A guide to the Canadian Medical Association Handbook on Clinical Practice Guidelines

Valerie A. Palda MD MSc, Dave Davis MD, Joanne Goldman MSc

At the fourth meeting of the Guideline International Network, held in Toronto in August 2007, world experts from 31 countries met to discuss the challenges and innovations afforded by clinical practice guidelines. Although each country faces unique local challenges to the implementation of effective health care, members of the guideline community have repeatedly and generously shared their solutions to these challenges, many of which are applicable in other countries. Nevertheless, it can be difficult for the local implementer with limited resources to harness this knowledge.

The Canadian Medical Association has long been a champion of enhancing the quality of clinical practice guidelines. The objective of the recently published Canadian Medical Association Handbook on Clinical Practice Guidelines (Box 1, Table 1) is to gather up-to-date, evidence-based, experience-driven guidance on how to use guidelines most effectively to improve patient care. This new handbook combines and updates the 1994 and 1997 documents “Guidelines for Canadian Clinical Practice Guidelines” and “Implementing Clinical Practice Guidelines: a Handbook for Practitioners.” The new handbook places the role of guidelines in health care into perspective, outlining where they are most useful. It helps the reader to decide if an existing guideline can be adapted or if a new guideline should be created, providing resources for both scenarios. The handbook also reviews evidence to guide those responsible for implementing recommendations through the bewildering array of available implementation strategies. In recognition of increasing demands for accountability and increasing emphasis on quality of care, the final chapter reviews the process of evaluating the effectiveness of guidelines.

The handbook reviews in depth the methodologic steps in the guideline process, including the biases inherent in guideline development, the struggle to write accurate recommendations and an approach to implementation.

Readers of the new handbook will be able to use it in 3 ways: as a review for the key parts of the process, with reference to the published evidence; as a source of practical approaches to complete the part of the cycle in which the user is involved; and as an illustrative inventory of resources and links. Guideline developers and implementers may be health care practitioners, administrators, health organizations or policy-makers. Experienced guideline developers and implementers may find within the handbook innovations from the international community that apply to their work. For guideline users, knowledge of what makes a guideline “good” allows selection of the best guidelines, saving time and potentially improving patient outcomes. It is our hope that the handbook may even inspire a few guideline users to become local champions of best practices.

Dealing with bias in guideline development

One of the significant changes in the field of guideline development has been widespread acceptance of a standardized methodology for the production of clinical practice guidelines. Just as the writing of randomized controlled trials has been informed by the CONSORT (Consolidated Standards of Reporting Trials) statement and the writing of systematic reviews improved by the QUORUM (Quality of Reporting of Meta-analyses) statement, so the AGREE (Appraisal of Guidelines Research and Evaluation) Collaboration has promoted validated criteria for writing clinical practice guidelines (www.agreecollaboration.org). The elements of the AGREE instrument are summarized in Table 2.

Each domain in the AGREE instrument reflects either a potential source of bias or issues related to clarity and understanding of the guideline. In trials, investigators aim to minimize bias to get closer to the “truth”: the same experiment conducted repeatedly should yield the same results. In clinical practice guidelines, minimizing bias should likewise lead various groups of developers who consider the same evidence to come up with similar recommendations. Sources of bias in the evidence-review process include the type of literature

From the Departments of Medicine (Palda), Health Policy Management and Evaluation (Palda, Davis), and Continuing Education and Professional Development, Faculty of Medicine (Goldman), University of Toronto, Toronto, Ont.; and the Association of American Medical Colleges (Davis), Washington, DC

Box 1: Canadian Medical Association Handbook on Clinical Practice Guidelines in a nutshell

The new Canadian Medical Association Handbook on Clinical Practice Guidelines combines 2 earlier documents, which have been updated and revised for use by individuals and groups interested in the development, adaptation, implementation and evaluation of clinical practice guidelines. The handbook reviews the key parts of each step of the guideline-development process, with appropriate reference to the published evidence. It suggests practical approaches to completing each component of the guideline cycle, using illustrative examples and referencing print and electronic resources to aid the reader.
search used, the method used to evaluate the quality of the literature, and editorial independence.

