

# Virtual Mentor

American Medical Association Journal of Ethics  
January 2015, Volume 17, Number 1: 41-48.

## STATE OF THE ART AND SCIENCE

### **Innovation in Surgery and Evidence Development: Can We Have Both at Once?**

Brett E. Youngerman, MD, and Guy M. McKhann II, MD

Surgical innovation has long enjoyed a privileged status in the evolving evidence-based medicine (EBM) paradigm. Since surgery is often considered to be as much an art as a science, many aspects of surgical practice have remained beyond the reach of formal evaluation. The culture of the surgical specialties, too, has traditionally valued innovation above standardization or external assessment. The success of this culture in producing rapid innovation over the last century has perhaps overshadowed surgery's collective missteps and allowed it to largely escape the purview of EBM and its proponents in government and industry. However, as gains slow and innovations raise costs without clear benefits, scrutiny has grown. There is now a concerted effort among patient groups, government, and private payers, as well as surgeons themselves, to bring surgery more fully into the evidence-based fold. Yet, for all these efforts, surgery's unique characteristics continue to present a number of challenges to the methods of EBM.

#### **Innovation in Surgery**

In the early days of surgery, patients desperate for a chance at a cure for dire or difficult-to-treat ailments sought out those offering the latest techniques. The landmark surgeons of the past were known for challenging the common perception about where in the body, whether it was the heart or the brain, it was possible to go safely and what was technically possible, from endoscopic surgery to transplantation.

Modern surgery has perpetuated a culture of innovation. Patients still equate the newest with the best and often seek it out [1], and surgeons and hospitals promote themselves by offering novel techniques. Surgery departments, societies, and journals reward innovation with career advancement, awards, and publications.

History is, of course, riddled with examples of excess exuberance in surgical innovation. Perhaps there is no better example than neurosurgery's enthusiastic adoption of prefrontal leucotomy for a wide range of psychiatric disorders in the 1940s. The only neurosurgical innovation to be recognized with a Nobel Prize, leucotomy was a highly valued treatment option at a time when there were few, if any, viable alternatives. However, leucotomy is remembered largely with condemnation because it was performed on tens of thousands of vulnerable patients without a clear understanding of its effects or appropriate standards of informed consent [2].

In more subtle ways, surgical innovation continues to blur the line between research and clinical practice. The 1979 Belmont Report attempted to distinguish research from innovative practice, based on the intended beneficiary [3]. According to the report, which forms the basis for current federal guidelines, practice is intended to benefit the individual patient and research is intended to test a hypothesis and produce generalizable knowledge. The report attempted to define the degree of innovation that marked a departure from standard clinical practice but nonetheless left a large gray area [4]. Institutional review boards are left to adjudicate individual cases, but, in surgery, innovation is a constant and iterative process in which many incremental changes can lead to a novel procedure before it is ever formally proposed for experimental evaluation [5, 6]. The decision to classify use of a new technique as research often falls to individual surgeons, and patients may be left without many of the systemic protections that we generally assume are in place in clinical care. The blurred boundary between research and innovative clinical practice in surgery highlights both ethical issues and the potential difficulty of applying EBM principles.

### **Evidence-Based Medicine (EBM)**

The definition of EBM is constantly evolving and can vary, but its general principles are at the foundation of modern medicine and, properly implemented, should be congruent with the goal of innovation in surgery. One widely accepted, if nonspecific, definition of EBM is “a systematic approach to clinical problem solving which allows the integration of the best available research evidence with clinical expertise and patient values” [7]. The idea is that, by drawing continuously updated conclusions from unbiased and reliable evidence, medicine will weed out false notions and identify beneficial and efficacious innovations.

Most of the surgical subspecialty societies have endorsed the broad outlines of EBM, and many have developed resources, training grants, and prestigious awards and positions to recognize and promote it [8]. However, skepticism about its implementation remains, and its current use in surgical practice is limited. The frequency of randomized controlled trials, the highest level of evidence in the EBM hierarchy, has remained low in surgical specialties [9]. Most of the recommendations that compose the countless surgical practice guidelines in the literature are based on low grades of evidence, at times little more than the personal experience of experts, and thus can be only weakly “endorsed” [10].

### **What Makes Surgery Different?**

Surgical interventions have a number of characteristics that make them more difficult to evaluate in a formal and reproducible way than other medical interventions. None of these features is unique to surgery, but their frequent appearance together collectively makes surgical interventions challenging for the EBM model.

