

Want More Trust in Medical Science? Embrace Uncertainty and Cut the Hype

John Mandrola, MD

April 17, 2017

Many of the same medical scientists who garner fame doing clinical trials also write practice guidelines.

Guidelines are too often infected with hubris. The sordid story of [dronedaron's](#) (*Multaq*, Sanofi) rise in atrial-fibrillation treatment guidelines exposes serious reliability issues with the important documents. In a study published in *JAMA Internal Medicine*, Italian researchers applied the Grading of Recommendations Assessment, Development and Evaluation (GRADE) method to the evidence base of dronedarone.^[11] This assessment of the evidence revealed excess deaths, adverse effects, and inefficacy of arrhythmia suppression with dronedarone. The Italian investigators also found that the current AF treatment guidelines were lacking in five of eight quality domains set out by the Institute of Medicine.

Another group of academics that threaten the public trust are screening evangelists. Screening is precarious because it puts doctors close to breaking the golden rule—*first, do no harm*. Doing things to people-without-complaints and promoting the slogan "screening saves lives" should require clearing the highest bar of evidence. The truth, though, is that the evidence does *not* support such zealous advocacy.

A systematic review of meta-analyses and randomized clinical trials that studied screening of asymptomatic adults for 19 diseases (39 tests, including mammography) found reductions in disease-specific mortality were uncommon and reductions in all-cause mortality were very rare or nonexistent.^[12] Prasad and colleagues explain how screening advocates conflate disease-specific death rates with overall mortality. "Using disease-specific mortality as a proxy for overall mortality," they wrote in the *BMJ*, "deprives people of information about their chief concern: reducing the risk of dying."^[13]

Through the Twitter feed of Dr Benjamin Mazer, I found this explanatory [quote](#): "If you are a patient contemplating some screening test, and the result of that test or treatment is no measurable reduction in the rate of death at some clinically relevant later point in time, then why have the test or treatment—unless the patient, for some reason, has a desire to die from condition A instead of condition B."

The late Dr David Sackett, the father of evidence-based medicine, famously wrote that preventive medicine displayed all three elements of arrogance: it was aggressively assertive, presumptuous, and overbearing.^[14] It's hardly controversial to posit less arrogance would improve confidence in medical science.

Can Journal Editors Help?

Finally, I think journal editors deserve mention, mainly because they have (some) power to make positive change. Three areas in need of improvement in medical journals are attention to the limitations of studies, peer review, and consistent reporting of results.

I would propose that authors have to list in the abstract—before the conclusions—a paragraph on limitations. Too many readers of the scientific literature never get to the full paper; as such, they don't see a study's limitations. Press releases from journals should also include the limitations of a study.

Peer review is another area in need of improvement. A Google search of "peer review" churns out a litany of articles describing how it is broken. Among the many problems, anonymous peer reviewers now have no incentive, other than goodwill, or malice, to do a good job. As a nonacademic, I don't claim to have *the* answer to fix peer review, but it's hard to see how a more transparent process could not help. Why not publish the peer review (and give academic credit) along with a paper?

Another easy improvement editors could insist on is a consistent presentation of results. It crushes trust when reading a study that expresses benefits in relative terms and harms in absolute terms. Clinicians caring for patients deserve to read about both benefits and harms of therapies in absolute—not relative—terms.

Worship of the P value is also a problem. Pharmacologist and statistician David Colquhoun warns that "if you use $P=0.05$ to suggest that you have made a discovery, you will be wrong at least 30% of the time." He adds: "If, as is often the case, experiments are underpowered, you will be wrong most of the time."^[5] This year, Nobel laureate Daniel Kahneman (author of *Thinking Fast and Slow*) admitted placing too much faith in underpowered studies.^[6]

Then there is the vast divide between *statistically* significant and *clinically* significant. The most recent blockbuster, the proprotein convertase subtilisin/kexin type 9 (PCSK9) inhibitor [evolocumab](#) (*Repatha*, Amgen), highlights this issue.^[7] If it takes over 27,500 patients for this lipid-lowering drug to show a 1.5% reduction in nonfatal events, and no mortality benefit, is this clinically important?

