

# Clinical Practice Guidelines in Anesthesiology: Adjusting Our Expectations

Mark D. Neuman, M.D., M.Sc., Jeffrey L. Apfelbaum, M.D.

Clinical practice guidelines are “statements that include recommendations intended to optimize patient care” based on “a systematic review of evidence and an assessment of the benefits and costs of alternative care options”<sup>1</sup> Clinical practice guidelines are valued resources for practitioners,<sup>2</sup> especially in anesthesiology, where providers often work in relative isolation from their peers. Triennial surveys of American Society of Anesthesiologists (Schaumburg, Illinois) members have repeatedly ranked American Society of Anesthesiologists practice parameters as a top benefit of membership.<sup>3</sup>

In this issue of *ANESTHESIOLOGY*, Laserna *et al.* review 2,280 recommendations from 60 clinical practice guidelines published by 26 anesthesia societies in North America and Europe over a 10-year period.<sup>4</sup> They find that half of all recommendations in these guidelines were based on the lowest level of evidence evaluated, typically case reports or consensus opinion. Fewer than one in six recommendations was based on the highest level of evidence—data from multiple randomized controlled trials or meta-analyses. The proportion of recommendations based on lower-level evidence did not change over the 10-year study period and did not vary based on the quality of the guideline development process, as judged against widely used standards for rigor and transparency.



**“Failing to understand the limitations of guidelines—particularly those based substantially on opinion or nonrandomized data—can have negative consequences for care and outcomes. Working to increase the number, diversity, and quality of randomized studies in anesthesiology represents the most certain way to improve clinical practice guidelines over time.”**

While these results merit attention, they are not surprising. Similar reviews in other specialties have consistently observed that most current guideline recommendations are not supported by the highest level of evidence.<sup>5–8</sup> Yet the findings of Laserna *et al.* are both important and noteworthy. This is not only because they call attention to a real need for more high-quality randomized trials to guide clinical decision-making in perioperative care. Beyond this, they offer an opportunity to reflect on how we can better align the expectations that practitioners, policy makers, and the public place on clinical practice guidelines with the types of guidance that they are actually able to deliver.

On one level, it may be tempting to view the findings of Laserna *et al.* as an opportunity to dismiss the value of perioperative guidelines. Given that half of the recommendations they identified were based on case reports, consensus opinion, or similar evidence, one might ask whether we should abandon clinical guidelines altogether. In our view, the answer to this question is a definitive no. While randomized trials do remain the most reliable means of comparing treatment

alternatives, other types of guidance—including guidance based on opinion—still have potential to be useful to practitioners. Problems in anesthesia practice often arise ahead of available trial evidence, and certain questions may not be easily amenable to study in randomized trials for a variety

Image: A. Johnson, Vivo Visuals.

This editorial accompanies the article on p. 31.

Accepted for publication April 1, 2021. From the Department of Anesthesiology and Critical Care, Perelman School of Medicine (M.D.N.), the Center for Perioperative Outcomes Research and Transformation (M.D.N.), and the Leonard Davis Institute for Health Economics (M.D.N.), University of Pennsylvania, Philadelphia, Pennsylvania; and the Department of Anesthesia and Critical Care, University of Chicago, Chicago, Illinois (J.L.A.).

Copyright © 2021, the American Society of Anesthesiologists. All Rights Reserved. *Anesthesiology* 2021; 135:9–11. DOI: 10.1097/ALN.0000000000003809

of reasons. For example, there are no major randomized trials demonstrating a positive effect on patient safety of using pulse oximetry in the operating room, yet few clinicians would question the transformative value of this monitor to anesthesiology practice.

