

Clinical practice guidelines: the need for greater transparency in formulating recommendations

Ian A Scott and Gordon H Guyatt

Clinical practice guidelines (CPG) serve to guide clinical practice and inform quality improvement programs by generating clinical standards and performance measures. Clinicians will only use guideline recommendations if they perceive them to be evidence-based, unambiguous, and feasible within routine care.¹ Despite the proliferation of CPGs, direct evidence of impact on quality of care or patient outcomes is limited.² Several explanations exist,³ including perceptions of bias in the development of guideline recommendations.⁴ Recently published claims⁵ and counter-claims⁶ of bias in Australian guidelines⁷ and related position statements⁸ regarding the management of acute coronary syndromes highlight problematic issues in CPG development. These include conflicts of interest (COI) of guideline panellists, validity and strength of recommendations, and involvement of external stakeholders and end users. We offer strategies for dealing with these issues in a transparent and explicit manner.

Challenges in CPG development and possible solutions

Conflicts of interest and panel representativeness

The quality of CPG bears little relation to the level of seniority or expertise of guideline authors.⁹ Guideline panellists often harbour COI that may not be fully evident, even to the panellists themselves, but which can potentially bias their recommendations.¹⁰ These conflicts include not only financial ties with industry but also practice reimbursement incentives, professional affiliations and practice specialisation, intellectual attachment to their own studies, ideas and innovations, and desire for academic recognition and career advancement.¹¹ The most entrenched conflict can be a disinclination to challenge or reverse strongly held beliefs. Using research evidence to make recommendations requires subjective interpretations, which will be influenced by the value structure of panel members.¹² Vulnerability to preconceptions is greatest for recommendations based on low-quality evidence — an increasingly frequent occurrence in contemporary CPG¹³ — although recommendations based on high-quality evidence are far from invulnerable.

Most current guidelines remain susceptible to COI, which can impinge on all stages of the CPG development process (Box 1). Many are published without peer review or, if contained in journal supplements, escape the standard of peer review applied to articles published in the parent journal.¹⁴ Moreover, many guidelines (79% in a recent survey of Australian guidelines¹⁵) fail to mention possible competing interests of guideline panellists. Even if COI are disclosed, guideline users may not adjust their perceptions of recommendations in response to such disclosures.¹⁶

Strategies for dealing with COI are outlined in Box 1,¹⁷⁻¹⁹ with key strategies being:

- Nominated panellists must disclose all industry-related professional activities, including research grants and speaker support, and, for the duration of guideline development, divest themselves of direct financial interests (stock ownership, board positions,

ABSTRACT

- A recently published critique of a set of Australian clinical practice guidelines (CPG) highlighted problematic issues in guideline development concerning conflicts of interest of guideline panellists, validity and strength of recommendations, and involvement of end users and external stakeholders.
- Management of financial or intellectual conflicts of interest requires: full disclosure; limitations on industry or agency financial support during guideline development; a representative panel that includes conflict-free members; and only conflict-free panellists to be involved in drafting guideline recommendations.
- Guideline panels should consider adopting the GRADE (Grading of Recommendations Assessment, Development and Evaluation) system to assist in determining the validity and strength of recommendations.
- Guideline panels should seek formal feedback from external stakeholders and end users.
- Enacting such policies aims to lend greater transparency and credibility to CPG, limit protracted and unhelpful interpretive debates, and promote wider use of CPG.

MJA 2011; 195: 29–33

consultancy agreements) in commercial companies with an interest in any guideline recommendation.