As for any systematic review, it is vital that a systematic search strategy be used in the development of clinical practice guidelines. Relying on experts’ recollection of the literature, as was common with consensus guidelines, is no longer sufficient. For example, Gilbert and associates\(^9\) compared historical recommendations with a systematic review of observational studies of the effect of infant sleeping position on sudden infant death syndrome. They found that by 1970, the literature demonstrated a statistically significantly increased risk of sudden infant death for sleeping on the front relative to sleeping on the back (pooled odds ratio 2.93, 95% confidence interval 1.15–7.47); however, guidelines did not consistently recommend the back-sleeping position until 1992. These authors concluded that use of systematic review techniques could have led to earlier recognition of the risks of sleeping on the front and might have prevented more than 10,000 infant deaths in the United Kingdom and at least 50,000 in Europe, the United States and Australasia.\(^9\) The recommendations in any clinical practice guideline should consider the results of the totality of the literature, giving greater weight to better-designed studies. One recent study\(^9\) found that applying 2 different quality-evaluation methods (Cochrane or best-evidence synthesis) led to different recommendations. Listing a level of evidence for each recommendation forces the guideline developer to identify the strength of the evidence supporting the statement.

Editorial independence, the sixth and last domain in the AGREE instrument (Table 2), asks the reader to evaluate conflict of interest. Financial conflict of interest has been the type of bias most widely discussed. One study of clinical practice guidelines published between 1991 and 1999 found that 87% of guideline authors had interactions with the pharmaceutical industry, 58% had received financial support for research, and 38% had been employees of or consultants for a pharmaceutical company.\(^9\) To score well in this domain of the AGREE instrument, the guideline must state not only that all group members have declared whether they have any conflict of interest (using standard forms), but also that

### Table 1: Summary of contents of the Canadian Medical Association Handbook on Clinical Practice Guidelines\(^3\)

| Chapter number and title | Description of content | Highlights |
|--------------------------|------------------------|------------|
| 1. Introduction to clinical practice guidelines | • Introduces the field of quality of health care and discusses the significance of the process by which knowledge, specifically evidence-based research, is incorporated into routine practice<br>• Discusses the key role of guidelines within this broader framework of quality of health care | • Knowledge-to-action framework<br>• Benefits and challenges of clinical practice guidelines<br>• Key websites and articles providing perspective in this area |
| 2. Adaptation and development of CPGs | • Clarifies the role of guidelines by answering the question, “Do you need a CPG?”<br>• Presents common salient elements in the adaptation and development of clinical practice guidelines<br>• Outlines a process for adapting guidelines for local or regional use<br>• Highlights key elements of guideline development<br>• Offers resources to assist with these initiatives | • Chart of major steps in creating a clinical practice guideline<br>• Practical considerations for focusing the topic<br>• Guidance on assembling a working group, including consideration of bias<br>• Description of ADAPTE process for adapting guidelines for local use<br>• Description of AGREE instrument for assessing quality<br>• Description of GRADE, SIGN and SORT taxonomies for making recommendations<br>• Websites for finding, adapting and appraising clinical practice guidelines |
| 3. Implementation of CPGs to change practice and outcomes | • Outlines the nature and challenges of knowledge translation, the process by which best-evidence guidelines are adopted by clinicians, patients and policy-makers<br>• Describes the stages in planning a guideline-implementation strategy: analyzing the context, reviewing the relevant literature and designing a strategy based on this information | • Guidance on assessing contextual barriers to and facilitators of change<br>• Review of known strategies for implementing guidelines<br>• Theories of behavioural and organizational change<br>• Steps to developing a local, targeted implementation plan<br>• Websites of experienced implementation groups |
| 4. Evaluation of CPGs | • Provides a brief overview of methods for evaluating clinical practice guidelines | • Definition of quality indicators<br>• Review of data collection techniques and outcome measures<br>• Websites and published articles to inform measurement strategies |

Note: ADAPTE = ADAPTE Collaboration, AGREE = Appraisal of Guidelines Research and Evaluation (collaboration), GRADE = Grading of Recommendations Assessment, Development and Evaluation (working group), SIGN = Scottish Intercollegiate Guidelines Network, SORT = Strength of Recommendation Taxonomy.
the views or interests of the funding body have not influenced the final recommendations. The American College of Chest Physicians guards against this type of bias by disallowing participation of any guideline panel member who does not complete a conflict of interest disclosure form and by careful review of each disclosure form. Conflicts declared by panel members are reviewed using a graded consideration based on the potential level of conflict, whether the conflict can be managed within established parameters and whether the panel member has expertise that would allow participation in a related area that does not involve the conflict. The disclosures that prove most difficult to evaluate receive full committee review.11