For scenarios not easily managed based on clinicians' personal experiences alone, having access to counsel from informed and experienced experts may be preferable to the alternative of no external guidance at all. More generally, clinical practice guidelines can serve functions that go beyond the specific content of their individual recommendations.<sup>9</sup> Guidelines can call attention to new or previously overlooked aspects of practice, and can act as a catalyst for efforts to understand and improve care for specific groups of patients. For example, the appearance of a new guideline on airway management signals not only "here is useful advice for managing difficult airway problems," but also "airway management is an important topic that merits focus and attention."<sup>10</sup>

Yet guidelines also have limitations that need to be taken seriously, as the findings of Laserna *et al.* remind us. Most importantly, guideline recommendations that are based on expert opinion, observational studies, or a single trial are more likely to be wrong when compared to recommendations based on multiple randomized trials. An analysis of changes over time in recommendations developed by major U.S. cardiology societies found that the likelihood of a given recommendation being downgraded, reversed, or omitted was more than three times greater for recommendations based on lower levels of evidence than for those based on data from multiple randomized trials.<sup>11</sup> More broadly, widespread approaches to ensuring rigor in the guideline development process may themselves be flawed.<sup>12</sup> The Grading of Recommendations Assessment, Development and Evaluation (GRADE) process, which was used to develop one in every three recommendations reviewed here, has been found to have poor interrater agreement.<sup>13</sup> And the Delphi method—a widespread approach to developing consensus recommendations—has potential to mold opinion *versus* simply collecting data.<sup>14</sup>

Failing to understand the limitations of guidelines—particularly those based substantially on opinion or nonrandomized data—can have negative consequences for care and outcomes. At the most fundamental level, recommendations do not make care better when they encourage treatments later discovered to be ineffective or potentially harmful, even if these recommendations were based on the best information available at the time. Premature adoption of recommendations or application beyond their intended scope may adversely impact care, particularly when scientific evidence is lacking<sup>15</sup> or when practitioners fail to consider individual patient needs.<sup>16</sup> Efforts to implement recommendations based on opinion or observational data can also interfere with the work of improving care by occupying attention and resources that could be used to promote adoption of interventions supported by more rigorous evidence. Translation

of evidence into practice is labor-intensive and often incomplete even for those interventions found to be effective in randomized trials.<sup>17,18</sup> As such, basing care improvement efforts on less-certain recommendations can run the risk of being wasteful or counterproductive, particularly if those recommendations are subsequently revised, reversed, or abandoned.<sup>15</sup> Last, recommendations based on lower levels of evidence can paradoxically interfere with efforts to make guidelines better over time to the extent that they may unintentionally influence what questions investigators, funders, and ethics review boards consider suitable for study in randomized trials. Even where guidelines acknowledge gaps in evidence, recommendations based on opinion or observational data alone may still come to be seen as establishing practices not yet tested in randomized trials as standards of care, potentially foreclosing opportunities for the types of studies needed to create more certain guidance in the future.

Working to increase the number, diversity, and quality of randomized studies in anesthesiology represents the most certain way to improve clinical practice guidelines over time. In the meantime, such guidelines are likely to remain valued sources of advice for practitioners, even in scenarios where their underlying evidence base may be incomplete or uncertain. Given this, those who produce and use guidelines should be mindful of the potential unintended consequences that may emerge from recommendations based on less *versus* more reliable forms of evidence, and take steps to mitigate such concerns. This may include actively choosing to refrain from making recommendations in areas where available evidence is limited or where misinterpretation of guidance could negatively impact care or constrain needed efforts to generate better evidence. When recommendations are made, guideline-producing organizations can aid efforts to improve care by providing full transparency for the basis of each recommendation and its associated level of uncertainty. Just as importantly, such organizations should work to promote appropriate expectations on the part of clinicians, policy-makers, and the public of guidelines as well-intentioned but imperfect tools designed to help clinicians tailor care to the specific needs of each patient.

## Competing Interests

Dr. Neuman is a past or current member of guideline writing committees for the American College of Surgeons (Chicago, Illinois) and the American Academy of Orthopedic Surgeons (Rosemont, Illinois). Dr. Apfelbaum is past Chair of the American Society of Anesthesiologists (Schaumburg, Illinois) Committee on Standards and Practice Parameters.

