 0.89 (CI 95% 0.78 — 1.02) Based on data from 6631 participants in 5 studies	130 per 1000	117 per 1000	Low Due to serious imprecision and inconsistency	Remdesivir may reduce mortality <div>Intervention</div>
		Difference: 13 fewer per 1000 (CI 95% 26 fewer — 2 more)			
Mechanical ventilation	Odds ratio 0.87 (CI 95% 0.77 — 0.99) Based on data from 6620 participants in 5 studies	116 per 1000	102 per 1000	Moderate Due to serious imprecision	Remdesivir probably reduces mechanical ventilation <div>Intervention</div>
		Difference: 14 fewer per 1000 (CI 95% 24 fewer — 1 fewer)			
Time to symptom improvement	Lower better Based on data from 2599 participants in 2 studies	9.9 days (Mean)	9.2 days (Mean)	Moderate Due to serious imprecision	Remdesivir probably has little or no impact on time to symptom improvement <div>No imp. diff.</div>
		Difference: 0.7 fewer (MD) (CI 95% 1.8 fewer — 0.6 more)			

**Figure 1** Example of GRADE's SoF table.<sup>28</sup> GRADE, grading of recommendations assessment, development, and evaluation; SoF, summary of findings.

and the certainty of evidence for every patient-important outcome previously selected and prioritised. The SoF table presents the absolute difference between intervention and comparator as number of events per 1000 for dichotomous outcomes and, for continuous outcomes, as mean difference.

Figure 1, for example, shows an example of a SoF table that supported recommendations in the living guideline on drugs for COVID-19.<sup>28</sup> On the top of the table, one can find information about the respective population (patients with severe COVID-19), intervention (remdesivir) and comparator (no remdesivir). The left column presents the outcomes, followed by relative effects, number of participants and studies included and absolute effect estimates. The following column shows the certainty of evidence rated for each outcome. Finally, the far-right column shows the plain language summary, which is a short sentence that conveys the direction and magnitude of effect and the certainty of evidence.<sup>29</sup> In the footnotes, authors may provide explanations for rating down the certainty of evidence.

The GRADE EtD framework allows readers to better understand what considerations led panellists to their final recommendation. In the SoF, the panel presents the magnitude of effect on desirable and undesirable outcomes and the overall certainty of evidence for each outcome. Justification of the direction and strength of recommendations will include succinct statements regarding underlying values and preferences, and the associated uncertainty and variability. Panels may also consider, in making their recommendations, issues of resource use, feasibility, acceptability and equity. Those judgements will assist in understanding how the panel moved from the evidence to a recommendation and, thus, informed their clinical, public health or policy recommendations.

While trustworthy guidelines will apply the tools and presentation aspects, we have summarised untrustworthy guidelines often fail to follow those strategies.

## CONCLUSION

Disparities between trustworthy evidence-based guidelines and less trustworthy guidelines frequently lie in one of the eight domains (table 1). Lack of clarity; a panel composition lacking diversity; lack of transparency regarding conflicts of interest; unclear or inappropriate clinical questions; failure to use a rigorously conducted summary of the evidence; failure to explicitly consider patient's values and preferences; inappropriate judgement of the strength of the recommendations and, finally, a complex or absent presentation of the judgements that led the panel to a recommendation—can potentially jeopardise the development and credibility of a practice guideline. For example, in the management of patients with heart failure, all the current clinical practice guidelines clearly describe the goals and target audience of the guidelines. However, most of them failed to describe the

panel composition, use the systematic review and rate the certainty of the evidence for the synthesis of the recommendations and trade-off between desirable and undesirable outcomes of an intervention.<sup>30</sup>

We advise readers that those eight domains should only serve as a framework for trustworthiness determination instead of a fixed tool to rule in or rule out guidelines. Trustworthiness is a continuum, and each dimension can be fully or partly met. The greater the extent to which a guideline meets all eight domains, and does so fully, the greater the trustworthiness.

Even though for guidelines that do not rely on the GRADE approach, some of these domains can be slightly or, maybe, significantly different, the core principle should be the same—the panel needs to be transparent when showing how the evidence led to a specific recommendation. Other methods may substitute, for example, SoF tables and EtD frameworks, but transparency still needs to be present. It is evident, however, that clinicians and policymakers will find more challenging and will spend more time trying to find these details in a guideline that does not follow the GRADE approach.

**Twitter** João Pedro Lima @JoaoPLogoLima

**Acknowledgements** This publication was supported by the Einstein Foundation Berlin as part of the Einstein Foundation Award for Promoting Quality in Research. The contents are those of the authors and do not necessarily represent the official views of, nor an endorsement by, the Einstein Foundation or the award jury.

**Contributors** JPL and WT equally contributed to this work. Study conception and design: JPL and GHG. Draft manuscript: JPL and WT. Reviewed and approved the final version of the manuscript: JPL, WT and GHG.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not applicable.

**Ethics approval** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## ORCID iD

João Pedro Lima <http://orcid.org/0000-0002-5486-7211>

## REFERENCES

- 1 Woolf SH, Grol R, Hutchinson A, *et al*. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ* 1999;318:527–30.
- 2 Guyatt GH, Oxman AD, Vist GE, *et al*. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.
- 3 Andrews J, Guyatt G, Oxman AD, *et al*. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *J Clin Epidemiol* 2013;66:719–25.
- 4 Brignardello-Petersen R, Carrasco-Labra A, Guyatt GH. How to interpret and use a clinical practice guideline or recommendation: users' guides to the medical literature. *JAMA* 2021;326:1516–23.

