MAYO CLINIC

Why Randomized Controlled Trials Are Needed to Accept New Practices: 2 Medical Worldviews

Vinay Prasad, MD

From the Medical Oncology Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD.

Clinical practice is currently at a crossroads, and 2 worldviews are in opposition. The first view is that new medical practices should only be instituted when large, well-designed randomized controlled trials (RCTs) have a benefit on hard outcomes (mortality or morbidity), and current standards of care should be systematically subjected to such testing. The second view is that it is not only reasonable but also essential to institute and embrace practices when there is a reasonable belief that the practice is beneficial on the basis of the best available evidence and not necessarily RCTs. If at a future date a practice is found not to work, it is reasonable to discontinue it, but strict adherence to the first view would paralyze clinical care and limit progress.

The second view (reasonable belief) has dominated medicine for most of the 20th century and continues to exert a powerful influence well into this one. The first view (evidence first) has been promoted[1-3] and continues to make inroads but also faces strong opposition from those who hold the second view. In this article, I define the 2 views and explore the questions they raise: Do they have common ground? Can evidence guide us? And, most important, which view should you hold?

The 2 Views

Dueling arguments on the appropriateness of using CYP2C19 genotype results for decisions regarding the use of clopidogrel indicate the contrast between the sides. Representing evidence first, Nissen[4] writes,

> no matter how promising, pharmacogenetic approaches to treatment must withstand the same scrutiny required of all therapeutic advances—careful evaluation through well designed randomized clinical trials.

His remark is immediately met with criticism. Shuldiner et al,[5] representing reasonable belief, write,

> We contend that until the results of such studies are available, clinicians must evaluate the full body of evidence and use their best clinical judgment…[Nissen's position] may be denying patients access to potentially life-saving individualized alternative therapies.

Nissen responds, "I could not more strongly disagree…."[6]

Both parties would not dispute the facts. Although some analyses[7] have found that CYP2C19 genotype status predicts clinical outcomes for patients taking clopidogrel, others have not.[8] Thus, as a profession, we may legitimately be unsure of whether CYP2C19 loss of function alleles worsen outcomes among patients taking clopidogrel and whether these patients would be better served by a drug that does not require conversion to active metabolite (such as prasugrel). Nissen believes that uncertainty warrants skepticism and that routine testing should not begin until trials indicate it improves outcomes. Shuldiner et al believe that the same uncertainly should be met with tolerance and that we should test in certain cases.

This disagreement of one's medical worldview has become common in medical literature but is seldom directly discussed. Commenting on data that suggest routine use of pulmonary artery catheters does not improve outcomes in patients with shock, Hall[9] writes,

> the extrication from widespread use of a technology that has not been adequately assessed is difficult and painful, and should be made unnecessary by the prospective testing of all technology introduced into critical care in the future

(view 1: evidence first). Meanwhile, commenting on 2 sham control trials of vertebroplasty for osteoporotic fractures, which found no benefit of the common practice, Weinstein[10] writes, "when faced with several choices for which the evidence is less than clear, patients

Mayo Clin Proc. ■ October 2013;88(10):1046-1050 ■ http://dx.doi.org/10.1016/j.mayocp.2013.04.026
www.mayoclinicproceedings.org ■ Published by Elsevier Inc on behalf of Mayo Foundation for Medical Education and Research

and doctors must thoroughly review the options together" (view 2: reasonable belief).

A full-fledged embrace of reasonable belief was recently made in an article arguing that we should move ahead and treat certain patients with smoldering multiple myeloma, a precursor to myeloma, lacking end organ damage.[11] The authors argue that the evidence-first view "may delay progress by a decade or more" and, as such, is a "disservice to some high-risk patients."