- Panellists are required to identify all sections of the draft guidelines for which they have COI. These conflicts are recorded in a COI grid maintained by the guideline chairperson.
- Methodologists free of financial or intellectual conflicts of interest share responsibility with content experts for collecting and interpreting evidence.
- Explicit processes must be used to assess evidence quality and link this directly with strength of recommendations.
- Only conflict-free panellists (both methodological and content experts) are involved in determining the direction (for or against a specific clinical action) and strength of recommendations.
- Lack of consensus around evidence quality or recommendations is resolved by explicit democratic processes (such as Delphi rounds and nominal group techniques) involving conflict-free panellists who have thoroughly reviewed the related evidence.
- Individuals should be invited to join guideline panels through an open, transparent application process centred on selection criteria that ensure an appropriate balance of content and methodological expertise. Such criteria may comprise extent of clinical experience with the topic in question, prior participation in undertaking critically appraised literature reviews, intended commitment in time and intellectual input into the guideline development process, and referee reports. For guidelines that deal with common conditions and are aimed at large, multidisciplinary

audiences, panel composition should reflect the spectrum of end users and avoid being dominated by a narrow spectrum of specialists.

The impact on guideline content if such policies were implemented and enforced is yet to be empirically determined,²⁰ but many organisations involved in guideline development have now adopted at least some of them as best practice for reducing the probability of conflicted panellists having undue influence.¹⁷⁻¹⁹

Credibility and methodological rigour

Clinicians lose confidence in CPG when separate guidelines on the same clinical topic from seemingly authoritative sources produce conflicting recommendations. For example, United States and European CPGs differ in their recommendations for use of anti-coagulants in acute coronary syndromes.²¹ The ways in which guideline panellists have interpreted and weighted the evidence and used it to formulate recommendations of different strength must be clearly communicated.

While various systems exist for rating evidence according to hierarchies of study design, with randomised controlled trials (RCTs) at the top, most contain no explicit processes for assessing evidence quality or linking it with recommendations.²² The Grading of Recommendations Assessment, Development and Evaluation (GRADE) system (Box 2) attempts to meet this need²³ and has advantages over other grading systems in that it:

- clearly separates quality of evidence and strength of recommendations;
- explicitly evaluates alternative management strategies;
- provides clear-cut, detailed criteria for downgrading and upgrading quality of evidence ratings related to different outcomes that are of importance to patients and which exclude surrogates;
- provides a transparent process for moving from evidence to recommendations and grading recommendations as strong or weak on the basis of clearly defined, pragmatic interpretive criteria;
- explicitly acknowledges patient preferences; and
- details potential resource use.

Adopting the GRADE system may assist in exposing and mitigating bias arising from COI, thus augmenting COI policies. Applying the GRADE quality of evidence classification to contemporary CPG suggests that many (in one study almost 50%²⁴) RCT-derived recommendations fail to meet a-priori definitions of high-quality evidence. Several examples now exist of how the GRADE system promotes the development of CPG recommendations that are more aligned with evidence quality.^{25,26}

More than 50 organisations worldwide have adopted the GRADE system, including the World Health Organization, the American College of Physicians, the Cochrane Collaboration, the Scottish Intercollegiate Guidelines Network (SIGN) and UpToDate (online clinical decision support system). In Australia, the National Health and Medical Research Council (NHMRC) has recently produced a revised draft schema for more explicit, structured grading of evidence quality and strength of recommendations which has several similarities — and differences — to the GRADE system.²⁷

Applicability

Guidelines commonly base their recommendations on trials involving selected populations and standardised interventions. These may not be applicable to unselected populations receiving

care from clinicians working under real-world constraints.³ Benefits reported in trials may not be reproducible in patient groups, such as older patients with multiple comorbidities, that are underrepresented in such studies or in clinical settings very different to those used in trials.²⁸

Guidelines tend to focus on single clinical conditions in isolation (such as heart failure or acute coronary syndromes) and may not adequately address situations where the management of commonly encountered comorbidities (such as asthma, diabetes or dementia) may conflict with, or override, recommendations for the index condition.

