

Editorials

Where Clinical Practice Guidelines Go Wrong

Henry C. Barry, MD, MS, Michigan State University, East Lansing, Michigan

Lisa Cosgrove, PhD, University of Massachusetts, Boston, Massachusetts

David C. Slawson, MD, University of North Carolina, Charlotte, North Carolina

Clinical practice guidelines, essential tools for evidence-based practice, have proliferated over the past few decades but remain controversial for several reasons, including discordance among guidelines, the influence of intellectual and financial conflicts of interest, and a lack of adherence to standards for developing trustworthy guidelines. In this editorial, we will use recent hypertension guidelines to illustrate guideline discordance and then propose how to identify trustworthy guidelines.

The best guidelines are developed using processes with several important attributes¹⁻³:

- Comprehensive, systematic evidence search
- Evidence linked directly to recommendations and a strength of recommendation
- Emphasis on patient-oriented rather than disease-oriented outcomes
- Transparency
- Minimal conflicts of interest
- Prospective validation
- Recommendations that offer flexibility in various clinical situations

Multiple organizations have issued guidelines on managing patients with hypertension. A recent analysis identified eight guidelines, released between 2010 and 2018, with 71 separate recommendations.⁴ However, three recommendations were unique to a single guideline, leaving 68 for the analysis. The recommendations included variations on blood pressure measurement methods, laboratory and other test indications, and treatment thresholds. Each recommendation was rated on the direction (e.g., do XYZ or do not do XYZ) and the overall strength. Twenty-two (32%) of the recommendations were concordant in both the direction and strength. However, 28 (41%) of the recommendations had inconsistent directions. For example, the guidelines recommended varying treatment targets (e.g., target blood pressure less than 150/90, 140/90, or 130/80 mm Hg, or systolic pressure less than 120 mm Hg) and varying preferred initial therapies (e.g., thiazides, angiotensin-converting

enzyme inhibitors, calcium channel blockers, beta blockers). Sensitivity analyses to determine the role of the underlying strength of ratings, source of recommendations, and importance of recommendations failed to reveal any causes for guideline inconsistency. We suggest that several guideline development factors not addressed by the authors are likely explanations:

- **Interpretation:** The developers of the guidelines value some outcomes more than others.
- **Speculation:** When evidence is lacking, experts provide their best guess.
- **Extrapolation:** Recommendations extrapolate beyond the research, whether to specific interventions (e.g., blood pressure targets) or to the study populations.
- **Representation:** The panel of developers lacks relevant stakeholders (e.g., patients, primary care physicians).
- **Oversimplification:** Guidelines apply a one-size-fits-all approach that ignores nuances about disease severity, patient comorbidity, age, and other contextual factors.
- **Overcomplication:** Guidelines address too many specific questions.
- **Application:** Guidelines do not consistently address feasibility and lack clarity on whether they apply to individuals or systems.
- **Money:** Developers have conflicts of interest, or they do not address the costs of guideline implementation.

The true measure of a guideline is whether clinicians who use and follow it have improved patient outcomes, preferably in the setting of a randomized trial. However, prospective validation of guidelines rarely occurs. Guidelines resulting in worse patient outcomes compared with other guidelines or watchful waiting should be retired. For example, a registry-based study found that the use of anticoagulant therapy for atrial fibrillation based on guidelines from the European Society of Cardiology was associated with a nearly three-fold increased risk of hemorrhage compared with treatment based on

FIGURE 1

	Yes	Can't tell	No
Relevance and utility			
The recommendations focus on improving patient-oriented outcomes, not disease-oriented outcomes, explicitly comparing benefits vs. harms to support clinical decision-making. <i>How to tell:</i> The recommendations are based on demonstrated direct benefits for patient outcomes and not biochemical markers or risk factors.		Stop	Stop
The recommendations are clear and actionable. <i>How to tell:</i> The recommendations provide explicit guidance. If there is no decision tree or algorithm, there should be sufficient detail to inform collaborative decision-making in your clinical setting.			
The patient populations and conditions are relevant to my clinical setting. <i>How to tell:</i> The guideline should explain the target conditions, target populations, practice settings, and audience to which the recommendations apply. Do the recommendations apply to your practice?			
Trustworthiness			
The guidelines are based on a systematic review of the research data. <i>How to tell:</i> Determine whether the recommendations are linked to a systematic review of the available literature. If there is no mention of a systematic literature search, the guideline is not trustworthy.		Stop	Stop
The recommendations important to you are based on graded evidence and include a description of the quality (e.g., strong, weak) of the evidence. <i>How to tell:</i> GRADE, SORT, U.S. Preventive Services Task Force, or other strong evidence-rating systems are used to grade the available evidence and the majority of the recommendations are supported by high-quality evidence.		Stop	Stop
The guideline development team includes a research analyst, such as a statistician or epidemiologist. <i>How to tell:</i> A research analyst (statistician, epidemiologist, or other qualified independent methodologist) is listed in the working group description, or an evidence review is conducted by a group separate from the guideline development group.			
Interpretation			
The chair of the guideline development group and a majority of the rest of the committee are free of declared financial conflicts of interest, and the guideline development group did not receive industry funding for developing the guideline. <i>How to tell:</i> Find and examine the conflicts of interest statement. It is usually at the end of the document.			
The guideline development group includes members from the most relevant specialties and includes other key stakeholders, such as patients, payer organizations, and public health entities, when applicable. <i>How to tell:</i> Guideline development groups should have representatives from applicable specialties and, when possible, patients or consumer advocacy groups.			
			Total

