



When should we adopt new technology into our practices?

Seth S. Leopold^{1,2}

Accepted: 19 July 2021 / Published online: 17 August 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

What makes orthopaedic surgery so satisfying to practice?

In my highly unscientific survey on this topic, the most common answer I hear is: “My patients usually get better.” The second is: “I enjoy using tools and new technology to make things as perfect as I can for my patients.”

We need to make sure that we don’t mistakenly infer that the second answer causes the first. As importantly, we need to be mindful not to assume that the reason some patients don’t get better is insufficiently advanced surgical technology.

Most people who undergo arthroplasty, for example, do well with surgery. Still, one in five patients who have total knee replacement are not satisfied with the result [1]. As the editor of a large, general-interest orthopaedic journal, I’ve lost count of the number of papers that begin with that unhappy 20%, and use it to justify the exploration of a new implant, navigation system, kinematic alignment approach, surgical robot, or other expensive, unproven tool.

In general, I think that effort is misdirected. As far as we now know, differences in outcomes scores among generally well-performing implants are negligible or nonexistent [2], the odds of a new implant lasting longer than an existing one is hardly better than a coin toss [3, 4], and no well-designed study about a novel implant-alignment tool, ligament-balancing approach, or technology-driven innovation has made a dent in patient-reported outcomes or implant durability. The best such studies—systematic reviews, network meta-analyses, registry reports, and long-term follow-up studies of randomized trials—have found no differences at all that a patient might perceive [5–9].

I believe the main causes of patient dissatisfaction and persistent pain after major elective orthopaedic surgery are

much simpler. For example, the proportion of patients in the United States with depression or anxiety is in the ballpark of 20% [10]; it’s pretty similar in Europe [11]. Incomplete management of depression and other manifestations of emotional distress (like anxiety disorders), as well as performing elective surgery on patients who are habituated to higher-dose narcotic analgesics—another known risk factor for persistent pain after full recovery—probably go a long way towards explaining why so many patients are not satisfied with their surgical results. They certainly make more sense to me than the fraction-of-a-degree improvements one might hope to get from a navigation system or a robot.

Why, then, do bright surgeons (and good journals) sometimes take the bait, and believe that these new tools are worth using?

Again, the answer is decidedly low-tech, if not downright unsexy: Human nature [11] as well as the common kinds of biases that cause us to overestimate our effectiveness in other areas [12] typically beset research about our newest tools. These include selection bias, transfer bias, and assessment bias, as well as the conflation of statistical significance with clinical importance (Table 1). All three of those kinds of bias are present, at least to some degree, in most observational orthopaedic research. And, importantly, they don’t offset one another. They work together to inflate the apparent benefits of new treatments and tools. Editors—people like me—need to help authors do a better job protecting readers from the misunderstandings and misinterpretations that these biases can cause.

Until we address those sources of bias, which is best done using in the context of randomized controlled trials (RCTs) of adequate follow-up duration, and until we refocus readers’ attention on clinical importance (rather than mere statistical significance), we are likely to view new approaches more favorably than we should. It’s easy to be fooled, or to fool oneself. I know that I have.

Some years ago, I published a comparative study of less-invasive TKA [13]; I chose patients and controls immediately on either side of the changeover date from the conventional approach to the new one, so there shouldn’t have

✉ Seth S. Leopold
sleopold@clinorthop.org

¹ Clinical Orthopaedics and Related Research, Philadelphia, Pennsylvania, USA

² Department of Orthopaedics and Sports Medicine, University of Washington School of Medicine, Seattle, WA, USA

Table 1 Common types of bias that affect apparent effect sizes in observational research

Type of bias	Examples	Influence on Study's Conclusions
Selection bias	<ul style="list-style-type: none"> • Surgeon-developers and surgeons using tools and devices for the first time will tend to pick “easier” procedures in which to try them out • Treatment and control groups differ at baseline with respect to factors that could influence the results of the treatments being compared, often with the novel treatment being used in less-complicated situations or patients with fewer comorbid conditions 	Inflates the apparent benefit of the new treatment
Transfer bias	<ul style="list-style-type: none"> • Follow-up insufficiently long to detect all relevant potential harms of the new treatment • Follow-up is insufficiently complete (loss to follow-up) • Differential loss to follow-up between treatment and control groups (the missing usually are not doing as well as the accounted for) 	Generally underestimates the harms of the new treatment, and may overstate its benefits
Assessment/Assessor bias	<ul style="list-style-type: none"> • The study's endpoints (such as implant alignment) may be inadequate surrogates for the endpoints of greatest clinical interest (such as pain, function, or premature loosening) • Use of non-validated outcomes tools • Assessment performed by individuals involved in care of the patients, rather than by independent study assessors 	Overstates effectiveness of the new treatment; assuming effectiveness based on surrogate endpoints can be misleading
Conflating statistical significance with clinical importance	<ul style="list-style-type: none"> • Reporting p values without effect sizes (patients and surgeons perceive effect sizes, not p values) • Focusing on small (“significant”) p values despite effects that are so small as to be clinically unimportant; large samples may produce small p values despite inconsequential effect sizes • Failing to define or report results in the context of the minimum clinically important differences 	Results in unsupported conclusions of clinical efficacy