Guideline panellists should assess the extent to which evidence of treatment benefit is consistent, or even exists, across different populations with different comorbidity spectra, in different settings and with different modes of treatment administration. The circumstances under which the magnitude of treatment benefit (and harm) is significantly enhanced or attenuated should be highlighted in the way recommendations are presented. Recommendations should, where appropriate, stratify populations according to disease risk and target treatments to those who will experience greatest net benefit.²⁹

Ideally, in developing the guideline, panels should seek feedback from a separate reference group of front-line clinicians who are likely to use the guidelines (also chosen by an open application process and subject to the same disclosure policies as panellists) regarding the impacts and feasibility of guideline recommendations and the extent to which proposed guidelines meet their care needs.

Impacts on health care systems

Guideline authors must avoid exercising power without responsibility in obliging clinicians and health services to enact recommendations and satisfy guideline-based performance measures with little regard to the added problems and pressures these may engender in terms of professional interactions, team functioning, organisational predispositions, resource availability and medico-legal considerations.^{3,30}

Following guideline release, a process of public consultation should exist (as it does for NHMRC guidelines) that allows a representative cross-section of health managers, quality improvement experts, and patient support groups to provide feedback on the wider environmental implications of specific CPG recommendations. In future iterations of the guideline, authors should respond to comments by reviewing and, where appropriate, modifying recommendations that have been identified as particularly problematic. Guideline authors may also consider asking other agencies to undertake formal cost-effectiveness analyses or modelling exercises related to recommendations that raise concerns about the availability and cost of resources.

Conclusion

In developing guidelines, transparent processes are needed that deal with potential COI, rate the quality of evidence and strength of recommendations, and address real-world needs of guideline users. The strategies outlined here, if adopted by guideline panels, may limit protracted interpretive debates and correct deficiencies that inhibit a wider use of CPG. While they potentially impose more effort, cost and delay in developing guidelines, we believe these imposts are outweighed by the minimisation of recom-

1 Steps in developing clinical practice guidelines (CPGs), potential conflicts of interest (COI) and potential solutions¹⁷⁻¹⁹

| Development steps and responsible body | Potential COI | Potential solutions |
|---|---|---|
| Sponsor to the guideline panel | | |
| Select topic for CPG | Panels dependent on industry funding propose topics that attract it | Public or professional sponsor nominates topics by explicit processes and recognised need. Prior NHMRC endorsement sought |
| Obtain financial and other forms of support | Panels at risk of undue influence from commercial sources if dependent on industry funding | No direct funding of guideline development groups from industry; funding from public or professional sources Industry only provides funding at arm's length to assist in printing or distribution of guideline at the sole discretion of public or professional sponsor Guidelines for which there is a recognised need but for which funding from professional sources is unavailable are funded by government or not-for-profit organisations |
| Nominate and appoint guideline panels | Commercial sponsor nominates panellists with whom it has pre-existing ties or relationships Guideline panels dominated by panellists who have pecuniary or other interests in guideline recommendations | Panellists only nominated by professional bodies Panellists appointed by an open and transparent application process with explicit selection criteria Panels include individuals with a range of professional and other backgrounds who adequately represent all clinical stakeholder groups |
| Guideline panel | | |
| Develop a governance structure and work plan | Guideline chair, co-chairs and panellists who have COI are appointed or elected | Adopt policies of open and full disclosure of COI, and require panellists to divest themselves of direct financial interests in commercial companies with a potential interest in any guideline recommendation |
| Specify clinical issues and outcomes of interest | Guideline content skewed towards issues and outcomes favoured by panellists with COI | Members of any craft group with a pecuniary interest are confined to an advisory role rather than being principal authors Guideline content and key issues are subject to roundtable discussion involving all panellists at a planning meeting following appointment of chair and co-chair |
| Conduct systematic review of the literature and rate the quality of evidence | Guideline panellists with COI are biased in their selection of studies, interpretation of results and rating of the quality of evidence | Systematic review of relevant evidence using standardised methods for selecting studies, analysing and rating evidence quality, identifying and evaluating benefits and harms, and presenting conclusions Methodologists involved in selecting, synthesising, interpreting and rating quality of evidence Explicit, formal system for grading quality of evidence |
| Draft recommendations | Panellists with COI are biased in the strength they confer to specific recommendations | Formal system for grading strength of recommendations according to quality of evidence, desirable and undesirable consequences of clinical actions, patient values and preference, and resource use Conflicted panellists abstain from drafting recommendations Explicit processes for resolving disagreements and achieving consensus |
| Guideline panel or sponsor | | |
| Distribute a draft for external and internal reviews | Panels or sponsor choose reviewers with whom they have pre-existing ties or relationships, and who may not be representative of guideline end users | Representative panel of guideline end users assembled by professional or public sponsor to critically review draft guidelines. No participants with COI to undertake review |
| Reviewers | | |
| Guideline review by external and internal reviewers | Reviewers with COI less critical of guideline development processes | No participants with COI to undertake review |
| Sponsor or journal or other bodies with direct interest in guideline as a reference document | | |
| Publish or disseminate the guideline | Commercial sponsor combines guideline dissemination with marketing strategies aimed at increasing sales Journal does not subject guideline to same level of rigorous peer review as original research Journals and guideline clearinghouses may not require disclosure of COI and funding sources of guideline panellists | Dissemination policy determined by professional or public sponsor; any association with commercial marketing campaigns explicitly forbidden All guidelines subject to the same level of peer review as original research published in journals Journal and guideline clearinghouse editors require all guidelines accepted for publication to describe COI policies, sources and amounts of funding for guideline, and relevant financial interests of panellists |
| Sponsor or guideline panel | | |
| Monitor new research findings and determine when a guideline should be updated | Commercial sponsor or panellists with COI slow in responding to new evidence that necessitates substantial revision of recommendations which may result in reduced product sales | Timing of guideline update should be determined by professional or public sponsor in response to ongoing monitoring of new evidence by non-conflicted guideline panellists who maintain a watching brief |

