

Should Guidelines Incorporate Evidence on Patient Preferences?

Craig A. Umscheid, MD, MSCE^{1,2,3}

¹Division of General Internal Medicine, University of Pennsylvania School of Medicine, Philadelphia, PA, USA; ²Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, Philadelphia, PA, USA; ³Center for Evidence-based Practice, University of Pennsylvania Health System, Philadelphia, PA, USA.

J Gen Intern Med 24(8):988–90

DOI: 10.1007/s11606-009-1055-0

© Society of General Internal Medicine 2009

Patient preferences are critically important to clinical decision making and patient care. Few clinicians practicing modern day medicine would argue otherwise. This is the case despite the fact that shared decision making can often be difficult for us to accomplish in our everyday practices.^{1–5} Traditional models for incorporating patient preferences into clinical decisions have described the intersection of patient values with clinical evidence and resource considerations at the point-of-decision (Figure 1).⁶ More recently, there have been calls to move beyond the mere integration of patient preferences at the point-of-decision to the more systematic integration of patient preference data into guidelines themselves.^{7,8}

In this issue of JGIM, Chong and colleagues advance this trend one step further by examining how well evidence on patient preferences is incorporated into clinical practice guidelines (CPGs).⁹ To do this, they critically appraise CPGs recommended by the Guidelines Advisory Committee (GAC) in Ontario, Canada. As described by Chong, “the GAC conducts annual surveys to identify clinical topics of importance in Ontario, performs literature searches to identify existing CPGs on those topics, and then asks community-based physicians trained in CPG appraisal to evaluate the guidelines using the internationally developed AGREE instrument.¹⁰ The GAC then endorses a single guideline on the basis of quality and relevance to Ontario practitioners.”⁹ To appraise how well evidence on *patient preferences* was sought and integrated into these guidelines, the authors adapted two instruments originally designed to assess the overall quality of CPGs—the AGREE instrument and a tool developed by Shaneyfelt and colleagues¹¹. The authors also quantified how much of the CPG text was actually devoted to patient preferences by counting words and references related to preference issues. For the purposes of comparison, a similar analysis was done to appraise how well evidence on *treatment effectiveness* was sought and integrated into these guidelines. Their results were not surprising. Although the guidelines included in the study were highly rated by GAC (mean 3.36 out of 4), relatively recent (published between 1999 and 2003), authored by government institutions or medical associations in the US or Canada, and were general guidelines for a disease, the scores for the quality of integration of patient preference data were low. The overall mean scores of the AGREE and

Shaneyfelt tools were statistically lower for the integration of *patient preference* information than *treatment effectiveness* information (mean AGREE scores on a scale of 0.25 to 1.00: 0.43 vs. 0.65 for *preference* and *effectiveness* data, respectively; mean Shaneyfelt scores on a scale of 0 to 1: 0.18 vs. 0.58 for *preference* and *effectiveness* data respectively; $p < 0.001$ for both comparisons). In addition, the percentages of text and references dedicated to *patient preferences* were significantly lower than those addressing *treatment effectiveness* (average percentage of the total word count: 4.6% vs. 24.2% for *preference* and *effectiveness* data, respectively; average percentage of references: 6.0% vs. 36.6% for *preference* and *effectiveness* data, respectively). The authors didn't find any differences in the integration of patient preference data when comparing guidelines that may be less preference sensitive (like sepsis guidelines) with those that may be more preference sensitive, such as those with poor quality effectiveness or safety data, or those that include many trade-offs between risks and benefits (like aggressive hypertension treatment in the elderly). Yet the study wasn't powered to find these differences. The authors did find that more recent guidelines were weakly correlated with higher scores for incorporation of patient preference data, which makes sense given the newness of the field of patient preference measurement. The authors conclude that their study “empirically demonstrates that high quality CPGs from a variety of disciplines generally do not systematically seek or integrate evidence on patient preferences.”⁹ They also nicely outline in a conceptual model where patient preferences most impact clinical decisions.

What concerns me about this study is the authors' implicit assumption that it is necessary to include a systematic review of the literature on patient preferences in every clinical guideline that's developed. Although the objective of this study is not to support this assumption, I think this should have been a primary objective. Instead of skipping to the question “*How well* do guidelines incorporate evidence on patient preferences?”, we should be asking the question that naturally precedes it—“*Should* guidelines be incorporating evidence on patient preferences?” Whereas evidence of efficacy, effectiveness and harm from population based research is necessary to inform clinical guidelines and clinical decisions at the point-of-care, it's unclear whether population-based patient preference evidence is necessary to inform individual patient decisions. Patients have their own individual preferences that will inform their clinical care, and those preferences will inform their care regardless of the existing population-based preference data (this is, of course, assuming that clinicians identify their patients' preferences, or that patients volunteer them^{1–5}). This is not to say that population-based patient preference data are without value. One could envision many scenarios where it