## Correspondence

Address correspondence to Dr. Neuman: [neumanm@pennmedicine.upenn.edu](mailto:neumanm@pennmedicine.upenn.edu)

## References

1. Graham R, Mancher M, Wolman D, Greenfield S, Steinberg E, Institute of Medicine (US) Committee on Standards for Developing Trustworthy Clinical Practice Guidelines: Clinical Practice Guidelines We Can Trust. Washington, DC, National Academies Press, 2011
2. Apfelbaum JL, Connis RT: The American Society of Anesthesiologists practice parameter methodology. *ANESTHESIOLOGY* 2019; 130:367–84
3. Apfelbaum JL, Connis RT, Nickinovich DG: 2012 Emery A. Rovenstine Memorial Lecture: The genesis, development, and future of the American Society of Anesthesiologists evidence-based practice parameters. *ANESTHESIOLOGY* 2013; 118:767–8
4. Laserna A, Rubinger DA, Barahona-Correa JE, Wright N, Williams MR, Wyrobek JA, Hasman L, Lustik SJ, Eaton MP, Glance LG: Levels of evidence supporting the North American and European perioperative care guidelines for anesthesiologists between 2010 and 2020: A systematic review. *ANESTHESIOLOGY* 2021; 135:31–56
5. Duarte-García A, Zamore R, Wong JB: The evidence basis for the American College of Rheumatology practice guidelines. *JAMA Intern Med* 2018; 178:146–8
6. Lee DH, Vielemeyer O: Analysis of overall level of evidence behind Infectious Diseases Society of America practice guidelines. *Arch Intern Med* 2011; 171:18–22
7. Tricoci P, Allen JM, Kramer JM, Califf RM, Smith SC Jr: Scientific evidence underlying the ACC/AHA clinical practice guidelines. *JAMA* 2009; 301:831–41
8. Fanaroff AC, Califf RM, Windecker S, Smith SC Jr, Lopes RD: Levels of evidence supporting American College of Cardiology/American Heart Association and European Society of Cardiology guidelines, 2008–2018. *JAMA* 2019; 321:1069–80
9. Weisz G, Cambrosio A, Keating P, Knaapen L, Schlich T, Tournay VJ: The emergence of clinical practice guidelines. *Milbank Q* 2007; 85:691–727
10. Hilgartner S, Bosk CL: The rise and fall of social-problems – a public arenas model. *Am J Sociol* 1988; 94:53–78
11. Neuman MD, Goldstein JN, Cirullo MA, Schwartz JS: Durability of class I American College of Cardiology/American Heart Association clinical practice guideline recommendations. *JAMA* 2014; 311:2092–100
12. Kavanagh BP: The GRADE system for rating clinical guidelines. *PLoS Med* 2009; 6:e1000094
13. Atkins D, Briss PA, Eccles M, Flottorp S, Guyatt GH, Harbour RT, Hill S, Jaeschke R, Liberati A, Magrini N, Mason J, O'Connell D, Oxman AD, Phillips B, Schünemann H, Edejer TT, Vist GE, Williams JW Jr; GRADE Working Group: Systems for grading the quality of evidence and the strength of recommendations II: Pilot study of a new system. *BMC Health Serv Res* 2005; 5:25
14. Dalkey N, Helmer O: An experimental application of the Delphi method to the use of experts. *Manag Sci* 1963; 9:458–67
15. Neuman MD, Bosk CL, Fleisher LA: Learning from mistakes in clinical practice guidelines: the case of perioperative  $\beta$ -blockade. *BMJ Qual Saf* 2014; 23:957–64
16. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J: Clinical guidelines: Potential benefits, limitations, and harms of clinical guidelines. *BMJ* 1999; 318:527–30
17. Lane-Fall MB, Cobb BT, Cené CW, Beidas RS: Implementation science in perioperative care. *Anesthesiol Clin* 2018; 36:1–15
18. Guise JM, Savitz LA, Friedman CP: Mind the gap: Putting evidence into practice in the era of learning health systems. *J Gen Intern Med* 2018; 33:2237–9