First, surgical procedures are not a single entity but rather a complex series of interventions, the outcomes of which depend on operator, team, and setting and may vary over time [11]. The risk-benefit profile of a procedure often varies from one

surgeon to another based on numerous and sometimes difficult-to-measure factors such as their technical skill, choice of technique, judgment, training, experience, or even something as vague as demeanor [12]. Factors beyond the individual surgeon's control such as anesthesia, support staff, hospital protocols, and pre- and postoperative care also play a role and are difficult to account for even in the largest multicenter trials. To further complicate matters, none of these unmeasured variables is static. Experience in particular has been shown time and again to play an important role in outcomes, both for an individual surgeon performing a given procedure and for that procedure more broadly as the field adopts it, and levels of experience change with time and exposure [13]. A procedure may also evolve so rapidly that by the time it is evaluated it bears only a distant resemblance to what is currently popular in practice. All of these variables make it difficult to formulate practice guidelines on the basis of evaluations of a surgical intervention.

A randomized trial of unruptured brain arteriovenous malformations (ARUBA) [14] exemplifies the difficulty of standardizing the study intervention in a surgical trial. The trial was conducted at 39 centers in 9 countries to increase enrollment and, it was hoped, to randomize variables related to setting. However, to accommodate the variability in practice at the time, the treatment arm included surgery, interventional embolization, and stereotactic radiosurgery, alone or in combination, without descriptions of how any of the procedures were performed at specific sites by particular surgeons [15]. In the end, the study had limited impact on practice because it was difficult for neurosurgeons to know if the results applied to their specific treatment algorithms.

Second, blinding, one of the hallmarks of high quality clinical trials, is often difficult or impossible in surgical trials. A surgeon knows whether he or she performed the trial surgery or placebo or sham surgery. Attempts to use placebos are ethically fraught because patients in the placebo arm must be exposed to considerable risk with no promise of potential benefit, and equipoise is necessarily lost [16]. Sham surgeries have been used when they could be performed with minimal risk, such as in the case of a high-profile trial of stem cell transplantation for Parkinson disease, in which the placebo consisted of a small incision and partial-thickness burr hole in the skull [17]. A later randomized trial of implantable deep brain stimulation electrodes for Parkinson disease took advantage of the “on-off” nature of the technology, giving all patients implants but placing some patients in a temporary placebo group by leaving theirs turned off and only later turning them on in an open-label arm of the trial [18]. However, sham surgery of any kind remains controversial and cases that afford opportunity for placebo control are the exception rather than the rule.

Third, surgical clinical trials tend to suffer from low enrollment, high crossover between the control and experimental groups, and poorly defined equipoise. Since there is no regulation restricting the availability of experimental surgical procedures to the trial setting, patients can often obtain them without participating in a trial. Strong surgeon and participant preferences have significant influence on the validity of clinical trials. For participants, the invasive and risky nature of surgery may

contribute to strong preferences, which may be amplified if the alternative is nonsurgical management or a placebo. Surgeons, for their part, may be less tolerant of uncertainty than their medical counterparts [19]. They may also feel greater accountability or have stronger opinions about the procedures they perform than physicians prescribing study drugs. Decisions of equipoise must be left to individual surgeons, and crossover between experimental and control groups must be permitted to avoid interfering with what an individual surgeon judges to be in the best interest of a given patient. All this can mean that it is difficult to recruit patients for randomized surgical trials, and the assignments individual participants do receive are often discarded, which makes analysis of the results extremely difficult.

In the expensive and highly anticipated Spine Patient Outcomes Research Trial (SPORT), which compared surgical and nonsurgical management of lumbar disc herniation, only about 50 percent of the patients assigned to receive surgery did so within the study period and 30 percent of those assigned to receive nonoperative management went on to receive surgery [20]. Due to the high rate of crossover, the authors could not draw any conclusions about the superiority of the treatments based on the intention-to-treat analysis.

Due to low enrollment, many surgical trials never even reach completion. The Early Randomized Surgery Epilepsy Trial (ERSET) [21] was halted after enrolling only 38 of 200 needed participants. Although the results of this limited study indicated that surgical treatment was superior to the best medical therapy for patients with newly diagnosed, intractable temporal lobe epilepsy, the study will be remembered more for its enrollment difficulties. Similarly, recruitment for the Radiosurgery or Open Surgery for Epilepsy (ROSE) Trial [22] was recently discontinued due to inability to enroll a sufficient number of participants with temporal lobe epilepsy who were willing to be randomized to radiosurgery versus temporal lobectomy.

