

Implications of Practice Variability

Daniel I. Sessler, M.D.

Clinicians strive to practice evidence-based medicine. The difficulty is that little routine care has actually been validated in robust clinical trials. Consequently, many generally accepted clinical approaches are neither supported nor refuted by available research, having instead developed piecemeal from improvements and clinician experience. Perhaps consequently, clinical practices quite reasonably vary considerably among clinicians within institutions, and even more across institutions and around the world. Insufficient knowledge is hardly limited to anesthetic management and extends to surgical practice and all other areas of medicine.

Even across major variations in practice, there is little convincing evidence for that one approach is preferable to another.¹ Consider, for example, the limited supporting evidence for (or compelling evidence against) stress testing or tomographic angiography, volatile anesthetic toxicity in neonates, neuraxial *versus* general anesthesia, intravenous *versus* volatile anesthesia, supplemental oxygen for prevention of surgical site infection, and targeted temperature management for most any indication except neonatal hypoxia. Even less evidence supports more subtle practice differences such as amount and type of intravenous fluid, intraoperative tidal volume, and positive end-expiratory pressure.

Given various approaches to a clinical problem, trials should be able to relatively easily identify the best. In fact, it has not been easy. Most major trials show that primary outcomes are similar with each tested treatment—an observation that applies to drugs, devices, clinical approaches, and health system modifications. For example, an analysis of trials funded by the National Heart, Lung, and Blood



“There is...no basis for giving clinical pathways a ‘free pass’ on evidence. Instead, they should be rigorously tested in context-sensitive robust and properly designed trials—just the way drugs and devices are.”

data, and supportive meta-analyses of small trials. A reasonable question is why well-designed and well-conducted major trials with statistically robust results so often demonstrate that primary results are similar with experimental and reference interventions.

Practice Variability

A potential explanation for large trials so often demonstrating comparable effects is that collectively, clinicians may have determined what matters and what does not. That understanding would drive clinicians to practice uniformly when it matters, while simultaneously allowing harmless deviations. This theory is supported by the fact that there are only rare examples where a major tenant of practice is completely overturned in the absence of novel treatments,

Institute showed that only 16% of the large (and expensive) trials with substantive clinical outcomes demonstrated meaningful treatment effects.² Recent perioperative examples include major superiority trials of nitrous oxide, clonidine, aspirin, short red cell storage, steroids for cardiac surgery, regional analgesia for cancer recurrence, intensive care unit checklists and goal-setting, and levosimendan.

Robust trials showing comparable effects of various treatments are valuable, especially if one treatment is easier to implement, less toxic, or less expensive than the alternative. Still, it is disconcerting that so many large trials (*e.g.*, more than 1,000 patients) fail to demonstrate strong evidence for a difference in treatments when differences were expected based on preclinical or other data, especially since such trials are typically based on compelling mechanisms, strong animal

Image: Adobe Stock.

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are available in both the HTML and PDF versions of this article. Links to the digital files are provided in the HTML text of this article on the Journal's Web site (www.anesthesiology.org).

Accepted for publication January 2, 2020. Published online first on February 11, 2020. From the Department of Outcomes Research, Anesthesiology Institute, Cleveland Clinic, Cleveland, Ohio.

Copyright © 2020, the American Society of Anesthesiologists, Inc. All Rights Reserved. Anesthesiology 2020; 132:606–8. DOI: 10.1097/ALN.0000000000003162

such as routinely using perioperative β blockers to generally avoiding them, or moving from avoiding childhood peanut exposure to prevent allergies to encouraging exposure for the same reason.

From this perspective, observed clinical variability may not represent suboptimal care. Instead, it may identify non-critical practice areas where reasonable variability minimally influences outcomes—and may explain why so many comparative effectiveness trials show various treatments to be comparable.

Is Practice Already Personalized?

An even more intriguing possibility is that clinicians already practice some form of personalized medicine. Thus, specific treatments may be assigned in ways that appear arbitrary and variable but are actually appropriate for individuals.³ Trial results generally represent averages among enrolled patients. Response variations within trial populations are always considered, but even large trials are usually underpowered for subgroup analyses. Furthermore, it may not be possible to categorize some subtle but important factors that drive treatment variation. Finally, complex patients with multiple morbidities are often excluded from trials—but are common in routine practice.

Small trials are often tightly controlled, which reduces variability by restricting enrollment and limiting clinicians' options for altering management. In contrast, large trials usually have more pragmatic designs that allow more management flexibility—and may not even record potentially important aspects of baseline risk and management. But to some degree, all clinical trials poorly recognize individual variation in response to interventions. Instead, they report results for the average enrolled patient—but that are potential wrong for subsets of the population.

Even within tightly controlled (usually small) trials, many of the outcomes are generated by a small subset of high-risk patients. An even smaller subset of patients may contribute most outcomes in large pragmatic trials with less restrictive enrollment criteria. The average reported benefit and number-needed-to-treat may therefore actually apply to only a fraction of qualifying patients, a well-known trial limitation called treatment-effect heterogeneity.^{4,5}

For example, many studies report that intraoperative and postoperative hypotension is associated with myocardial injury. However, baseline risk is a far stronger determinant of perioperative myocardial injury than hypotension. (Hypotension remains interesting because, unlike baseline risk, it is modifiable.) Generally healthy young patients are at little risk of myocardial injury. Efforts to prevent moderate hypotension in such patients may actually worsen care because there are costs and risks of vasopressor and fluid administration. In contrast, limited randomized data suggest that hypotension prevention reduces major complications in high-risk patients.