Many journals now refuse to publish clinical practice guidelines unless statements of conflict of interest are available, and many guideline developers ask participants in the guideline development process to fill out standard conflict of interest forms. A description of how potential conflicts have been addressed is often lacking, however. Given that 7% of the authors surveyed by Choudhry and associates10 stated that their own relationships influenced recommendations, and 19% thought that their coauthors’ recommendations were influenced, this form of bias remains a significant challenge to the reliability of clinical practice guidelines. Other potential sources of bias also exist, such as long-term service to government committees or private insurers, participants’ previously established “stake” in an issue, the way in which developers make their living and personal experiences.12

Developing recommendations: grading systems

Many guideline developers provide a legend to explain how they came to each of the recommendations. The strength of recommendations is often categorized according to a specific “grading system,” which usually considers only levels of evidence but sometimes also addresses other factors that might influence the strength of the recommendation, such as the magnitude of the therapeutic risk reduction and the magnitude of potential harms and benefits for possible outcomes. The GRADE (Grading of Recommendations, Assessment, Development and Evaluation) system is an international effort to standardize the approach to making recommendations (www.gradeworkinggroup.org). The GRADE approach assigns evidence “quality” at 1 of 4 levels — very low, low, moderate or high — on the basis of specific criteria. Each recommendation is then based on a judgment of net benefits, including whether the net benefits are positive, negative or uncertain. This and other grading schemes13-14 that are less explicit but still provide transparency are compared in Table 3. In 2003, the GRADE Working Group itself acknowledged that there is no published evidence on how best to communicate grades of evidence and recommendations.15

Guideline developers often modify grading systems to reflect their specific needs. The US Preventive Services Task Force, an active participant in the GRADE Working Group, has chosen to maintain its own recommendation grading system16 to reflect this organization’s more narrow focus on prevention, in contrast to the broader clinical scope of the GRADE Working Group. Until sufficient evidence accumulates to demonstrate the superiority of one system over another, concentrating on making recommendations clear and reflective of the evidence is a reasonable approach for most guideline developers.

Table 2: The AGREE instrument: domains and descriptions of key items*

| Domain                        | Key Items                                                                 |
|-------------------------------|---------------------------------------------------------------------------|
| Scope and purpose             | - Objective(s) of the guideline is (are) specifically described           |
|                               | - Clinical question(s) is (are) specifically described                    |
|                               | - Target patients are specifically described                             |
| Stakeholder involvement       | - Guideline-development group includes all relevant professional groups   |
|                               | - Patients’ views and preferences have been sought                        |
|                               | - Target users of the guideline are clearly defined                       |
|                               | - Guideline has been piloted by target users                             |
| Rigour of development         | - Systematic methods were used to search for evidence                     |
|                               | - Criteria for selecting the evidence are clearly described               |
|                               | - Methods used for formulating the recommendations are clearly described |
|                               | - Health benefits, side effects and risks were considered in formulating the recommendations |
|                               | - There is an explicit link between the recommendations and the supporting evidence |
|                               | - Guideline has been externally reviewed by experts before its publication |
|                               | - A procedure for updating the guideline is provided                      |
| Clarity and presentation      | - Recommendations are specific and unambiguous                            |
|                               | - Different options for management of the condition are clearly presented |
|                               | - Key recommendations are easily identifiable                            |
|                               | - Guideline is supported with application tools                          |
| Applicability                 | - Potential organizational barriers in applying the recommendations are discussed |
|                               | - Potential cost implications of applying the recommendations have been considered |
|                               | - Guideline presents key review criteria for monitoring and audit purposes |
| Editorial independence        | - Guideline is editorially independent from the funding body               |
|                               | - Conflicts of interest of the guideline-development group have been recorded |

*Details about each question and more specific explanations are available in the AGREE Instrument (www.agreecollaboration.org/pdf/agreeinstrumentfinal.pdf).
Implementation of guidelines

In the end, even well-designed guidelines in the same area will occasionally differ in their recommendations. Developers should therefore consider, during the development phase of any guideline, its ease of implementation (implementability). The recent GLIA (GuideLine Implementability Appraisal) instrument takes the developer through a series of validated questions that ask about factors known to predict the relative ease of implementation of guideline recommendations. The currently recommended approach to implementation is summarized in Box 2. Implementation strategies may be most effective when they are targeted to locally identified facilitators and barriers to implementation. Barriers may be effectively identified through a process as simple as structured reflection by the implementation group. Many implementation strategies have shown modest benefit, and multiple strategies often work better than single ones. Although the sheer number of possible implementation strategies precludes their description here, the interested reader is directed to the handbook, which reviews the major implementation strategies that have been assessed in the literature.

