

Clinical practice guidelines on trial



Robert S.A. Hayward, MD, MPH

Résumé

LE MOMENT EST-IL VENU DE LAISSER TOMBER les guides de pratique clinique (GPC)? Comme le signalent le Dr Graham Worrall et ses collègues (page 1705), les données probantes qui indiquent que les GPC ont eu une incidence favorable sur les résultats des patients dans le contexte des soins primaires sont rares. Ils signalent toutefois que sur 13 études d'interventions fondées sur des GPC qu'ils ont examinées, toutes sauf 2 présentent de graves lacunes méthodologiques. Dans cet éditorial, l'auteur décrit à quoi pourrait ressembler une étude idéale sur un guide, indique si une telle étude est possible et précise les principales menaces à la validité. Tant qu'on n'aura pas fait d'études qui satisfont à ces critères plus rigoureux, nous ne pourrions conclure que le mouvement des guides est lui-même mal guidé.

It is distressing to health care researchers that so much of the evidence they produce remains unheeded in day-to-day practice. At the same time, the potential consumers of this evidence are distressed by the backlog of clinical problems for which evidence is confusing, conflicting, inapplicable or unavailable. They observe that evidence must be interpreted and applied in contexts complicated by comorbidity, costs and other concerns that do not figure in clinical trials.

Clinical practice guidelines (CPGs) attempt to bridge the gap between producers and consumers of health care research. Good guidelines start with a specific clinical question, articulate relevant issues, seek and synthesize sound evidence, assign values to outcomes, generate recommendations and try to influence what clinicians do in the hope that reduced practice variation, lower costs and improved health outcomes will result.¹⁻³

Despite the fact that CPGs are prominent and prolific, early evidence of their impact is patchy. Some studies of guideline interventions suggest that they improve practice,⁴⁻⁶ while others show little change in practice and outcomes.^{7,8} In this issue, Dr. Graham Worrall and colleagues (page 1705) show there is little evidence that guidelines have had a positive impact on patient outcomes in primary care settings. They express concern that such informational "technologies" may be embraced before they have been properly evaluated. In reviewing guideline studies, the authors regard CPGs as a "therapy," targeted at physicians, that should be judged in much the same way as other therapeutic interventions.⁸ Under this magnifying glass, guidelines, guideline interventions and evaluations of guidelines all come up wanting.⁹

Should we blow the whistle on guidelines? Developing guidelines is time-consuming and expensive. Well respected and widely endorsed guidelines have been promoted in primary care for well over a decade,¹⁰ but if all this hard work has failed to improve patient outcomes are we wasting our time? Does implicit policy development work just as well as explicit development? Should we shift our attention to other practice management strategies?

Before giving up on guidelines, we should consider what an ideal guideline study might look like, whether such a study could be done, what the main threats to validity would be and what less-than-ideal studies we would be inclined to believe.

Editorial

Éditorial

Dr. Hayward is with the Departments of Medicine and of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ont.

Can Med Assoc J 1997;156:1725-7



Guideline trials

Studies of health care are concerned with problems, populations, exposures and outcomes. Evaluations of health services start with a defined problem in a defined population. The population is exposed to an intervention to test the hypothesis that the intervention reduces unwanted outcomes and increases desirable ones. Let us take influenza immunization as an example. Here, the intervention is simple. There are few exclusion criteria to limit the population of interest, an injection is easy to document, and an important outcome can be measured by rates of hospital admission for pneumonia. Trials of influenza vaccination in elderly people can be judged fairly on grounds of patient selection, successful randomization, blinded outcome assessment, completeness of follow-up, effect size, statistical significance and applicability of the results.

However, most primary care interventions are more complex. For example, in the management of hypertension, asthma, diabetes or obesity it is harder to define the population (eligible patients must be detected and those with complex comorbidity must be excluded), the intervention (there are usually multiple interventions) and the outcome (there may be many intermediate outcomes).