- 5 Miller J, Petrie J. Development of practice guidelines. *Lancet* 2000;355:82–3.
- 6 Grilli R, Magrini N, Penna A, *et al.* Practice guidelines developed by specialty societies: the need for a critical appraisal. *Lancet* 2000;355:103–6.
- 7 Zerah L, Bun RS, Guillo S, *et al.* A prescription support-tool for chronic management of oral antithrombotic combinations in adults based on a systematic review of international guidelines. *PLoS One* 2019;14:e0211695.
- 8 Guyatt GH, Alonso-Coello P, Schünemann HJ, *et al.* Guideline panels should seldom make good practice statements: guidance from the GRADE working group. *J Clin Epidemiol* 2016;80:3–7.
- 9 Schünemann HJ, Zhang Y, Oxman AD, *et al.* Distinguishing opinion from evidence in guidelines. *BMJ* 2019;366:l4606.
- 10 World Health Organization. *WHO handbook for guideline development, 2nd ed.* Geneva: World Health Organization, 2014.
- 11 Graham R, Mancher M, Miller Wolman D, *et al.*, eds. Institute of medicine Committee on standards for developing trustworthy clinical practice G. In: *Clinical Practice Guidelines We Can Trust. Copyright 2011 by the National Academy of Sciences. All rights reserved.* Washington (DC): National Academies Press (US), 2011.
- 12 Kunz R, Fretheim A, Cluzeau F, *et al.* Guideline group composition and group processes: article 3 in integrating and coordinating efforts in COPD guideline development. *Proc Am Thorac Soc* 2012;9:229–33.
- 13 Guyatt G, Akl EA, Hirsh J, *et al.* The vexing problem of guidelines and conflict of interest: a potential solution. *Ann Intern Med* 2010;152:738–41.
- 14 Boyd EA, Akl EA, Baumann M, *et al.* Guideline funding and conflicts of interest: article 4 in integrating and coordinating efforts in COPD guideline development. *Proc Am Thorac Soc* 2012;9:234–42.
- 15 Qaseem A, Forland F, Macbeth F, *et al.* Guidelines international network: toward international standards for clinical practice guidelines. *Ann Intern Med* 2012;156:525–31.
- 16 Guyatt GH, Oxman AD, Kunz R, *et al.* GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol* 2011;64:395–400.
- 17 Wilt TJ, Guyatt G, Kunz R, *et al.* Deciding what type of evidence and outcomes to include in guidelines: article 5 in integrating and coordinating efforts in COPD guideline development. *Proc Am Thorac Soc* 2012;9:243–50.
- 18 Uhlig K, Leff B, Kent D, *et al.* A framework for crafting clinical practice guidelines that are relevant to the care and management of people with multimorbidity. *J Gen Intern Med* 2014;29:670–9.
- 19 Bucher HC, Guyatt GH, Cook DJ, *et al.* Users' guides to the medical literature: XIX. applying clinical trial results. A. how to use an article measuring the effect of an intervention on surrogate end points. *JAMA* 1999;282:771–8.
- 20 Guyatt GH, Oxman AD, Kunz R, *et al.* GRADE guidelines: 8. rating the quality of evidence—Indirectness. *J Clin Epidemiol* 2011;64:1303–10.
- 21 Murad MH, Montori VM, Ioannidis JPA, *et al.* How to read a systematic review and meta-analysis and apply the results to patient care: users' guides to the medical literature. *JAMA* 2014;312:171–9.
- 22 Guyatt G, Oxman AD, Akl EA, *et al.* GRADE guidelines: 1. introduction—GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64:383–94.
- 23 Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet* 2017;390:415–23.
- 24 Andrews JC, Schünemann HJ, Oxman AD, *et al.* GRADE guidelines: 15. going from evidence to recommendation—determinants of a recommendation's direction and strength. *J Clin Epidemiol* 2013;66:726–35.
- 25 Goodman SM, Miller AS, Turgunbaev M, *et al.* Clinical practice guidelines: incorporating input from a patient panel. *Arthritis Care Res (Hoboken)* 2017;69:1125–30.
- 26 Zeng L, Helsing LM, Bretthauer M, *et al.* A novel framework for incorporating patient values and preferences in making guideline recommendations: guideline panel surveys. *J Clin Epidemiol* 2023.
- 27 Alonso-Coello P, Oxman AD, Moberg J, *et al.* GRADE evidence to decision (Etd) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: introduction. *BMJ* 2016;353:i2089.
- 28 Lamontagne F, Agarwal A, Rochweg B, *et al.* A living WHO guideline on drugs for COVID-19. *BMJ* 2020;370:m3379.
- 29 Santesso N, Glenton C, Dahm P, *et al.* GRADE guidelines 26: informative statements to communicate the findings of systematic reviews of interventions. *J Clin Epidemiol* 2020;119:126–35.
- 30 Sephien A, Dayto DC, Reljic T, *et al.* A systematic critical appraisal of clinical practice guidelines in heart failure using the AGREE II tool. *Am J Cardiol* 2023;202:192–8.