The difficulty comes when the presumed benefits of hypotension avoidance—or any other treatment—are uncritically applied across broad populations. Even worse, some institutions may define quality by adherence to protocols, including enhanced recovery pathways, rather than by outcomes that actually matter to patients and healthcare systems. Confusing process with outcomes is dangerous because causal relationships between the two are often lacking, especially in individual patients. Hogan's "malignant hypercompliance" may therefore actually worsen care.⁶ Furthermore, the approach makes systems rigid and stifles innovation.

Clinicians intuitively recognize that trial results are driven by subsets of enrolled patients and may apply poorly to the entire study population, much less to other groups. Clinicians also understand that individual patients are complicated and that treatments for one condition may worsen another. A natural consequence is that thoughtful evaluation of multiple considerations will promote patients-specific treatments. Tailoring treatments to specific patients increases practice variability—but it is variability that may well improve care.

Clinical Pathways and Enhanced Recovery

Medicine is slowly shifting to value-based care, with value being defined as quality divided by cost. Everyone agrees that the quality of medical care should be high. But there is also an assumption in some quarters that practice variability is inherently suboptimal. For example, a fundamental precept of clinical pathways, including enhanced recovery, is that reducing variability will improve outcomes. There is some logic to the assertion, and reducing variability in clinical practices may often represent an improvement opportunity. For example, consistent practices and procedures presumably reduce misunderstandings and may speed patients' flow through a hospitalization. Undoubtedly, some aspects of various clinical pathways are evidence-based and clearly effective, including avoiding nasogastric tubes, early mobilization and feeding, and minimally invasive surgery.

Patients, though, are far more complicated than widgets being stamped out on an assembly line. There is little doubt that variability in industrial processes is harmful, but it is also true that the raw material is uniform and the process identical for each widget. The opposite is true in medicine: Much or most of the variability in patient care results from differences among patients, and the risk of complications is overwhelmingly driven by baseline patient risk rather than care patterns. Whether simply preventing clinical practice variability reduces major complications remains largely theoretical or based on scientifically weak observational analyses. It seems at least equally likely that much variation in clinical practice is harmless—or possibly even beneficial to the extent that clinicians recognize and respond appropriately to infinite variations in the human condition.

Of course, the supposition that practice variability is harmless or even beneficial hardly precludes comparative effectiveness trials, which remain important and much needed.⁷ When results from such trials are clear, they should be rapidly implemented; no one disputes that clinicians should practice evidence-based medicine when robust context-appropriate evidence exists. But, in the meantime, the clinical community should not assume that reducing practice variability *per se*, as required by many clinical pathways, will necessarily improve outcomes. Caution is especially warranted when pathway elements have been uncritically added, and when pathways are extrapolated to procedures and populations for which they were never tested.⁸ Restrictive pathways may even worsen outcomes in some patients. There is thus no basis for giving clinical pathways a “free pass” on evidence. Instead, they should be rigorously tested in context-sensitive robust and properly designed trials—just the way drugs and devices are.

Key citations are included in this version. A version with full citations is provided as Supplemental Digital Content (<http://links.lww.com/ALN/C234>).

Competing Interests

The author is not supported by, nor maintains any financial interest in, any commercial activity that may be associated with the topic of this article.

Correspondence

Address correspondence to Dr. Sessler: DS@OR.org

References

1. Onuoha OC, Arkoosh VA, Fleisher LA: Choosing wisely in anesthesiology: The gap between evidence and practice. *JAMA Intern Med* 2014; 174:1391–5
2. Gordon D, Taddei-Peters W, Mascette A, Antman M, Kaufmann PG, Lauer MS: Publication of trials funded by the National Heart, Lung, and Blood Institute. *N Engl J Med* 2013; 369:1926–34
3. McIsaac DI, Wijeyesundera DN, Bryson GL, Huang A, McCartney CJL, van Walraven C: Hospital-, anesthesiologist-, and patient-level variation in primary anesthesia type for hip fracture surgery: A population-based cross-sectional analysis. *ANESTHESIOLOGY* 2018; 129:1121–31
4. Kent DM, Alsheikh-Ali A, Hayward RA: Competing risk and heterogeneity of treatment effect in clinical trials. *Trials* 2008; 9:30
5. Kent DM, Rothwell PM, Ioannidis JP, Altman DG, Hayward RA: Assessing and reporting heterogeneity in treatment effects in clinical trials: A proposal. *Trials* 2010; 11:85
6. Hogan K: Malignant hypercompliance. *ANESTHESIOLOGY* 2017; 126:759–62
7. Devereaux PJ, Chan MT, Eisenach J, Schricker T, Sessler DI: The need for large clinical studies in perioperative medicine. *ANESTHESIOLOGY* 2012; 116:1169–75
8. Memtsoudis SG, Poeran J, Kehlet H: Enhanced recovery after surgery in the United States: From evidence-based practice to uncertain science? *JAMA* 2019; 321:1049–50