been much of a temporal effect. If anything, I thought, the learning curve of the new approach should have dampened its apparent benefits in the study; after all, I was much more experienced with the old methods of exposing a knee and cutting the bones. It seemed to me that this was the fairest comparison short of an RCT. Despite that, less-invasive technique still looked better.

This resulted in some local notoriety and even a call from an editor at the *New England Journal of Medicine*, who asked me to write it up for their audience [14]. Perhaps not surprisingly, this caused the phone to ring in the office, along with a resultant boost to my practice. But the advantages of the new approach we observed may not have been—in fact, probably were not—caused by the approach at all. In parallel with the change to the less-invasive surgical approach, a number of improvements in analgesia, anesthesia, and therapy occurred. The new surgical approach probably represented a modification of the old midvastus approach, which is somewhat more patella-friendly than the medial parapatellar approach we used in the control group. I discounted the RCTs of the day that disagreed with my own work as not being well done, or as having been done by people using subtly different (and to my arrogant eye, less-effective) approaches from my own. I was wrong to have done that.

Though I felt our study used the best approach short of an RCT, it was in fact not an RCT, and it caused me to overstate the efficacy of the new intervention. In time, partners using plain-vanilla approaches—but better analgesia, anesthesia, and therapy—achieved the same results I had. I had misled myself, and perhaps others. I suspect I'm not the only clinician-scientist to have fallen into this trap.

New technology almost always adds costs. It often adds time. It always carries the potential for unintended consequences (sometimes called “revenge effects”), which sometimes are surprising and severe [15, 16]. There usually is an associated learning curve, which often is associated with real harms to real people. Given this, I'm surprised by how often studies find small differences favoring new treatments (or even find no differences), yet still recommend the novel approach. I suppose it is natural to try to see the good in something, but here, I fear that impulse may be misguided. Skepticism may be the better posture, unless the approach being replaced is genuinely problematic or unreliable. When we're talking about expensive, potentially risky, or time- and resource-consuming interventions—and all surgical tools fall into one or several of those categories—they need to prove their value in definitive, explicit ways that our patients can perceive. Absent that, studies should recommend against their use, and surgeons should not use them.

Practicing according to those principles has not always resulted in me choosing the best-possible implants and tools, but it's almost always caused me to choose good ones, and

it's helped me to steer clear from innumerable disasters (and product recalls). For example, I was a late adopter of highly crosslinked polyethylene, and the difference of a few years from when it was available to when I began to use it surely resulted in some of my patients receiving a bearing that wore more quickly than it might have. Some of them may yet receive a revision for this incremental difference, though probably not many will. More importantly for my patients, as I waited things out, and used conventional metal-on-polyethylene bearings until crosslinked polyethylene proved itself into its second decade of service, I was able to spare my patients three ceramic bearing recalls, metal-on-metal THA with all of its complications, hip resurfacing, navigated THA, patient-specific implants, and robotic surgery; together, a sundry assortment of interventions that generally fell somewhere on the continuum of the expensive and helpful to the toxic and injurious.

There are genuinely unsolved problems in orthopaedics, even in specialties that generally serve our patients well, such as arthroplasty. We do not have great approaches for patients with extensor mechanism disruption after revision TKA, nor do we have a consistently winning solution for patients who experience chronic pelvic discontinuity after THA. In situations like those, we should be more open to surgical innovations. But we should also be open in those circumstances to counseling patients against surgery of any sort, and helping them to adapt to their disabilities, which may be better than offering someone a ninth revision procedure when the previous eight have only caused more pain.

Until or unless a new approach, implant, or tool is proven superior in ways that patients can perceive we should not use it.

Acknowledgements The author thanks Anthony M. Calabro, MA, whose suggestions improved this essay.

Funding No funding has been received for the editorial.

Declarations

Conflict of interest The author has no relevant financial or non-financial interests to disclose.

Ethical approval Ethical approval not applicable.

Informed consent Informed consent not applicable.