Guidelines for complex interventions are hard to build. Developers must rely on frameworks such as “causal pathways” to help organize evidence about the effectiveness of such interventions and to identify potential difficulties.¹¹ For example, practitioners may over- or underdiagnose the condition; diagnoses may or may not lead to 1 or more interventions (which may or may not alter intermediate and final outcomes); and harmful effects of detection and treatment may occur at various points. The strength of recommendations can be compromised by weak evidence or skewed values at any step in the causal pathway.

Studies of complex guideline interventions are vastly more complex. To prove that a CPG leads to better patient outcomes, both “causal” and “policy” pathways must not break at any of many steps. The guideline must be and *be seen to be* valid, important and applicable. The CPG implementation strategy must be clear, specific and distinct from a control intervention, and the processes and outcomes measured should be appropriate to both the guideline and its application. More specifically:

- The goal of the CPG should match the goals of the targeted patients and practitioners.
- The clinical problem should be relevant and important.
- The clinical-decision model should include options and outcomes that do not oversimplify, distort or otherwise limit the applicability of the CPG in typical care settings.

- Rules of evidence and values must be credible.
- Analysis of evidence and values should be appropriate, timely, complete and valid.
- Resulting recommendations should differ from prevailing practices and implicit policies.
- The necessary skills, equipment and processes should be available to the target population.
- Practitioners should have confidence in the recommendations and their developers.
- Practitioners should be inclined to implement the recommendations at the right time and place.
- The specific implementation of the CPG should be defined, and expected changes in health care processes and outcomes declared.
- Implementation of the CPG should follow the recommendations exactly.
- Inappropriate, variable or unwanted health care practices or outcomes should be of sufficient magnitude that an interventional study could cause and detect improvement.
- Assessment of associations between outcome and intervention must be free of significant bias.

CPG implementation studies should include a control group. This can be problematic. The control intervention could be “usual” policy implementation practices, the “next best” implementation strategy or no intervention. If a CPG represents the best evidence about an important problem, it may be unethical not to give the control group the benefit of the CPG recommendations. Similarly, if there is any risk that the CPG does not reflect best evidence, it may be unethical to use it in the experimental group.

The control group should match the experimental group in all respects save the specific CPG recommendations. Unfortunately, even a randomized study may not ensure that this happens. Hawthorne effects — that is, the knowledge that one is participating in a study — can contaminate the experimental group, systematically exposing it to factors unrelated to guidelines and not experienced by the control group. For example, some guidelines advocate or require local adaptation of recommendations. This involvement of practitioners in policy discussions may itself reduce practice variation independently of any specific recommendation. Even technologic interventions such as computerized reminder systems can exert positive effects by causing clinicians to set aside time for such things as preventive care, or negative effects by virtue of being impractical or unappealing. To test the clinical validity of a specific CPG, the experimental and control groups should be exposed to the same reminder system but different recommendations.

Studies that evaluate CPGs may use informational interventions, such as reminder systems or education ses-



sions. One could argue that in effect these studies assess informational interventions and would yield the same result if the subjects did not know that a CPG was at issue. Results from studies of guideline-based computerized reminder systems are consistent with results from similar studies done in the pre-guideline era. To test the incremental impact of evidence packaged as CPGs, both the experimental and the control groups could be exposed to the same recommendations, labelled as a guideline in the experimental group but not so labelled in the control group.

Given the variable context in which CPGs are implemented and evaluated, randomized controlled trials with quantitative outcome assessment may fail to capture what is important about the success or failure of guidelines to improve health and health care. It is possible that qualitative research methods, with careful attention to the changing experience of clinicians and patients, may yield more immediately useful insights. If so, levels of evidence based on a hierarchy of study designs, as is now popular, may not be an appropriate metric in trials of CPGs.

Guideline trials on trial

If the intent of evidence-based guidelines is to summarize knowledge about the efficiency and effectiveness of health care, then what can we deduce from the statement that there is "little evidence that the use of CPGs produces significant changes in clinical outcomes in primary care"? Despite the developer's best efforts, it is possible for a guideline to distort a clinical problem in a way that leads to unintended or unwanted outcomes when the guideline is implemented. A guideline meant to cut costs could actually increase overall costs. It is important to test and validate CPGs.