| Strategy       | Strengths                                                                                                                                                                                                 | Weaknesses                                                                                                                                                                                                 |
|----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| GRADE system   | • Working group is an international collaboration interested in developing a common grading system to address limitations and draw on strengths of existing systems  
• System sequentially assesses quality of evidence, balance between risks and benefits, and judgment about the strength of recommendations  
• Final recommendation categories are simple: **Strong**: recommendations reflect evidence that the benefits do, or do not, outweigh risks and burdens  
**Weak**: recommendations reflect evidence that the benefits, risks and burdens are finely balanced, or there is appreciable uncertainty about the balance; furthermore, the recommendation is classified as “weak” if, across the range of patient values, fully informed patients are liable to make different choices | • Application is complicated  
• Often difficult for recommendation developers to use formulaic approaches to global judgments about evidence |
| SIGN method    | • Represents a collaboration to improve the quality of health care for patients in Scotland by reducing variation in practice and outcomes, through the development and dissemination of national clinical guidelines  
• Levels of evidence (1++, 1+, 1–, 2++, 2+, 2–, 3 or 4) depend on type and quality of study design; grade of recommendation (A, B, C or D) reflects assigned level of evidence  
• “Considered judgement” forms are used to help guideline development if decisions must be made according to experience as well as knowledge of evidence and underlying methods; forms address quantity, quality and consistency of evidence, generalizability of study findings, directness and clinical impact | • System lacks transparency; no rationale provided to clarify which factors are weighted more heavily for any particular recommendation  
• Use of numbers and letters may not be intuitive |
| SORT taxonomy  | • Developed by US family medicine and primary care journals and the Family Practice Inquiries Network to address the need for a single consistently applied taxonomy of evidence  
• Emphasizes patient-oriented outcomes (i.e., “outcomes that matter to patients and help them live longer or better lives, including reduced morbidity, mortality or symptoms, improved quality of life or lower cost”14)  
• Rates quality of individual studies as follows: 1 = good-quality patient-oriented evidence, 2 = limited-quality patient-oriented evidence, 3 = other evidence  
• Grades strength of recommendations by letters: A recommendations are based on consistent, good-quality patient-oriented evidence; B recommendations are based on inconsistent or limited-quality patient-oriented evidence; C recommendations are based on consensus, usual practice, opinion, disease-oriented evidence or case series for studies of diagnosis, treatment, prevention or screening | • Limited guidance for developers on how to classify studies within the numeric categories (1, 2 or 3)  
• Use of numbers and letters may not be intuitive |

Note: GRADE = Grading of Recommendations, Assessment, Development and Evaluation system (www.gradeworkinggroup.org), SIGN = Scottish Intercollegiate Guidelines Network method (www.sign.ac.uk), SORT = Strength of Recommendation Taxonomy.14
Other guideline manuals

Many guideline development organizations have manuals outlining their methods for interested readers. For users who do not have the resources to develop their own clinical practice guidelines, help is available for finding, evaluating and adapting existing guidelines for local use (Box 3). Other groups have provided guides for implementation and chart-type tools to guide practice (Box 3). For those wanting to test how well their guideline has worked, evaluation strategies are harder to find: usually the process of designing and performing a guideline evaluation involves a review of primary studies of guideline evaluation to determine which evaluation strategy would be most suitable.

Other organizations that offer extensive English-language collections of resources and references include Australia’s federal guideline agency, the New Zealand Guidelines Group, the Scottish Intercollegiate Guidelines Network and the UK National Institute for Health and Clinical Excellence (Box 3). The Guidelines International Network provides a wealth of international resources (Box 3). To our knowledge, until now, no Canadian organization has brought all these resources together in a single document, but these are all listed in the Canadian Medical Association Handbook on Clinical Practice Guidelines.

What the future holds for guidelines and the handbook

Current research related to clinical practice guidelines includes studies of the role of patient involvement in guideline development, the validation of tools to enhance the implementability and evaluation of guidelines, and examination of the balance between studying and implementing guidelines. As the field of guideline methodology matures, developers will also struggle with providing guidance in the context of multidisciplinary approaches to care and with addressing the needs of patients who have multiple chronic conditions.

The new handbook, like clinical practice guidelines themselves, should be considered a living document, responsive to changes in the literature and feedback from guideline implementers. It will therefore need regular updating to incorporate advances in knowledge about clinical practice guidelines. We welcome any comments that readers of this article may have.

Competing interests: None declared.