2 The Grading of Recommendations Assessment, Development and Evaluation (GRADE) system²³

GRADE proposes that the quality of evidence associated with each outcome of importance to patients be evaluated separately. The GRADE system classifies quality of evidence into 4 levels: high, moderate, low or very low. Evidence from randomised controlled trials (RCTs) begins as high quality, but may be rated down if trials demonstrate one of five categories of limitations. Observational studies begin as low-quality evidence, but may be rated up if associated with one of three categories of special strengths.

| Study design | Quality of evidence | Lower/higher level of quality if: |
|---------------------|--|---|
| Randomised trial | <ul style="list-style-type: none"> High (further research is very unlikely to change our confidence in the estimate of effect) Moderate (further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate) | <ul style="list-style-type: none"> Risk of bias (serious [-1]; very serious [-2]) Inconsistency (serious [-1]; very serious [-2]) Indirectness (serious [-1]; very serious [-2]) Imprecision (serious [-1]; very serious [-2]) Publication bias (likely [-1]; very likely [-2]) Large effect (large [+1]; very large [+2]) Evidence of a dose-response gradient (+1) All plausible confounding <ul style="list-style-type: none"> ➤ would reduce a demonstrated effect (+1) ➤ would suggest a spurious effect when results show no effect (+1) |
| Observational study | <ul style="list-style-type: none"> Low (further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate) Very low (any estimate of effect is very uncertain) | |

Reasons for rating down quality of evidence

Risk of bias: Quality will be lower if most of the evidence from available RCTs is compromised by limitations such as: lack of allocation concealment; lack of blinding (particularly if outcome assessment is highly susceptible to bias); large losses to follow-up; failure to analyse patients in the groups to which they were randomised; premature termination for benefit; or failure to report outcomes (often those for which no effect was observed).

Inconsistent results: Widely differing estimates of treatment effect across studies suggest true differences in underlying treatment effect, and if investigators fail to identify a plausible explanation, quality of evidence decreases. Variability may arise from differences in populations, interventions, or outcomes.