Tests of guidelines, however, are difficult to design and execute. A failure to demonstrate improved outcomes is uninformative unless the guideline is valid, the intervention is appropriate to the guideline, both guideline and intervention are acceptable to patients and practitioners, and the outcomes can in fact be improved by changes in practice. Moreover, as my colleagues and I discuss in this issue (page 1715), we know that the same CPG recommendations can be welcomed or shunned by physicians depending upon who is listed as the guideline developer.¹² Many physicians first encounter guidelines in the context of peer review, utilization management and quality control programs — experiences that they may not perceive positively. Indeed, the "political" baggage associated with a guideline, or with all guidelines, may be enough to compromise a study.

A sobering observation by Worrall and colleagues is that all but 2 of the 13 guideline intervention studies

they reviewed had serious methodologic flaws. None would satisfy the more exacting requirements for guideline intervention studies that I have outlined. There are many impediments to an "ideal" CPG intervention study, and the current literature appears to be marred by a lack of solid information about the efficacy of information dissemination strategies and practice management interventions. A whole new generation of guideline-based software, care maps and disease management systems is only now mature enough to admit systematic evaluation. There is an urgent need for new knowledge about how CPGs can be packaged and delivered to maximize their acceptance by clinicians and their impact on care. Until this evidence appears, we cannot conclude that the guideline movement is itself misguided.

The author acknowledges the support of the Physicians' Services Incorporated Foundation.

References

1. Institute of Medicine. Field MJ, Lohr KN, editors. *Clinical practice guidelines: directions for a new program*. Washington: National Academy Press; 1990.
2. Hayward RSA, Wilson MC, Tunis SR, Bass EB, Guyatt G, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. VIII: How to use clinical practice guidelines. A: Are the recommendations valid? *JAMA* 1995;274:570-4.
3. Wilson MC, Hayward RSA, Tunis SR, Bass EB, Guyatt G, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. VIII: How to use clinical practice guidelines. B: What are the recommendations and will they help you in caring for your patients? *JAMA* 1995; 274:1630-2.
4. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet* 1993;342:1317-22.
5. Weingarten S, Stone E, Hayward RSA, Tunis S, Pelter M, Huang H, et al. The adoption of preventive care practice guidelines by primary care physicians: Do actions match intentions? *J Gen Intern Med* 1995;10(3):138-44.
6. Weingarten SR, Reidinger MS, Conner L, Riedinger MS, Lee TH, Hoffman I, et al. Practice guidelines and reminders to reduce duration of hospital stay for patients with chest pain: an interventional trial. *Ann Intern Med* 1994;120:257-63.
7. Freemantle N, Harvey EL, Wolf F, Grimshaw JM, Grilli R, Bero LA. Printed educational materials to improve the behaviour of health care professionals and patient outcomes. In: Bero L, Grilli R, Grimshaw J, Oxman A. *Collaboration on effective professional practice module of the Cochrane Database of Systematic Reviews* [updated 3 Mar 1997]. Available in the Cochrane Library. The Cochrane Collaboration; issue 2. Oxford (UK): Update Software; 1997 [updated quarterly].
8. Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *Can Med Assoc J* 1995;153:1423-31.
9. Sackett DL; Haynes RB; Guyatt GH, Tugwell P. *Clinical epidemiology: a basic science for clinical medicine*. 2nd ed. Toronto: Little, Brown & Co; 1991.
10. Hayward RSA, Steinberg EP, Ford DE, Roizen MF, Roach K. Preventive care guidelines: 1991. *Ann Intern Med* 1991;114:758-83.
11. Battista RN, Fletcher SW. Making recommendations on preventive practices: methodological issues. *Am J Prev Med* 1988;4(4 suppl):53-67.
12. Tunis SR, Hayward RSA, Wilson MC, Rubin HR, Bass EB, Johnston M, et al. Internists' attitudes about clinical practice guidelines. *Ann Intern Med* 1994;120:956-63.

Reprint requests to: Dr. Robert S.A. Hayward, Department of Clinical Epidemiology and Biostatistics, McMaster University Faculty of Health Sciences, Rm. 3H7, 1200 Main St. W, Hamilton ON L8N 3Z5; fax 905 546-0401; haywardr@fhs.csu.mcmaster.ca