

Editorial

Editorial: When “Safe and Effective” Becomes Dangerous

Seth S. Leopold MD

Commercials might claim a product is “safe and effective,” but most research studies should not. Small, focused studies may deem one treatment more effective than another, but the problem arises when authors of such studies claim safety based on the observation that few (or no) patients were hurt by some intervention. Such claims may be misleading.

It is almost impossible to evaluate safety and efficacy in the same study.

There are many reasons for this, but the key ones are (1) the elements of study design one needs to evaluate efficacy are different from those to evaluate safety, and (2) demonstrating safety requires evaluation of many more patients than does demonstrating efficacy.

There are several important study-design differences between safety and efficacy studies. Whereas efficacy studies can focus on specific endpoints like “Does ligament reconstruction decrease the likelihood of subsequent inversion injury to the tibiotalar joint?”, safety studies must be open to the possibility that many different kinds of adverse effects might occur, not all of which will be immediately evident at the time of treatment or even at the conclusion of an efficacy study. For example, the discovery of systemic effects of local procedures like THA [10], in particular THAs with metal-on-metal bearings [6], drug interactions or unexpected complications from pharmacologic treatments [12], and unanticipated modes of failure [3] all have changed our views about potentially promising treatments, in some cases, even after shorter-term efficacy trials have immodestly claimed safety. This last point is important: While efficacy can be demonstrated quickly,

we often do not learn about the harms our interventions cause until much later.

Because serious complications of our treatments are generally and thankfully uncommon, safety studies must be much larger in order to have a fair likelihood of detecting them. For example, if we examine the studies supporting the multimodal analgesia approaches now in common use after orthopaedic surgery, we note they tend to have two things in common: They apply at least several classes of medications to a population, and they almost always are small [4, 11]. Because the study populations often include older patients, many of whom also take other medications, we take a risk when we infer safety from small studies designed to evaluate efficacy, and apply a complex protocol to complex population on a large scale. One study [11] on the efficacy of a particular multimodal analgesia approach, which also claimed it “confirmed the safety” of its protocol, prescribed at least five drug classes (including two different NSAIDs), and involved a cocktail containing drugs from three classes, which was injected into six different kinds of tissue around the knee. This study was powered to detect a clinically important difference in patients’ pain, and with a total of only 42 patients, was able to detect such a

Editor-in-Chief, *Clinical Orthopaedics and Related Research*®

The author certifies that he, or any members of his immediate family, has no commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

The opinions expressed are those of the writers, and do not reflect the opinion or policy of *CORR*® or the Association of Bone and Joint Surgeons®.

S. S. Leopold MD (✉)
Clinical Orthopaedics and Related
Research, Philadelphia, PA 19103, USA
e-mail: sleopold@clinOrthop.org

Editorial

difference. But with fewer than four dozen patients, efficacy studies of this sort should make no claims about safety; contrast this with a key meta-analysis that concluded the NSAID rofecoxib (Vioxx) was unsafe — that study required data from over 20,000 patients [2] in order to draw definitive conclusions, and it still was controversial [5].

Registries offer another potential window into the safety of some of the tools we use; they do this by accessing data from large populations of patients [1]. The postmarketing surveillance required by the FDA includes a database consisting of hundreds of thousands of new reports of confirmed or possible device-associated serious illnesses, deaths, and malfunctions drawn from the experiences of millions of patients every year [8]. While the FDA has definitions for the kinds of evidence required to declare a device to be “safe” [9], we also now know that many devices thus vetted turn out not to be safe at all [7], placing the burden back on us, as clinicians, to know the difference between a study that demonstrates safety and one that demonstrates efficacy.

It’s important to remember, though, that studies whose size, scope, and duration genuinely permit answering safety questions generally do so at the expense of patient-level detail about efficacy. To evaluate hip scores after

femoroacetabular impingement surgery, the likelihood of return to sport after shoulder arthroscopy, or range of motion after basilar joint arthroplasty of the thumb, smaller trials often suffice, and may allow for a more granular examination of the dataset. Questions like those often can be answered by studies enrolling anywhere between a few dozen and a couple hundred patients.

But if small studies of efficacy fail to identify any patients who were harmed by the intervention, one should not conclude that those interventions are safe. Safety and efficacy both are important, but evaluating each requires a different kind of study. Beware of studies that, like commercials, claim a treatment to be both “safe and effective.”

Acknowledgments I would like to thank Lee Beadling BS, Paul A. Manner MD, Clare M. Rimnac PhD, and Montri D. Wongworawat MD for their reviews of and thoughtful suggestions about this essay, which vastly improved it.

References

- Huang DC, Tatman P, Mehle S, Gioe TJ. Cumulative revision rate is higher in metal-on-metal THA than metal-on-polyethylene THA: Analysis of survival in a community registry. *Clin Orthop Relat Res.* 2013;471:1920–1925.
- Jüni P, Nartey L, Reichenbach S, Sterchi R, Dieppe PA, Egger M. Risk of Cardiovascular events and Rofecoxib: Cumulative meta-analysis. *Lancet.* 2004;364:2021–2029.
- Lee DH, Ryu KJ, Song HR, Han SH. Complications of the intramedullary skeletal kinetic distractor (ISKD) in distraction osteogenesis. [Published online ahead of print March 7, 2014]. *Clin Orthop Relat Res.* DOI: 10.1007/s11999-014-3547-4.
- Lee KJ, Min BW, Bae KC, Cho CH, Kwon DH. Efficacy of multimodal pain control protocol in the setting of total hip arthroplasty. *Clinics in Orthopedic Surgery.* 2009;1:155–160.
- Nyberg J. Discontinuation of Vioxx. *Lancet.* 2005;365:24–25.
- Prentice JR, Clark MJ, Hoggard N, Morton AC, Tooth C, Paley MN, Stockley I, Hadjivassiliou M, Wilkinson JM. Metal-on-metal hip prostheses and systemic health: A cross-sectional study 8 years after implantation. *PLoS One.* 2013;8:e66186.
- Reito A, Puolakka T, Elo P, Pajamäki J, Eskelinen A. High prevalence of adverse reactions to metal debris in small-headed ASR™ hips. *Clin Orthop Relat Res.* 2013;471:2954–2961.
- United States Food and Drug Administration. Available at: <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm>. Accessed April 25, 2014.
- United States Food and Drug Administration. Available at: <http://www.accessdata.fda.gov/scripts/cdrh/>

Editorial

- [cfdocs/cfCFR/CFRSearch.cfm?FR=860.7](#). Accessed April 28, 2014.
10. Urban RM, Jacobs JJ, Tomlinson MJ, Gavrilovic J, Black J, Peoc'h M. Dissemination of wear particles to the liver, spleen, and abdominal lymph nodes of patients with hip or knee replacement. *J Bone Joint Surg.* 2000;82-A:457–476.
 11. Vendittoli PA, Makinen P, Drolet P, Lavigne M, Fallaha M, Guertin MC, Varin F. A multimodal analgesia protocol for total knee arthroplasty: A randomized, controlled study. *J Bone Joint Surg.* 2006;88-A:282–289.
 12. Ward WG Sr, Carter CJ, Wilson SC, Emory CL. Femoral stress fractures associated with long-term bisphosphonate treatment. *Clin Orthop Relat Res.* 2012;470:759–765.