Indirectness of evidence: In comparing effects of two active treatments, randomised head-to-head trials constitute high-quality evidence. Indirect comparisons of the magnitude of effects seen in separate placebo-controlled trials of each treatment constitute lower quality evidence. Another type of indirectness arises if there are important differences between the populations (eg, elderly v non-elderly), interventions (eg, low v high dose) and outcomes (patient-important v surrogate) measured in trials and those under consideration in the guideline.

Imprecision: When studies include relatively few patients and few events and thus have wide confidence intervals, quality of evidence decreases.

Publication bias: Failure to report studies that typically show no effect reduces evidence quality. Such publication bias is more likely when only a small number of trials, all funded by industry, are available.

Reasons for rating up quality of evidence

Large and consistent effect sizes: If several large and methodologically strong observational studies report a very large effect size and confounding is unlikely to explain all or most of the apparent benefit, quality of evidence can be rated up (eg, hip replacement in severe osteoarthritis or dialysis for end-stage renal failure).

Presence of a dose-response gradient: Where intensity of intervention (eg, dose, duration, or parenteral v oral method of administration) shows a correlation with effect size, the quality of evidence may increase.

Accounting for all plausible confounding: Where investigators have accounted for all plausible biases which might decrease the magnitude of an apparent effect or create a spurious effect when results show no effect, the quality of evidence increases.

Grading strength of recommendations

The GRADE system grades recommendations as "strong" or "weak" based on four determinants: quality of evidence, trade-off between desirable and undesirable consequences, variability in patient values and preference, and resource use. When desirable effects of an intervention clearly outweigh undesirable effects, or vice versa, and estimates are based on high-quality evidence, the recommendation is strong. When trade-offs are less certain (lower quality evidence or desirable and undesirable effects closely balanced), the recommendation is weak. Also, the greater the variation in values and preferences of patients (and/or informed proxies), or the greater their uncertainty, the more likely a weak grading is warranted. Similarly, the more uncertain it is that an intervention represents a wise use of resources (eg, a marginal net benefit of a very resource-intensive intervention), the lower the likelihood of a strong grading. ◆

mendations that are biased, poorly substantiated or insensitive to patient and clinician needs and which, if followed, may have far-reaching deleterious effects on clinical practice.

Competing interests

None identified.

Author details

Ian A Scott, FRACP, MHA, MEd, Director¹
Gordon H Guyatt, MD, MSc, FRCP(C), Professor²

¹ Department of Internal Medicine and Clinical Epidemiology, Princess Alexandra Hospital, Brisbane, QLD.

² Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada.

Correspondence: ian_scott@health.qld.gov.au

References

- Grol R, Dalhuijsen J, Thomas S, et al. Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. *BMJ* 1998; 317: 858-861.

2 Lugtenberg M, Burgers JS, Westert GP. Effects of evidence-based clinical practice guidelines on quality of care: a systematic review. *Qual Saf Health Care* 2009; 18: 385-392.

3 Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* 1999; 282: 1458-1465.

4 Choudry NK, Stelfox HT, Detsky AS. Relationships between authors of clinical practice guidelines and the pharmaceutical industry. *JAMA* 2002; 287: 612-617.

5 Forge B. The "Acute coronary syndromes: consensus recommendations for translating knowledge into action" position statement is based on a false premise. *Med J Aust* 2010; 192: 696-699.

6 Thompson PL. The invasive approach to acute coronary syndrome: true promise or false premise? [editorial] *Med J Aust* 2010; 192: 694-695.

7 Acute Coronary Syndrome Guidelines Working Group. Guidelines for the management of acute coronary syndromes 2006. *Med J Aust* 2006; 184: S1-S32.

8 Brieger D, Kelly A-M, Aroney C, et al; on behalf of the National Heart Foundation ACS Implementation and Advocacy Working Group. Acute coronary syndromes: consensus recommendations for translating knowledge into action. *Med J Aust* 2009; 191: 334-338.

9 Burgers JS, Cluzeau FA, Hanna SE, et al. Characteristics of high-quality guidelines: evaluation of 86 clinical guidelines developed in ten European countries and Canada. *Int J Technol Assess Health Care* 2003; 19: 148-157.