Published online July 3, 2009

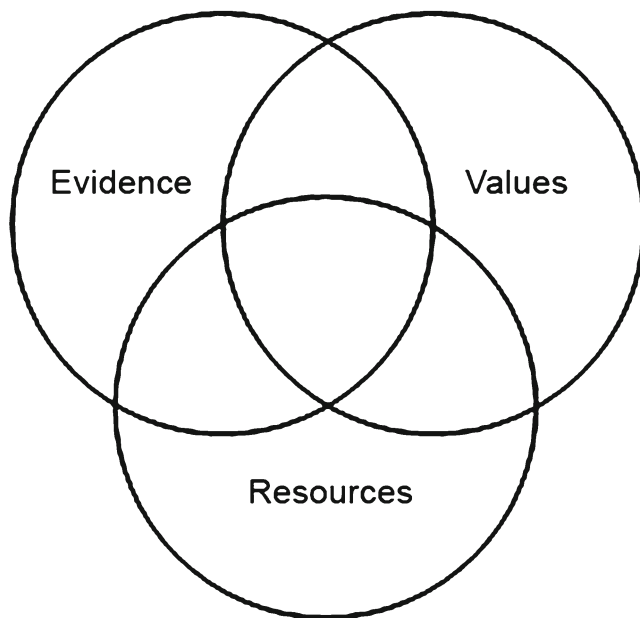


Figure 1. Venn diagram depicting the intersection of evidence, resources and values in the medical decision-making process. From J.A. Muir Gray. *Evidence-based Healthcare*. Second Edition. 2001

could help formulate recommendations or directly affect the impact of a guideline. For example, patient preference data could help us understand how many patients might refuse guideline recommended care, and this data could help inform pay-for-performance metrics. Or guidelines that incorporate patient preferences into their recommendations could be written for patients and could be used by patients to direct their own medical care. But the benefits of integrating population-based patient preference data into guidelines should be examined and demonstrated, and not assumed.

Besides the lack of evidence around the value of including patient preference data in guidelines, there are real risks (and opportunity costs) to including this type of data. First, there is enough of a knowing-doing gap¹² that it seems additional types of data and data synthesis could further obscure the clarity of guideline recommendations, resulting in less usable guidelines and a growing knowing-doing gap. As the length and complexity of guidelines grow, they potentially become less accessible to medical and nursing leaders and practitioners on the ground, decreasing their impact. In their discussion, Chong and colleagues make this point by describing a study where the simplification of hypertension treatment into an abridged algorithm allowed more uncomplicated patients to reach blood pressure targets than using CPG-based practice.¹³ In addition, increasing complexity and growing guideline standards make the development of targeted, systematic and timely guidelines even more challenging, bogging down the translation of evidence into practice. This is a more vexing problem when we consider how quickly guidelines can become out of date.¹⁴ Requiring guidelines to systematically include patient preference data could also be a low-yield pursuit, as such data is likely nil for most guideline topics. It's difficult enough to find valid data on efficacy, effectiveness and safety. We'd only be creating whole new areas devoid of data. And I'm not sure that we would want to promote this area of research

to the detriment of comparative effectiveness research, which is sorely needed. Lastly, Shaneyfelt and colleagues have told us that many guideline developers aren't even adhering to the current standards of guideline development¹¹, so adding more standards could detract from efforts to improve guideline development based on the current standards. Interestingly, in Shaneyfelt's study, a higher percentage of guidelines addressed patient preferences than specified methods for identifying scientific evidence (21.5% versus 16.8%).¹¹ So before we start deducting points from guidelines for not including patient preference data, let's at least examine whether the inclusion of such data makes a difference in guideline recommendations or guideline impact. Opportunity costs exist for every additional requirement we place on guidelines.

Accepting the authors' assumption that patient preference information should be systematically integrated into CPGs, I'm not sure that the authors' measures of "preference data integration" are valid. Many "state of the art" guidelines apply a strength to each of their recommendations, and those recommendation schemes most often delineate "preference sensitive" or weaker recommendations from stronger recommendations which are less "preference sensitive."^{15,16} Thus, considerations about patient preferences are directly integrated into such rating schemes so that clinicians and other decision makers (including payers) know for which recommendations individual patient preferences are critical and where they may be less critical. To my understanding, the authors' measures of "preference data integration" do not take the subtlety of these schemes into account.