References

1. Howells N, Murray J, Wylde V et al (2016) Persistent pain after knee replacement: do factors associated with pain vary with

- degree of patient dissatisfaction? *Osteoarthr Cartil* 24:2061–2068. <https://doi.org/10.1016/j.joca.2016.07.012>
2. Kahlenberg CA, Lyman S, Joseph AD et al (2019) Comparison of patient-reported outcomes based on implant brand in total knee arthroplasty: a prospective cohort study. *Bone Joint J* 101-B:48–54. <https://doi.org/10.1302/0301-620X.101B7.BJJ-2018-1382.R1>
 3. Leopold SS (2020) Editor's spotlight/take 5: does knee prosthesis survivorship improve when implant designs change? Findings from the Australian orthopaedic association national joint replacement registry. *Clin Orthop Relat Res* 478:1152–1155. <https://doi.org/10.1097/CORR.0000000000001248>
 4. Lewis PL, Graves SE, de Steiger RN et al (2020) Does knee prosthesis survivorship improve when implant designs change? Findings from the Australian orthopaedic association national joint replacement registry. *Clin Orthop Relat Res* 478:1156–1172. <https://doi.org/10.1097/CORR.0000000000001229>
 5. Bouché P-A, Corsia S, Dechartres A et al (2020) Are there differences in accuracy or outcomes scores among navigated, robotic, patient-specific instruments or standard cutting guides in TKA? A network meta-analysis. *Clin Orthop Relat Res* 478:2105–2116. <https://doi.org/10.1097/CORR.0000000000001324>
 6. Budhiparama NC, Lumban-Gaol I, Ifran NN et al (2019) Does accelerometer-based navigation have any clinical benefit compared with conventional TKA? A systematic review. *Clin Orthop Relat Res* 477:2017–2029. <https://doi.org/10.1097/CORR.0000000000000660>
 7. Kim Y-H, Yoon S-H, Park J-W (2020) Does robotic-assisted TKA result in better outcome scores or long-term survivorship than conventional TKA? a randomized, controlled trial. *Clin Orthop Relat Res* 478:266–275. <https://doi.org/10.1097/CORR.0000000000000916>
 8. McAuliffe MJ, Beer BR, Hatch JJ et al (2019) Impact of image-derived instrumentation on total knee arthroplasty revision rates: an analysis of 83,823 procedures from the Australian orthopaedic association national joint replacement registry. *J Bone Joint Surg Am* 101:580–588. <https://doi.org/10.2106/JBJS.18.00326>
 9. Young SW, Walker ML, Bayan A et al (2017) The Chitranjan S. ranawat award : no difference in 2-year functional outcomes using kinematic versus mechanical alignment in TKA: a randomized controlled clinical trial. *Clin Orthop Relat Res* 475:9–20. <https://doi.org/10.1007/s11999-016-4844-x>
 10. Johns Hopkins Medicine. Mental health disorder statistics world health organization (Regional Office for Europe). Depression in Europe: Facts and Figures
 11. Leopold SS (2019) Editorial: are we all better-than-average drivers, and better-than-average kissers? Outwitting the kruger-dunning effect in clinical practice and research. *Clin Orthop Relat Res* 477:2183–2185. <https://doi.org/10.1097/CORR.00000000000000948>
 12. Leopold SS (2017) Editorial: getting the most from what you read in orthopaedic journals. *Clin Orthop Relat Res* 475:1757–1761. <https://doi.org/10.1007/s11999-017-5371-0>
 13. King J, Stamper DL, Schaad DC, Leopold SS (2007) Minimally invasive total knee arthroplasty compared with traditional total knee arthroplasty. Assessment of the learning curve and the post-operative recuperative period. *J Bone Joint Surg Am* 89:1497–1503. <https://doi.org/10.2106/JBJS.F.00867>
 14. Leopold SS (2009) Minimally invasive total knee arthroplasty for osteoarthritis. *N Engl J Med* 360:1749–1758. <https://doi.org/10.1056/NEJMct0806027>
 15. Peltola M, Malmivaara A, Paavola M (2013) Learning curve for new technology?: a nationwide register-based study of 46,363 total knee arthroplasties. *J Bone Joint Surg Am* 95:2097–2103. <https://doi.org/10.2106/JBJS.L.01296>
 16. St Mart J-P, de Steiger RN, Cuthbert A, Donnelly W (2020) The three-year survivorship of robotically assisted versus non-robotically assisted unicompartmental knee arthroplasty. *Bone Joint J* 102-B:319–328. <https://doi.org/10.1302/0301-620X.102B3.BJJ-2019-0713.R1>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.