Early release, published at www.cmaj.ca on March 4, 2013. Subject to revision.

How can Canadian guideline recommendations be tested?

Ananda Chatterjee MD, Onil Bhattacharyya MD, Navindra Persaud MD MSc

The developers of clinical practice guidelines invest substantial resources to produce and disseminate recommendations in an attempt to improve patient care. This process involves reviewing large bodies of sometimes conflicting evidence, converting the evidence into actionable and feasible recommendations, and deciding how best to communicate those recommendations to clinicians. But does this effort lead to better outcomes for patients?

We suggest that guidelines should be tested before widespread dissemination. Most guidelines are not tested, and some have caused harm in the past.[1,2] In this article, we outline some practical methods for testing clinical practice guidelines, including public consultation, clinical vignettes and performance measures. If guidelines are shown to improve outcomes, physicians may use them with greater confidence.

Potential for harm

Few guidelines have been tested to determine their effect on patient care or clinical outcomes despite calls for testing. For example, *The Journal of the American Medical Association*'s Users' Guides to the Medical Literature series includes the question: "Has the guideline been subjected to peer-review and testing?"[3] More than 20 years ago, the Institute of Medicine decried the lack of attention to the effect of guidelines on patient care:

> …Relatively few steps were underway or planned to evaluate the impact of guidelines on the cost, quality and outcomes of care and on the patient and practitioner satisfaction. This neglect is unfortunate because the effectiveness of guidelines cannot be taken for granted.[4]

Current tools for assessing guidelines (Appraisal of Guidelines for Research and Evaluation and GuideLine Implementability Appraisal)[5,6] focus on methodologic issues related to the gathering of evidence and development of guidelines, but they do not consider whether the guidelines have been tested.

Most Canadian clinical practice guidelines have not been tested. In contrast, several clinical prediction rules developed in Canada (e.g., the Ottawa Ankle Rules) have been tested in a variety of clinical settings and in multiple countries.[7]

Elements of the untested Canadian clinical practice guidelines may be harmful in certain clinical contexts, just as guidelines in other countries have caused harm. Linden and Schotte[1] found that general practitioners exposed to depression guidelines prescribed doxepin at higher doses to patients with mild depression; these patients had lower rates of symptom improvement than patients cared for by physicians in the control group. The authors concluded that "guidelines should be empirically tested before being called 'evidence based'." A Dutch guideline that combined recommendations related to asthma and chronic obstructive pulmonary disease led to an appropriate decrease in inhaled corticosteroid prescriptions for patients with chronic obstructive pulmonary disease, but they also led to an inappropriate decrease in inhaled corticosteroid prescriptions for patients with asthma.[2]

Difficulty assessing effect on care

It would be impractical to conduct randomized controlled trials of guideline use on clinical outcomes for every version of every guideline. Most of the outcomes of interest (e.g., macrovascular complications of diabetes) take years or decades to accrue in significant numbers. Even for outcomes that appear sooner (e.g., asthma exacerbations), field trials would still be too expensive for guideline developers who struggle to find resources to support the development and dissemination of

Key points

- Most clinical practice guidelines are not tested either before or after widespread dissemination, so their effect on clinical outcomes is unknown.
- It is difficult to determine the effect of clinical practice guidelines on clinical outcomes because trials are impractical.
- Guidelines can be tested using public consultations and clinical vignettes before widespread dissemination and by use of performance measures after dissemination.
- Clinicians should preferentially use clinical practice guidelines that have been tested.

Competing interests: None declared.

Disclaimer: Navindra Persaud is an associate editor for *CMAJ* and was not involved in the editorial decision-making process for this article.

This article has been peer reviewed.

Correspondence to: Navindra Persaud, nav.persaud@utoronto.ca

***CMAJ* 2013. DOI:10.1503/cmaj.121830**

guidelines. Control group contamination would also be a problem when guidelines are widely disseminated. Even if all of these obstacles were overcome, the results of randomized controlled trials would not necessarily show which parts of guidelines were helpful and which were not. More practical methods for assessing the effect of guideline recommendations on care are needed.

Potential solutions

We suggest public consultation, clinical vignette studies and performance measures as 3 pragmatic methods for testing clinical practice guidelines. These methods can be used before publication of guidelines or after dissemination to help determine if the recommendations are acceptable, understandable and helpful to clinicians and patients. The strengths and limitations of each method (Table 1) should be considered when deciding which is most appropriate.

Public consultation allows clinicians to comment on draft guidelines before widespread dissemination. Clinicians who think that the recommendations are confusing, impractical or unlikely to be helpful can have their feedback incorporated before the guidelines are finalized. The National Institute of Clinical Studies in Australia displays draft guidelines on its website and has a systematic submission process for comments and suggestions by the scientific community; the guideline authors then review and incorporate the suggestions before the final document is complete.[8]

Clinical vignette studies can be used to test the understandability of guidelines before publication. Clinical practice guideline recommendations are the result of a complex process involving a number of people, and users might misinterpret the guidelines. Clinical vignette performance is a validated and accurate measure of the process of care.[9] Different versions of a recommendation (e.g., text v. algorithm) can be put head-to-head to determine which one leads to better clinical decisions. This allows rapid feedback to guideline developers about contentious recommendations. Shekelle and colleagues[10] used clinical vignettes in randomized controlled trials to compare specific and vague guidelines for ordering electrodiagnostic tests for patients with back pain; they found that vague guidelines increased the frequency of inappropriate test ordering, whereas specific guidelines decreased inappropriate test ordering.

