

Are Medical Guidelines Better Than Flipping a Coin?

A recent article by Prasad et al. (1) in the Mayo Clinic Proceedings reviewed all original articles published over 10 years (2001-2010) in the New England Journal of Medicine (NEJM). Articles were classified on the basis of whether they addressed a medical practice, whether they tested a new or existing therapy, and whether results were positive or negative. Most striking was that of the 363 articles examining standards of care, 146 (40.2%) reversed that practice, whereas 138 (38.0%) reaffirmed it. The remaining percentage remained inconclusive.

As pointed out in an accompanying editorial, the NEJM is widely read, has high visibility and has a large influence on the mass media and medical practitioners (2). However, the effect of articles published in the NEJM, Lancet and JAMA, the top 3 general medical journals in terms of impact factor, are markedly inflated (3,4). Presumably, a randomized trial published in these journals must be true because these are the “best” medical journals.

Prasad's conclusions that the NEJM reversed accepted medical practice about half the time would be consistent with the Cochrane Review of Clinical Trials. El Dib et al. (5) concluded in 2004 that there is insufficient evidence to endorse the examined interventions 47.8% of the time. A repeat evaluation in 2011 showed that the percentage of insufficient evidence remained about the same (6).

Now before anyone gets too upset, I happen to agree that NEJM, Lancet and JAMA are probably the best and most influential medical journals. Authors send their best work to these journals because they are widely read. The editors choose articles based on their importance and whether the work is new, innovative, or contradicts accepted medical practice. All of this makes these journals the most influential.

Not surprisingly, authors of guidelines give more credibility to these higher impact journals. In other words, a randomized trial done in the NEJM is more likely to influence a guideline writing committee than a trial from the Southwest Journal of Pulmonary and Critical Care. Looking at Bob Raschke's recent journal club reviewing 6 landmark randomized controlled trials that were eventually reversed, 5 were from the NEJM or JAMA (7). Several of the outcomes from these studies were the basis for guidelines.

Guideline writing committees really cannot do better than the medical literature. However, if half the established standards of care are wrong as Prasad suggests, half the guidelines based on these standards of care are also wrong. Should we require higher levels of evidence before practice guidelines are recommended—perhaps at least two, or in cases of marginal effects, even more trials. To me the overwhelming answer has to be yes.

Lee and Vielemeyer (8) found that only 14% of the Infectious Disease Society of America (IDSA) guidelines are based on level I evidence (data from >1 properly randomized controlled trial). Much of this 14% and the 86% that are below level I evidence will eventually be proven wrong. I doubt that other medical societies are performing much better. Serving on a guideline writing committee is a compliment paid by professional colleagues. However, as Lee and Vielemeyer point out, the guidelines tend to be more opinion than science. This is especially true when the data supporting standards of care is weak, nonexistent or conflicting. Experts often rationalize that an answer is needed, even when the correct response might be “I don’t know”.

All this points out that reading and interpreting medical literature is difficult. It takes knowledge, experience and a healthy dose of skepticism. Experts relying on the best evidence frequently get it wrong. Improvement lies in the intellectual honesty of the guidelines committees and research. Well designed clinical trials are usually expensive and time-consuming, not what health care administrators want to hear in a time of restricted budgets. However, can we afford not to invest in getting it right?

Richard A. Robbins, MD*

References

1. 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):790-8. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Ioannidis JP. How many contemporary medical practices are worse than doing nothing or doing less? *Mayo Clin Proc.* 2013;88(8):779-81. [\[CrossRef\]](#) [\[PubMed\]](#)
3. ISI Web of Science. Journal Citation Reports. Available at: <http://thomsonreuters.com/journal-citation-reports>. Accessed August 7, 2013.
4. Siontis KC, Evangelou E, Ioannidis JP. Magnitude of effects in clinical trials published in high-impact general medical journals. *Int J Epidemiol.* 2011;40(5):1280-91. [\[CrossRef\]](#) [\[PubMed\]](#)
5. El Dib RP, Atallah AN, Andriolo RB. Mapping the Cochrane evidence for decision making in health care. *J Eval Clin Pract.* 2007;13(4):689-692. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Villas Boas PJ, Spagnuolo RS, Kamegasawa A, et al. Systematic reviews showed insufficient evidence for clinical practice in 2004: what about in 2011? The next appeal for the evidence-based medicine age. *J Eval Clin Pract.* 2013;19(4):633-7. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Raschke RA. August 2013 critical care journal club: less is more. *Southwest J Pulm Crit Care.* 2013;7(3):162-4. [\[CrossRef\]](#)
8. 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. [\[CrossRef\]](#) [\[PubMed\]](#)

*The opinions expressed are those of the author and do not necessarily reflect the opinion or policies of the Arizona, New Mexico, or Colorado Thoracic Societies, the Mayo Clinic, or most guideline writing committees.