Contributors: All of the authors made substantial contributions to the content and framework of the article, revised it critically for important intellectual content and provided final approval of the version to be published.
REFERENCES

1. Guidelines International Network. Collaboration in clinical practice guidelines — local, national and international cooperation in synthesizing and applying best evidence. Proceedings of the 4th G-I-N (Guidelines International Network) Conference; 2007 Aug 22–25; Toronto. Available: www.g-i-n.net/index.cfm?fuseaction=news&fusesubaction=docs&documentID=57 (accessed 2007 Oct 1).

2. Rich P. CMA helping doctors navigate guidelines maze. Ottawa: Canadian Medical Association; 2004. Available: www.cma.ca/index.cfm?ci_id=4409&la_id=1 (accessed 2007 Oct 1).

3. Davis D, Goldman J, Palda VA. Canadian Medical Association handbook on clinical practice guidelines. Ottawa: Canadian Medical Association; 2007.

4. Guidelines for Canadian clinical practice guidelines. In: CMA Infobase: clinical practice guidelines. Ottawa: Canadian Medical Association. Available: mdm.ca/cpgsnew/cpgs/cpgc-e.htm (accessed 2007 Oct 1).

5. Implementing clinical practice guidelines: a handbook for practitioners. In: CMA Infobase: clinical practice guidelines. Ottawa: Canadian Medical Association; 1997. Available: mdm.ca/cpgsnew/cpgs/handbook/index.htm (accessed 2007 Oct 1).

6. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA 1996;276:637-9.

7. Moher D, Cook DJ, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUORUM statement. Lancet 1999;354:1896-900.

8. Gilbert R, Salanti G, Harden M, et al. Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002. Int J Epidemiol 2005;34:877-87.

9. van der Velde G, van Tulder M, Cote P, et al. The sensitivity of review results to analyses of randomised controlled trials: the QUORUM statement. Lancer 1999;354:1896-900.

10. HS. Dr. William Marsden Award in Medical Ethics

11. Herreras C, Lewis SZ, Harris DA, et al. American College of Chest Physicians evidence-based guideline methodology. Presented at the 41st G-I-N (Guidelines International Network) Conference; Toronto; 2007 Aug 22–25.

12. Detkay AS. Sources of bias for authors of clinical practice guidelines [editorial]. CMAJ 2006;175:1033.

13. SIGN 50: a guideline developer’s handbook. Edinburgh: SIGN — Scottish Intercollegiate Guidelines Network; 2004. Available: www.sign.ac.uk/guidelines/fulltext/50/index.html (accessed 2007 Oct 1).

14. Ebell MH, Siwek J, Weiss BD, et al. Simplifying the language of evidence to improve patient care. Strength of Recommendation Taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature. J Fam Pract 2004;53:111-20.

15. Schünemann HJ, Best D, Vist G, et al; GRADE Working Group. Letters, symbols and words: how to communicate grades of evidence and recommendations. CMAJ 2003;169:677-80.

16. Barton MB, Miller T, Wolf T, et al.; U.S. Preventive Services Task Force. How to read the new recommendation statement: methods update from the U.S. Preventive Services Task Force. Ann Intern Med 2007;147:123-7.

17. Shiffman RN, Dixon J, Brandt C, et al. The GuideLine Implementability Appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. BMC Med Inform Decis Mak 2005;5:23.

18. Fretheim A, Oxman AD, Flottorp S. Improving prescribing of antihypertensive and cholesterol-lowering drugs: a method for identifying and addressing barriers to change. BMC Health Serv Res 2004;4:43.

19. Grimshaw JM, Thomas RE, MacLennan G, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004;8(3):iii-iv, i-72.

20. The guidelines manual. London (UK): National Institute for Health and Clinical Excellence; 2007. Available: www.nice.org.uk/page.aspx?o=422956 (accessed 2007 Oct 1).

21. Harris RP, Helfand M, Woolf SH, et al; Methods Work Group, Third US Preventive Services Task Force. Current methods of the US Preventive Services Task Force: a review of the process. Ann J Prev Med 2001;20(3 Suppl):21-35.

22. Shojania KG, McDonald KM, Wachter RM, et al. Series overview and methodology. Vol 1 of Closing the quality gap: a critical analysis of quality improvement strategies [Technical review 9]. Rockville (MD): Agency for Healthcare Research and Quality; 2004. Contract no 290-02-0017 to the Stanford University — UCSF Evidence-based Practices Center. AHRQ Publication no 04-0051-1.

Correspondence to: Dr. Valerie A. Palda, Department of Medicine, University of Toronto, St. Michael’s Hospital, Rm. CC4-154, 30 Bond St., Toronto ON M5B 1W8; fax 416 864-5541; va.palda@utoronto.ca