10 Van der Weyden MB. Clinical practice guidelines: time to move the debate from the how to the who [editorial]. *Med J Aust* 2002; 176: 304-305.

11 Detsky AS. Sources of bias for authors of clinical practice guidelines. *CMAJ* 2006; 175: 1033-1035.

12 Shrier I, Boivin JF, Platt RW, et al. The interpretation of systematic reviews with meta-analysis: an objective or subjective process? *BMC Med Inform Decis Mak* 2008; 8: 19.

13 Tricoci P, Allen JM, Kramer JM, et al. Scientific evidence underlying the ACC/AHA clinical practice guidelines. *JAMA* 2009; 301: 831-841.

14 MacDonald N, Downie J. Editorial policy: industry funding and editorial independence. *CMAJ* 2006; 174: 1817.

15 Buchan HA, Currie KC, Lourey EJ, Duggan GR. Australian clinical practice guidelines — a national study. *Med J Aust* 2010; 192: 490-494.

16 Silverman GK, Loewenstein GF, Anderson BL, et al. Failure to discount for conflict of interest when evaluating medical literature: randomised trial of physicians. *J Med Ethics* 2010; 36: 265-270.

17 Institute of Medicine. Conflicts of interest and development of clinical practice guidelines. In: Conflict of interest in medical research, education and practice. Washington, DC: IOM, National Academies Press, 2009: 189-215.

18 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: 738-741.

19 Greenhalgh T. Papers that tell you what to do (guidelines). In: How to read a paper. The basics of evidence-based medicine. 4th ed. London: Wiley-Blackwell, BMJ Books 2010: 132-148.

20 Boyd EA, Bero LA. Improving the use of research evidence in guideline development. 4. Managing conflicts of interests. *Health Res Policy Syst* 2006; 4: 16.

21 Eikelboom JW, Guyatt G, Hirsh JW. Guidelines for anticoagulant use in acute coronary syndromes. *Lancet* 2008; 371: 1559-1561.

22 Atkins D, Eccles M, Flottorp S, et al; for the GRADE Working Group. Systems for grading the quality of evidence and the strength of recommendations 1. Critical appraisal of existing approaches. *BMC Health Serv Res* 2004; 4: 38.

23 Guyatt GH, Oxman AD, Vist GE, et al; for the GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; 336: 924-926.

24 McAlister FA, van Diepen S, Padwal R, et al. How evidence based are the recommendations in evidence-based guidelines? *PLoS Med* 2007; 4: e250-e258.

25 Schunemann HJ, Jaeschke R, Cook DJ, et al; on behalf of the ATS Documents Development and Implementation Committee. An official ATS statement: Grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. *Am J Resp Crit Care Med* 2006; 174: 605-614.

26 Djulbegovic B, Trikalinos TA, Roback J, et al. Impact of quality of evidence on the strength of recommendations: an empirical study. *BMC Health Serv Res* 2009; 9: 120-125.

27 National Health and Medical Research Council. NHMRC additional levels of evidence and grades for recommendations for developers of guidelines. Stage 2 consultation. Early 2008 – end June 2009. http://www.nhmrc.gov.au/guidelines/consult/consultations/add_levels_grades_dev_guidelines2.htm (accessed Jun 2010).

28 Vitry AI, Zhang Y. Quality of Australian clinical guidelines and relevance to the care of older people with multiple comorbid conditions. *Med J Aust* 2008; 189: 360-365.

29 Farquhar CM, Kofa EW, Slutsky JR. Clinicians' attitudes to clinical practice guidelines: a systematic review. *Med J Aust* 2002; 177: 502-506.

30 Haycox A, Bagust A, Walley T. Clinical guidelines: the hidden costs. *BMJ* 1999; 318: 391-393.

Provenance: Not commissioned; externally peer reviewed.

(Received 28 Jul 2010, accepted 9 Mar 2011)