Performance measures should be included in guidelines so that the developers can solicit feedback as to how the measures change after the guidelines are implemented. This is a practical way to determine the effect of guideline recommendations on patient outcomes. For example, Hypertension Canada's Outcomes Research Task Force observed that an 84.4% increase in prescriptions for antihypertensive medications between

Table 1: Strengths and limitations of 3 pragmatic methods for testing clinical practice guidelines

Method	Strengths	Limitations	Guideline developers currently using this method
Public consultations	• Incorporates feedback from frontline clinicians • Increases acceptance from the medical community	• Extends the length of time for guideline development	• National Institute of Clinical Studies (Australia) • United States Preventive Services Task Force • Scottish Intercollegiate Guidelines Network • National Institute for Health and Clinical Excellence (UK)
Clinical vignettes	• Requires minimal resources • Provides rapid feedback • Allows different formats of recommendations to be tested	• Requires a group of physicians as study participants • Must be carefully constructed and linked to explicit outcomes or evidence-based guidelines • May not reflect actual behaviours of clinicians	• None
Performance measures	• Quality of medical care can be gauged	• Feedback is available only after guidelines have been implemented • Performance measures do not always reflect the quality of care • Unmeasured aspects of care could worsen	• Canadian Task Force on Preventive Health Care • Canadian Stroke Network • National Institute for Health and Clinical Excellence (UK)

1999 and 2006 was associated with decreases in mortality from stroke, heart failure and myocardial infarction.[11] Carefully selected performance measures can be used to gauge the effect on quality of care, as practised by the National Institute for Health and Clinical Excellence in the United Kingdom.[12] Feedback about the effect of a previous version of a guideline on performance measures should be incorporated in later versions.

Conclusion

Clinicians should preferentially use clinical practice guidelines that have been tested. Existing methods (e.g., public consultation, performance measures) and new methods (e.g., clinical vignettes) for testing guidelines should be regularly applied to Canadian guidelines to ensure that clinical practice guidelines improve patient care.

References

1. Linden M, Schotte K. A randomized controlled clinical trial comparing "guideline exposed" and "guideline naive" physicians in respect to dosage selection and treatment outcome with doxepin in depressive disorders. *Pharmacopsychiatry* 2007;40: 77-81.
2. Martens JD, Winkens RA, van der Weijden T, et al. Does a joint development and dissemination of multidisciplinary guidelines improve prescribing behaviour: a pre/post study with concurrent control group and a randomised trial. *BMC Health Serv Res* 2006; 6:145.
3. Hayward RS, Wilson MC, Tunis SR, et al. Users' guides to the medical literature. VIII. How to use clinical practice guidelines. A. Are the recommendations valid? The Evidence-Based Medicine Working Group. *JAMA* 1995;274:570-4.
4. Field MJ, Lohr MJ, editors. *Clinical practice guidelines: directions for a new program*. Washington (DC): National Academy Press; 1990.
5. Brouwers M, Kho ME, Browman GP, et al. AGREE II: Advancing guideline development, reporting and evaluation in healthcare. *CMAJ* 2010;182:E839-42.
6. Shiffman RN, Dixon J, Brandt C, et al. The GuideLine Implementability Appraisal: development of an instrument to identify obstacles to guideline implementation. *BMC Med Inform Decis Mak* 2005;5:23.
7. Stiell IG, Bennett C. Implementation of clinical decision rules in the emergency department. *Acad Emerg Med* 2007;14:955-9.
8. National Health and Medical Research Council. Canberra (AU): Commonwealth of Australia; National Institute of Clinical Studies. Available: www.nhmrc.gov.au/guidelines/public-consultations (accessed 2012 Oct. 15).
9. Peabody JW, Luck J, Glassman P, et al. Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. *Ann Intern Med* 2004;141:771-80.
10. Shekelle PG, Kravitz RL, Beart J, et al. Are nonspecific guidelines potentially harmful? A randomized comparison of the effect of nonspecific versus specific guidelines on decision making. *Health Serv Res* 2000;34:1429-48.
11. Bolli P, Campbell NRC. Do recommendations for the management of hypertension improve cardiovascular outcome? The Canadian experience. *Int J Hypertens* 2011;2001:410754.
12. Campbell SM, Kontopantelis E, Hannon K, et al. Framework and indicator testing protocol for developing and piloting quality indicators for the UK quality and outcomes framework. *BMC Fam Pract* 2011;12:85.

Affiliations: From the Keenan Research Centre in the Li Ka Shing Knowledge Institute (Chatterjee, Bhattacharyya, Persaud), St. Michael's Hospital, Toronto, Ont.; the Department of Family and Community Medicine (Bhattacharyya, Persaud), St. Michael's Hospital; and the Department of Family and Community Medicine (Bhattacharyya, Persaud) University of Toronto, Ont.

Contributors: Navindra Persaud conceived the commentary. Ananda Chatterjee drafted the manuscript, which was revised by all of the authors. All authors have approved the final version of the document.