In recent years, support for evidence first has increased. In an editorial on the increasing rate of radiation exposure due to medical imaging, Lauer[12] writes, "we need to adopt a new paradigm…were we to insist that all, or nearly all, procedures be studied in well-designed trials, we could answer many critical clinical questions within a short time." Dhruva and Redberg[13] concur. They argue that sex-specific data (indicating benefit in women) should be acquired and examined before insurers cover novel practices: "Data on therapeutic benefit should be gathered before Food and Drug Administration approval and Medicare coverage, as patterns of care become established quickly after approval."[13] Discussing the emergence of computer-aided detection mammography, Kerlikowske[14] writes, "Health care providers should not adopt new technologies without first demanding scientific evidence beyond that required for FDA [Food and Drug Administration] approval…[that] the benefits outweigh the harms." Regarding screening for coronary artery disease with electron-beam computed tomography, Chen and Krumholz[15] write,

> Evidence that EBCT [electron-beam computed tomography] is effective in reducing mortality or morbidity is arguably the most important barrier that this technology must clear before its routine use can be justified.

Problems With the Second View

Although reasonable belief is compelling and may strike many as the pragmatic position, there are deep problems with continuing to abide by it. One problem occurs in the inevitable situations in which our judgment regarding the best available evidence is later found to be in error. Previously, we have called such contradicted practices *medical reversals*.[2,3,16] A review of 35 trials in a high-impact medical journal, which tested standard of care, found that 16 (46%) were reversals.[16] Another study of 363 current medical practices found that 146 (40.2%) were medical reversals.[17] Elsewhere, my colleagues and I considered the costs for both human health and the health care system of medical reversal and concluded that the calculus strongly favors doing everything possible to minimize reversal (ie, view 1).[2]

The other major problem with reasonable belief is that what counts as reasonable, the best judgment of current knowledge, is inherently slippery, allowing the pervasive influence of conflicted parties (those with a financial interest in the results, such as industry or investigators with professional bias) to affect our practice. There is perhaps no better example of this than the dominance of the selective serotonin receptor inhibitors in the treatment of depression. Although the efficacy of selective serotonin receptor inhibitors is determined on the basis of hundreds if not thousands of RCTs,[18] nearly none were cleanly conducted, well-designed studies that assessed long-term end points in diverse and unfiltered groups and whose results might have reliably guided clinical decisions. Nevertheless, several of these drugs became billion dollar industries unto themselves.[18] In 2008, two meta-analyses suggested that for many patients the drugs had no benefit greater than placebo.[19,20] Poor study design, short follow-up, selective and distorted reporting, and building a supporting scientific myth have been implicated as factors that contributed to false enthusiasm.[18] The case of smoldering myeloma illustrates possible professional bias. Leaders in the translation science of myeloma are those who call for an embrace of reasonable belief, confident in their pathophysiologic understanding. It is not surprising that skeptics, not proponents, require RCTs to convince them that earlier treatment improves meaningful outcomes.

The last problem with reasonable belief is that it lacks consistency. What is reasonable on the basis of best available evidence to one physician might not be to another. Some may consider the use of inferior vena cava filters necessary, to be considered, or absolutely useless in patients with deep venous thrombosis, pulmonary embolism, or both (in patients receiving and

not receiving anticoagulation). Currently, there are no well-designed (double-blinded, hard outcome) studies to provide guidance. For the inferior vena cava filter, a lack of evidence is likely the reason why the 4 major sets of guidelines deviate so greatly from one another.[21,22]

In other cases, expert guidelines may provide consistency but may be skewed toward action because of financial conflicts of interest. Several sets of national guidelines have been faulted for this reason.[23,24] Lastly, if we relegate judgment to individual practitioners, practice may be biased by reimbursement (particularly among specialists whose salaries are tied to intervention). Under current payment structures, it may be truly impossible to practice view 2 impartially. Likely, this is the state of medicine.