In terms of the authors' modification of different quality scales to measure the quality of the integration of preference data, I think they do an excellent job of describing their efforts in the study's appendix, but I have concerns about the validity of the quality criteria they've developed. First, some of the quality criteria, like the AGREE item requiring explicit links between evidence and recommendations, rely on the availability of patient preference data of reasonable quality—yet, such data may not be available for topics other than the common examples of PSA screening and anticoagulation for atrial fibrillation. Second, some of the quality criteria for preference integration are dependent on methods that have yet to be developed or for which there are no generally accepted methods. This is the case for the modified Shaneyfelt criteria that "formal methods were used to combine evidence on preference data" and "benefits and harms of preference data have been quantified."⁹

The word and reference counts also have no face validity as a quality measure of patient preference data integration. Such a measure neglects the availability or quality of such data for a given guideline. This important limitation is described in the authors' discussion, but they address it only semantically by suggesting that the counts are not measuring the quality of data integration, but are merely measuring the quantity. But it is clear from the manuscript that these quantitative data are being used to support an argument that the percentages of references devoted to patient preferences are too low. I would argue that it's quite unclear that the percentages of references for patient preference data are too low. They might be higher than many would anticipate. The percentages of patient preference references in the guidelines examined may actually be much higher than those in standard databases of the medical literature, like MEDLINE or EMBASE. Thus, the

interpretation that these reference counts reflect poor integration may not be accurate.

Instead of adapting the AGREE or Shaneyfelt scales to measure the quality of integration of patient preference data, or counting words or references, the authors could have developed and used more valid measures. For example, the authors could have described the number of preference-related references that could have potentially been included in CPGs but weren't based on a search of the patient preference literature. This would have accounted for the lack of patient preference data available for most guideline topics. The authors could have also estimated the additional time needed to search for these references, extract the data, and synthesize the data in the evidence analysis and into the guideline recommendations. In addition, to determine whether the exclusion of such data was meaningful, the authors could have described whether recommendations would change if the excluded data were in fact included.

Outside the scope of Chong's study, there is another set of preferences already included in every guideline that is developed—the preferences of the guideline developers themselves. These preferences mold the evidence review and the resulting recommendations. As such, these preferences need to be explicitly described in guidelines so that the consumers of those guidelines can understand and adjust for the values and potential biases of the guideline developers. But that's another editorial.^{8,17,18}

Corresponding Author: Craig A. Umscheid, MD, MSCE, Center for Evidence-based Practice, University of Pennsylvania Health System, 3535 Market Street, Mezzanine, Suite 50, Office 108, Philadelphia, PA 19104, USA (e-mail: craig.umscheid@uphs.upenn.edu).

REFERENCES

1. **Barry MJ.** Involving patients in medical decisions: how can physicians do better? *JAMA*. 1999;282(24):2356–7.
2. **Braddock CH 3rd, Edwards KA, Hasenberg NM, Laidley TL, Levinson W.** Informed decision making in outpatient practice: time to get back to basics. *JAMA*. 1999;282(24):2313–20.
3. **Coulter A.** Paternalism or partnership? Patients have grown up—and there's no going back. *BMJ (Clinical research ed)*. 1999;319(7212):719–20.
4. **Gravel K, Legare F, Graham ID.** Barriers and facilitators to implementing shared decision-making in clinical practice: a systematic review of health professionals' perceptions. *Implement Sci*. 2006;1:16.
5. **Hanson JL.** Shared decision making: have we missed the obvious? *Arch Intern Med*. 2008;168(13):1368–70.
6. **Muir Gray JA.** Evidence-based healthcare. 2nd ed. Philadelphia: Elsevier; 2001. 11.
7. **Krahn M, Naglie G.** The next step in guideline development: incorporating patient preferences. *JAMA*. 2008;300(4):436–8.
8. **Shaneyfelt TM, Centor RM.** Reassessment of clinical practice guidelines: go gently into that good night. *JAMA*. 2009;301(8):868–9.
9. **Chong CAKY, Chen I, Naglie G, Krahn MD.** How well do guidelines incorporate evidence on patient preferences? *J Gen Intern Med*. 2009; doi:10.1007/s11606-009-0987-8
10. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Quality & Safety in Health Care*. 2003;12(1):18–23.
11. **Shaneyfelt TM, Mayo-Smith MF, Rothwangl J.** Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. *JAMA*. 1999;281(20):1900–5.
12. **McGlynn EA, Asch SM, Adams J, et al.** The quality of health care delivered to adults in the United States. *N Engl J Med*. 2003;348(26):2635–45.
13. **Feldman RD, Zou GY, Vandervoort MK, Wong CJ, Nelson SA, Feagan BG.** A simplified approach to the treatment of uncomplicated hypertension: a cluster randomized, controlled trial. *Hypertension*. 2009;53(4):646–53.
14. **Shekelle P, Eccles MP, Grimshaw JM, Woolf SH.** When should clinical guidelines be updated? *BMJ (Clinical research ed)*. 2001;323(7305):155–7.
15. **Guyatt GH, Oxman AD, Vist GE, et al.** GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924–6.
16. **Guyatt GH, Oxman AD, Kunz R, et al.** Going from evidence to recommendations. *BMJ*. 2008;336(7652):1049–51.
17. **Strech D, Tilburt J.** Value judgments in the analysis and synthesis of evidence. *J Clin Epidemiol*. 2008;61(6):521–4.
18. **Sniderman AD, Furberg CD.** Why guideline-making requires reform. *JAMA*. 2009;301(4):429–31.