Reviewer's report

Title: Twelve years of clinical practice guideline development, dissemination and evaluation in Canada (1994 to 2005)

Version: 1 Date: 29 October 2008

Reviewer: Antoine Boivin

Reviewer's report:

This study seeks to describe how guidelines are being developed, implemented, and evaluated in Canada. It uses a survey questionnaire which was used in a study published by Graham and colleagues in 2003, which covered guidelines produced between 1994 and 1999. This paper addresses two new questions which were not covered in the 2003 article: 1) how guideline development, dissemination, and evaluation have changed over time (between 1994-1999 and 2000-2005), and 2) are there differences between frequent and infrequent developers with regard to how they produced, evaluated, and disseminated guidelines over the 12-year study period?

This study asks questions that are relevant and useful. The use of surveys to monitor the process of guideline development and implementation is extremely relevant to the Canadian context. Guideline development in Canada and the United States, unlike many countries in Europe, is characterized by a mosaic of different organizations and it is very difficult for observers to get a general picture of “who is doing what”. Furthermore, the historical perspective added by this study is useful as a number of important developments have occurred in the past decade in the guideline world, such as the emergence of international quality standards for CPG development (AGREE), and new research developments on knowledge translation and guideline implementation strategies. Finally, there seems to be recent trends toward greater centralization of guideline development is some parts of Canada (ex. the creation of the Institut National d’Excellence en Santé, in Quebec), which makes the comparison between frequent and infrequent guideline developers timely.

Although I believe this paper could be of interest to researchers in the field and make a good contribution to the guideline development and implementation, I do have some reservations regarding how the survey is being analysed and reported, as described below.

Major Compulsory Revisions

1. Response rate

One of the main concern I have with the study relates to how the response rate is reported. I am not clear with the authors’ rationale to calculate their response rate from the number of listed guidelines rather than the number of guideline development groups or guideline development organizations to whom surveys
were sent. The authors mention that they calculated response rates “as with the previous study”, but the 2003 paper by Graham and al. does not provide further explanations.

There are two different levels of guideline development embedded in this survey: 1) guideline development organizations; 2) guideline development groups producing individual guidelines. One could assume that the process of guideline development/implementation may be different between individual guidelines produced within the same organisations, which could provide some justification for using the total number of guidelines produced as a denominator. This however, would not justify counting the English and French version of the same guideline. Also, because the authors compare frequent and infrequent guideline development organizations in the second part of their study, this level of analysis is also relevant and should be reported when calculating response rate.

The current reported response rate could be potentially misleading for readers. In my opinion, 1) a first response rate should be calculated by counting translated versions of guidelines only once, just like they are being analyzed; 2) response rates for guideline development organizations should also be reported as it relates with one of the main research question in this paper.

2. Study population and comprehensiveness of the CMA database

As the authors appropriately point out in the discussion section, the comprehensiveness of the CMA database is critical to infer results. However, I do not share CMA’s belief that their database “represents the vast majority of CPGs published in Canada”. At least two areas would be of concern, namely the coverage of the database for francophone guidelines, and those produced for non-physicians. The Canadian Medical Association is an association whose membership is restricted to physicians and, although its database is available to everybody, guidelines produced specifically for non-physicians may be disseminated elsewhere. Secondly, although the vast majority of physicians in the English-speaking provinces of Canada are member of the CMA, less than 50% of practicing physicians in the francophone province of Quebec are registered member of this association. For example, I was unable to retrieve from the CMA database a set of eleven antibiotic guidelines on various conditions widely used in Quebec and produced in 2005 by the Conseil du Médicament du Québec (one of the largest governmental organization in the province involved in guideline development, available at www.cdm.gouv.qc.ca). This could partly explain why so few CPGs published only in French are available in the survey. I would suggest that some of these limitations of the CMA database be incorporated in the discussion section.

3. Survey instrument

Nor the 2003 or the current paper describe how the survey instrument was developed and validated. This is important to help the reader interpret how reported methods of guideline development and implementation may correlate with the actual process used by guideline developers. I am wondering, for example, if the apparent decrease from 99.6 to 94.3% in the number of guideline
development groups who use a “scientific literature review” may simply reflect change in reporting and an evolution of what developers consider to be a “scientific” literature review, rather than an actual lowering of the standards of CPG development, as suggested by the authors in the discussion. In any case, it would be useful if the authors could include the two page questionnaire as an appendix to their paper so that readers may assess for themselves the face validity of the questions asked.

Minor Essential Revisions

4. In line with the previous comments on study population, table 3-4-5 could be clarified by specifying what the “n” refers to (eg. guidelines/guideline development groups) as opposed to the “n” for guideline development organization in table 2.

5. The confidence intervals of many of the results presented in table 2 overlap without including mean values and it is therefore difficult for the reader to assess the statistical significance of the observed differences. It would be useful to add a "change column" in table 2 (as in tables 3, 4, and 5). Tables 2 to 5 should also include the p values for the difference (with H0 being no difference between infrequent and frequent developers or between older and recent guidelines), rather than only mention that some results are "non significant" (which cannot be interpreted unless your threshold to assess statistical significance is known).

Discretionary Revisions

6. The first sentence of the second paragraph in the background section could be changed to emphasize the fact that we refer to a previously conducted/published study and not the current one.

7. The 2003 paper reported on the composition of guideline development groups, including the presence of patient representatives. I understand that committee member characteristics were “not available for analysis for the 2000-2005 period", which is unfortunate, given the increasing prominence of patient involvement in guideline development in the international literature: was this component dropped from the survey? For which reason?

8. It may be useful to describe (in the appendix section? in a footnote?) the criteria that were used to classify guideline organizations and distinguish between a governmental and para-governmental organization, or between a professional, a medical, and a health association.

9. I would not be afraid to describe in more details the relevance and contribution of this study in the introduction or discussion section.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the
statistics.

**Declaration of competing interests:**

I declare that I have no competing interests