The First View Will Not Prevent All Contradiction

It is worth noting that evidence first will not be able to altogether prevent erroneous practices from entering the mainstream. A key example is drotrecogin alfa (activated). In 2001 the Food and Drug Administration licensed drotrecogin alfa for the treatment of septic shock and severe sepsis on the basis of an RCT (PROWESS [Protein C Worldwide Evaluation in Severe Sepsis][25]), which found an improvement in all-cause mortality. A decade later, the manufacturer voluntarily recalled the drug after another large RCT[26] failed to substantiate its benefit. The case of drotrecogin alfa suggests that some reversal may be unavoidable. Randomized controlled trials are not infallible, yet they remain the strongest truth claim in the medical sciences.[27]

How to Evaluate the 2 Views

As a thought experiment, we might consider what it would be like to formally compare the 2 views to each other. Each potentially has strengths and weaknesses. Evidence first has the advantage of limiting the spread of ultimately mistaken medicine, and reasonable belief has the benefit of allowing for the early adoption of ultimately beneficial practices. Unfortunately, it would be difficult, if not impossible, to compare the 2 worldviews by RCT. Although we might imagine randomizing hospitals or insurance companies and following hard outcomes, such a study would be fraught with crossover as optimistic patients transfer to reasonable belief providers. In addition, the mere existence of reasonable belief medical systems would diminish the pressure on developers to perform studies to meet the standards of evidence first. A smaller market share that comes easily may be more enticing than twice that hard earned. For these reasons it is unlikely that we can ever truly know which view of medicine is better in the rigorous way that evidence first demands.

Will a High Bar Dissuade Innovation?

A perennial argument against higher up-front standards for approval is that it would dissuade innovation,[28] but there is a distinction worth making here. Innovation is only valuable if new products improve meaningful outcomes. Novel therapies that do not do so but merely favorably affect surrogates are deceptive and wasteful. If manufacturers are confident that their products are truly beneficial, a high bar does not change anything.

Perhaps the crux of the matter here is the related problem of market exclusivity, in which patent holders have sole production rights and may recoup initial outlays. Assessing meaningful end points requires longer study periods, which will cost time on industry patents. Additional market exclusivity, for products with meaningful gains, may be one way to remedy this.

Common Ground

Whether new practices should be required to show up-front benefit in well-designed, properly controlled RCTs that examine hard end points remains the central disagreement between the 2 views; however, when it comes to existing practices, the 2 views do not disagree. Both sides agree that we need to systematically appraise current clinical practice that is hitherto untested. Proponents of reasonable belief do not believe that we should remain ignorant, and proponents of evidence first do not believe that we should immediately abandon all unproven medicine. As noted by Berwick and Hackbarth,[29] with health care costs nearing $2.5 trillion, any substantial changes in spending (by strict adherence to view 1) may disrupt the entire US economy. Both groups agree that clinical practice cannot markedly change overnight, and at the same time, to answer important questions, we must prioritize trial enrollment. We might

Mayo Clin Proc. ■ October 2013;88(10):1046-1050 ■ http://dx.doi.org/10.1016/j.mayocp.2013.04.026
www.mayoclinicproceedings.org

find disagreement in the amount of funds we should invest in such ventures, the speed at which we should adjudicate uncertainty, and at what point we should suspend practices because data suggest their lack of benefit or presence of harms outweighing benefits. However, on the question of principle, they agree.

When RCTs Are Not Possible

Many contend that RCTs powered for hard end points are not feasible, ethical, or practical for certain clinical questions, but each year, these examples have dwindled. Rare diseases were long considered untestable by RCT, but now robust multicenter trial data exist for conditions as rare as 0.7 to 2 per million persons per year.[30] One author[31] relied on appendectomy for acute appendicitis as an example of an intervention that would never be tested in an RCT.[31] Now, 2 RCTs[32,33] suggest that the surgery may not be required for many patients with this condition. Others contend that biomarkers will soon be reliable enough to serve as an infallible surrogate, perfectly predicting treatment response and long-term outcomes.[34] Although it is not impossible to imagine such a future, these biomarkers will first have to be validated by traditional empirical studies under a variety of conditions.[35]

Conclusion

When it comes to one's medical worldview, there is no neutral position. Every clinician makes the choice daily. For most of us, our decisions are capricious, uneven, and arbitrary. For interventions we spurn, we proudly assert that "there is no evidence." For others that we favor, we stress that "there are no negative studies." For other interventions with promising rationale and negative empirical trials, we argue that null data are flawed.[36] Philosophically, this isn't sure footing.