Publication bias is another *trust-in-the-evidence* problem I learned of later in my career. In a highly cited study published in the *New England Journal of Medicine*, Dr Erick Turner and colleagues showed that positive drug trials (of an antidepressant medication) were much more likely to be published. Of 74 FDA-registered studies, they found 37/38 positive trials were published while negative or questionable results were, with three exceptions, either not published (22 studies) or published in a way that conveyed a positive outcome (11 studies).^[7]

These findings are not an outlier. Ross and colleagues analyzed more than 600 NIH-funded trials and reported that a third of trials remained unpublished more than 4 years after completion.^[8] Not surprising is that nonpublication of trials is more common among those with industry

funding.^[9] Selective publication deserves attention because not knowing about unpublished negative studies leads to overconfidence in our therapies.

The Trust Problem With Overselling Medical Science

Not all of the trust problem with medical science is due to overconfidence on the part of clinicians. Bias and hype from medical researchers are also a problem.

Science does not do itself. Humans—bent on having a successful academic career—do science. This means positive results can become the goal rather than the pursuit of scientific truth.

Statistician Thomas Fleming noted that the "psychological influence of the driving goal to establish benefit is sufficiently subtle that it often is not recognized by investigators who have this goal." He pointed to a review of trial protocols in which bias came through in the language of the written objectives.^[10] Rather than write that a study's objectives are "to *determine whether* the experimental intervention is safe and effective," many authors, he observed, write objectives in this way: "to *establish* the experimental intervention is safe and effective."

Conclusion

My view of medical science is similar to my view of patient care. Slow and conservative is better. Don't oversell; share the data more than the interpretation of the data. Consumers of medical science, patients and clinicians alike, will make better decisions with more data and less hype. We can handle the truth: science is supposed to move slowly.

- Hoffmann TC, Del Mar C. Clinicians' expectations of the benefits and harms of treatments, screening, and tests: A systematic review. *JAMA Intern Med.* 2017;177:407-419. [Abstract](#)
- Ioannidis JP. Why most published research findings are false. *PLoS Med.* 2005;2:e124. [Article](#)
- Ioannidis JA. Contradicted and initially stronger effects in highly cited clinical research. *JAMA.* 2005;294:218-228. [Article](#)
- Horton R. Offline: What is medicine's 5 sigma? *Lancet.* 2015;385:1380. [Article](#)
- Colquhoun D. An investigation of the false discovery rate and the misinterpretation of p-values. *R Soc Open Sci.* 2014;19;140216. [Article](#)
- Schimmack U, Heene M, and Kesavan K. Reconstruction of a train wreck: How priming research went off the rails. *Replicability Index*, February 2, 2017. [Article](#)

- Turner EH, Matthews AM, Linardatos E, et al. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. 2008;358:252-260. [Article](#)
- Ross JS, Tse T, Zarin DA, et al. Publication of NIH funded trials registered in ClinicalTrials.gov: Cross sectional analysis. *BMJ*. 2012;344:d7292. [Article](#)
- Jones CW, Handler L, Crowell KE, et al. Non-publication of large randomized clinical trials: Cross sectional analysis. *BMJ*. 2013;347: f6104. [Article](#)
- Fleming TR. Clinical trials: Discerning hype from substance. *Ann Intern Med*. 2010;153:400-406. [Article](#)
- Iannone P, Haupt E, Flego G, et al. Dronedarone for atrial fibrillation: The limited reliability of clinical practice guidelines. *JAMA Intern Med*. 2014;174:625-629. [Article](#)
- Saquib N, Saquib J, Ioannidis JP. Does screening for disease save lives in asymptomatic adults? Systematic review of meta-analyses and randomized trials. *Int J Epidemiol*. 2015;44:264-277. [Article](#)
- Prasad V, Lenzer J, Newman DH. Why cancer screening has never been shown to "save lives"—and what we can do about it. *BMJ*. 2016;352:h6080. [Article](#)
- Sackett DL. The arrogance of preventive medicine. *CMAJ*. 2002;167:363-364. [Article](#)