Scoring

Any "Stop" items: guideline not useful.

"No" answers for other items: 0 to 1 = useful guideline; 2 = may not be useful; 3 to 5 = not useful.

G-TRUST = guideline trustworthiness, relevance, and utility scoring tool; GRADE = Grading of Recommendations, Assessment, Development and Evaluations; SORT = Strength of Recommendation Taxonomy.

G-TRUST tool for clinicians to independently assess the trustworthiness and utility of guidelines. This scoring system applies only to guidelines written or updated in the past five years.

Copyright © 2017 Allen F. Shaughnessy, PharmD, MMedEd, and Lisa Cosgrove, PhD. Used with permission.

guidelines from the American College of Chest Physicians, which did not recommend anticoagulant therapy for older people with a CHADS₂ (congestive heart failure, hypertension, age, diabetes, stroke) score of 0.⁵

What should clinicians do? First, they should be skeptical. Next, they should identify a trusted source of guidelines that follow good practices. For example, a collection of Practice Guidelines published in *American Family Physician* is available at <https://www.aafp.org/afp/practguide>. The Emergency Care Research Institute Guidelines Trust (<https://guidelines.ecri.org/>) is a non-profit entity that provides a repository of guidelines accompanied by a scorecard that grades the guidelines against the standards developed by the National Academy of Medicine, which are described in this Editorial. Although this resource is free, it requires registration, and it is unclear how often the guidelines are updated.

A more rigorous proposition is that clinicians grade guidelines themselves using the Guideline Trustworthiness, Relevance, and Utility Scoring Tool (G-TRUST)⁶ or a similar tool. The G-TRUST (*Figure 1*) is a simple tool for clinicians to independently assess the trustworthiness and utility of guidelines. Although its validation has been limited, its face validity is high because it prioritizes patient-oriented outcomes and adherence to good guideline development processes. Finally, clinicians need to be prepared to stop the use of guidelines found to be ineffective or harmful.

Sir Muir Gray of the United Kingdom National Screening Committee famously stated, “All screening programs do harm; some do good as

well.” This is also true for clinical practice guidelines. To do more good than harm to patients, guideline panels must at a minimum use transparent and inclusive processes and eliminate the process factors that result in discordant and potentially biased guidelines.

Address correspondence to Henry C. Barry, MD, MS, at barry@msu.edu. Reprints are not available from the authors.

Author disclosure: No relevant financial relationships.

References

1. Lin KW, Slawson DC. Identifying and using good practice guidelines. *Am Fam Physician*. 2009;80(1):67-70. Accessed November 9, 2021. <https://www.aafp.org/afp/2009/0701/p67.html>
2. Institute of Medicine Committee on Standards for Developing Trustworthy Clinical Practice Guidelines. Graham R, Mancher M, Wolman DM, et al., eds. *Clinical Practice Guidelines We Can Trust*. National Academies Press; 2011. Accessed November 9, 2021. <https://www.ncbi.nlm.nih.gov/books/NBK209539/>
3. Qaseem A, Kansagara D, Lin JS, et al. The development of clinical guidelines and guidance statements by the clinical guidelines committee of the American College of Physicians: update of methods. *Ann Intern Med*. 2019;170(12):863-870.
4. Alper BS, Price A, van Zuuren EJ, et al. Consistency of recommendations for evaluation and management of hypertension. *JAMA Netw Open*. 2019;2(11):e1915975.
5. Andrade AA, Li J, Radford MJ, et al. Clinical benefit of American College of Chest Physicians versus European Society of Cardiology guidelines for stroke prophylaxis in atrial fibrillation. *J Gen Intern Med*. 2015;30(6):777-782.
6. Shaughnessy AF, Vaswani A, Andrews BK, et al. Developing a clinician friendly tool to identify useful clinical practice guidelines: G-TRUST. *Ann Fam Med*. 2017;15(5):413-418. Accessed November 9, 2021. <https://www.annfammed.org/content/15/5/413.long> ■