I propose that the discussion of one's medical worldview must be made explicit. Students should be taught about the 2 views. Researchers should formally disclose their camp. I favor evidence first because reasonable belief can easily be hijacked by special interests; because primum non nocere has meant many things over millennia, but in the modern world it can only mean minimizing reversal[37]; and because, in an age in which health care costs will soon reach 25% of the our nation's gross domestic product,[38] only evidence first may save our economy and our profession.

ACKNOWLEDGMENTS

The views and opinions of Dr Prasad do not necessarily reflect those of the National Cancer Institute or National Institutes of Health.

Correspondence: Address to Vinay Prasad, MD, Medical Oncology Branch, National Cancer Institute, National Institutes of Health, 10 Center Dr, 10/12N226, Bethesda, MD 20892 (vinayak.prasad@nih.gov).

REFERENCES

1. Redberg RF, Walsh J. Pay now, benefits may follow—the case of cardiac computed tomographic angiography. *N Engl J Med*. 2008;359(22):2309-2311.
2. Prasad V, Cifu A, Ioannidis JP. Reversals of established medical practices: evidence to abandon ship. *JAMA*. 2012;307(1):37-38.
3. Prasad V, Cifu A. Medical reversal: why we must raise the bar before adopting new technologies. *Yale J Biol Med*. 2011;84(4):471-478.
4. Nissen SE. Pharmacogenomics and clopidogrel: irrational exuberance? *JAMA*. 2011;306(24):2727-2728.
5. Shuldiner AR, Vesely MR, Fisch A. CYP2C19 genotype and cardiovascular events [letter]. *JAMA*. 2012;307(14):1482; author reply 1485.
6. Nissen SE. CYP2C19 genotype and cardiovascular events—reply. *JAMA*. 2012;307(14):1485.
7. Mega JL, Simon T, Collet JP, et al. Reduced-function CYP2C19 genotype and risk of adverse clinical outcomes among patients treated with clopidogrel predominantly for PCI: a meta-analysis. *JAMA*. 2010;304(16):1821-1830.
8. Holmes MV, Perel P, Shah T, Hingorani AD, Casas JP. CYP2C19 genotype, clopidogrel metabolism, platelet function, and cardiovascular events. *JAMA*. 2011;306(24):2704-2714.
9. Hall JB. Use of the pulmonary artery catheter in critically ill patients. *JAMA*. 2000;283(19):2577-2578.
10. Weinstein JN. Balancing science and informed choice in decisions about vertebroplasty. *N Engl J Med*. 2009;361(6):619-621.
11. Rajkumar SV, Merlini G, San Miguel JF. Haematological cancer: redefining myeloma. *Nat Rev Clin Oncol*. 2012;9(9):494-496.
12. Lauer MS. Elements of danger—the case of medical imaging. *Minn Med*. 2009;92(12):40-41.
13. Dhruva SS, Redberg RF. Clinical trial enrollment and progress in women's health [letter] . *JAMA*. 2011;305(12):1197; author reply 1197-1198.
14. Kerlikowske K. A call for evidence of benefits outweighing harms before implementing new technologies: comment on "Diffusion of Computer-Aided Mammography After Mandated Medicare Coverage." *Arch Intern Med*. 2010; 170(11):990-991.
15. Chen J, Krumholz HM. How useful is computed tomography for screening for coronary artery disease? screening for coronary artery disease with electron-beam computed tomography is not useful. *Circulation*. 2006;113(1):125-146.
16. Prasad V, Gall V, Cifu A. The frequency of medical reversal. *Arch Intern Med*. 2011;171(18):1675-1676.
17. Prasad V, Vandross A, Toomey C, et al. A decade of reversal: an analysis of 146 contradicted medical practices. *Mayo Clin Proc*. 2013;88(8):790-798.
18. Ioannidis JP. Effectiveness of antidepressants: an evidence myth constructed from a thousand randomized trials? philosophy, ethics, and humanities in medicine: *Philos Ethics Humanit Med*. 2008;3:14.

19. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. 2008; 358(3):252-260.
20. Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med*. 2008;5(2):e45.
21. Minichiello T. Efforts to optimize patient benefit from inferior vena cava filters: comment on "Retrieval of Inferior Vena Caval Filters After Prolonged Indwelling Time." *Arch Intern Med*. 2011; 171(21):1955-1956.
22. Prasad V, Rho J, Cifu A. The inferior vena cava filter: how could a medical device be so well accepted without any evidence of efficacy? *JAMA Intern Med*. 2013;173(7):493-495.
23. Mendelson TB, Meltzer M, Campbell EG, Caplan AL, Kirkpatrick JN. Conflicts of interest in cardiovascular clinical practice guidelines. *Arch Intern Med*. 2011;171(6):577-584.
24. Neuman J, Korenstein D, Ross JS, Keyhani S. Prevalence of financial conflicts of interest among panel members producing clinical practice guidelines in Canada and United States: cross sectional study. *BMJ*. 2011;343:d5621.
25. Bernard GR, Vincent JL, Laterre PF, et al. Efficacy and safety of recombinant human activated protein C for severe sepsis. *N Engl J Med*. 2001;344(10):699-709.
26. Ranieri VM, Thompson BT, Barie PS, et al. Drotrecogin alfa (activated) in adults with septic shock. *N Engl J Med*. 2012; 366(22):2055-2064.
27. Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124.
28. Wood AJ. A proposal for radical changes in the drug-approval process. *N Engl J Med*. 2006;355(6):618-623.
29. Berwick DM, Hackbarth AD. Eliminating waste in US health care. *JAMA*. 2012;307(14):1513-1516.
30. Fassnacht M, Terzolo M, Allolio B, et al. Combination chemotherapy in advanced adrenocortical carcinoma. *N Engl J Med*. 2012;366(23):2189-2197.
31. Worrall J. Why there's no cause to randomize. *Br J Philos Sci*. 2007;58:451-488.
32. Vons C, Barry C, Maitre S, et al. Amoxicillin plus clavulanic acid versus appendicectomy for treatment of acute uncomplicated appendicitis: an open-label, non-inferiority, randomised controlled trial. *Lancet*. 2011;377(9777):1573-1579.
33. Styrud J, Eriksson S, Nilsson I, et al. Appendectomy versus antibiotic treatment in acute appendicitis: a prospective multicenter randomized controlled trial. *World J Surg*. 2006;30(6):1033-1037.
34. Kolata, G. In treatment for leukemia, glimpses of the future. NY Times. http://www.nytimes.com/2012/07/08/health/in-gene-sequencing-treatment-for-leukemia-glimpses-of-the-future.html. Accessed July 18, 2013.
35. Prasad V, Bonow RO. The cardiovascular biomarker conundrum. *JAMA*. 2011;306(19):2151-2152.
36. Siontis GC, Tatsioni A, Katritsis DG, Ioannidis JP. Persistent reservations against contradicted percutaneous coronary intervention indications: citation content analysis. *Am Heart J*. 2009;157(4):695-701.
37. Prasad V, Cifu A. A medical burden of proof: towards a new ethic. *Biosocieties*. 2012;7:72-87.
38. Congressional Budget Office. The 2012 long-term budget outlook. http://cbo.gov/publication/43288. Accessed June 20, 2012.