### **EBM's Implications for Surgical Innovation**

Given the limitations of evidence-based medicine in surgery, there is appropriate concern among surgeons that the rising use of available evidence to make coverage and reimbursement decisions could result in problematic requirements or guidelines. The Leapfrog Group, an independent group of corporations and public agencies that provide health benefits for their employees, and private insurers have begun to use evidence-based guidelines to assess the quality of care offered by clinicians [23]. The Centers for Medicare and Medicaid Services (CMS) is linking a portion of reimbursement to performance on evidence-based quality measures [24]. Current efforts are focused on simple, well-studied, primarily perioperative practices like use of antibiotics or beta blockers. Overly prescriptive attempts to further standardize surgical practice based on the limited body of evidence, however, are likely to be met with resistance. If only those procedures meeting excessively high evidentiary standards were to be covered by insurers, this would have serious consequences for care quality and surgical innovation.

The demand for better evidence is likely to continue growing, however. Numerous prestigious groups have outlined paths forward [25, 26]; the common emerging theme is a more concerted effort at developing evidence without stifling innovation.

There is a need for both improved observational and experimental evidence. For observational data, single-center cases series are being replaced by audited national registries with standardized definitions of variables, outcomes, and complications [27, 28]. Such registries, while administratively burdensome and not immune to the biases inherent in observational data (i.e., selection bias, information bias, and confounding), can vastly improve the available evidence in surgical specialties at a fraction of the cost of clinical trials. Ongoing, large-scale data collection is changing the way new statistical information is incorporated into evidence-based decision making [29]. CMS has demonstrated willingness to allow participation in these databases, designed by surgical subspecialty societies, to meet its national quality reporting requirements [30].

Experimental methods are constantly evolving to meet the realities of surgical practice. Randomized controlled trials will remain the gold standard of evidence, but new trial designs are being attempted that can incorporate learning curves and patient and physician preferences [26]. Financial incentives that encourage evidence development rather than simply denying payment for unproven techniques will also be critical. There are substantial new government funding streams directed at answering clinical questions, such as the Patient Centered Outcomes Research Institute (PCORI) [31]. Innovative coverage schemes, like Medicare's coverage with evidence development [32] and reimbursement for the routine costs of clinical trials [33], aim to encourage the collection of data as innovation occurs rather than mandating burdensome levels of evidence collection before new ideas can receive financial support.

### **Conclusions**

Innovation is at the heart of surgery's culture and mission. The growing demand for evidence in support of clinical practice poses significant challenges for surgery. Given that surgical interventions depend significantly on the surgeon, patient, and setting, attempts to measure outcomes and standardize decision making are difficult to integrate and, not surprisingly, viewed unfavorably [4]. Nonetheless, the goals of evidence-based medicine are ultimately supportive of innovation that aims to maximize patient well-being. With prudent observational and experimental research designs and thoughtful financial and policy support it should be possible to simultaneously promote innovation and evidence development.

### **References**

1. Reitsma AM, Moreno JD. Ethics of innovative surgery: US surgeons' definitions, knowledge, and attitudes. *J Am Coll Surg*. 2005;200(1):103-110.
2. Wind JJ, Anderson DE. From prefrontal leukotomy to deep brain stimulation: the historical transformation of psychosurgery and the emergence of

neuroethics. *Neurosurg Focus*. 2008;25(1):E10.  
doi:10.3171/FOC/2008/25/7/E10.

3. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. April 18, 1979. <http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html>. Accessed October 4, 2014.
4. Mastroianni AC. Liability, regulation and policy in surgical innovation: the cutting edge of research and therapy. *Health Matrix Clevel*. 2006;16(2):351-442.
5. Strasberg SM, Ludbrook PA. Who oversees innovative practice? Is there a structure that meets the monitoring needs of new techniques? *J Am Coll Surg*. 2003;196(6):938-948.
6. Barkun JS, Aronson JK, Feldman LS, et al; Balliol Collaboration. Evaluation and stages of surgical innovations. *Lancet*. 2009;374(9695):1089-1096.
7. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312(7023):71-72.
8. Maier RV. What the surgeon of tomorrow needs to know about evidence-based surgery. *Arch Surg*. 2006;141(3):317-323.
9. Chang DC, Matsen SL, Simpkins CE. Why should surgeons care about clinical research methodology? *J Am Coll Surg*. 2006;203(6):827-830.
10. Naylor CD. Grey zones of clinical practice: some limits to evidence-based medicine. *Lancet*. 1995;345(8953):840-842.
11. Ergina PL, Cook JA, Blazeby JM, et al; Balliol Collaboration. Challenges in evaluating surgical innovation. *Lancet*. 2009;374(9695):1097-1104.
12. McLeod RS, Wright JG, Solomon MJ, Hu X, Walters BC, Lossing AI. Randomized controlled trials in surgery: issues and problems. *Surgery*. 1996;119(5):483-486.
13. Cook JA, Ramsay CR, Fayers P. Statistical evaluation of learning curve effects in surgical trials. *Clin Trials*. 2004;1(5):421-427.
14. Mohr JP, Parides MK, Stapf C, et al; International ARUBA Investigators. Medical management with or without interventional therapy for unruptured brain arteriovenous malformations (ARUBA): a multicentre, non-blinded, randomised trial. *Lancet*. 2014;383(9917):614-621.
15. Russin J, Spetzler R. Commentary: the ARUBA trial. *Neurosurgery*. 2014;75(1):e96-e97. doi:10.1227/NEU.0000000000000357.
16. Bonchek LI. Sounding board. Are randomized trials appropriate for evaluating new operations? *N Engl J Med*. 1979;301(1):44-45.
17. Freeman TB, Vawter DE, Leaverton PE, et al. Use of placebo surgery in controlled trials of a cellular-based therapy for Parkinson's disease. *N Engl J Med*. 1999;341(13):988-992.
18. Deuschl G, Schade-Brittinger C, Krack P, et al; German Parkinson Study Group, Neurostimulation Section. A randomized trial of deep-brain stimulation for Parkinson's disease. *N Engl J Med*. 2006;355(9):896-908.

19. McCulloch P, Kaul A, Wagstaff GF, Wheatcroft J. Tolerance of uncertainty, extroversion, neuroticism and attitudes to randomized controlled trials among surgeons and physicians. *Br J Surg*. 2005;92(10):1293-1297.
20. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. *JAMA*. 2006;296(20):2451-2459.
21. Engel J Jr, McDermott MP, Wiebe S, et al; Early Randomized Surgical Epilepsy Trial (ERSET) Study Group. Early surgical therapy for drug-resistant temporal lobe epilepsy: a randomized trial. *JAMA*. 2012;307(9):922-930.
22. University of California, San Francisco. Radiosurgery or Open Surgery for Epilepsy Trial (ROSE). <http://clinicaltrials.gov/show/NCT00860145>. Accessed October 12, 2014.
23. Birkmeyer JD, Dimick JB. Potential benefits of the new Leapfrog standards: effect of process and outcomes measures. *Surgery*. 2004;135(6):569-575.
24. Ryan AM, Press MJ. Value-based payment for physicians in Medicare: small step or giant leap? *Ann Intern Med*. 2014;160(8):565-566.
25. McClennan MB, McGinnis JM, Nabel EG, Olsen LM; Institute of Medicine. *Evidence-Based Medicine and the Changing Nature of Healthcare: 2007 IOM Annual Meeting Summary*. Washington, DC: National Academies Press; 2008.
26. McCulloch P, Altman DG, Campbell WB, et al; Balliol Collaboration. No surgical innovation without evaluation: the IDEAL recommendations. *Lancet*. 2009;374(9695):1105-1112.
27. Edwards FH, Grover FL, Shroyer AL, Schwartz M, Bero J. The Society of Thoracic Surgeons National Cardiac Surgery Database: current risk assessment. *Ann Thorac Surg*. 1997;63(3):903-908.
28. Asher AL, McCormick PC, Selden NR, Ghogawala Z, McGirt MJ. The National Neurosurgery Quality and Outcomes Database and NeuroPoint Alliance: rationale, development, and implementation. *Neurosurg Focus*. 2013;34(1):e2. doi:10.3171/2012.10.FOCUS12311.
29. Angevine PD, McCormick PC. Measuring clinical practice: science and methods. *Neurosurg Focus*. 2013;34(1):e3. doi:10.3171/2012.10.FOCUS12299.
30. Centers for Medicare and Medicaid Services. 2014 Physician Quality Reporting System Qualified Clinical Data Registries. <http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/Downloads/2014QCDRPosting.pdf>. Accessed December 2, 2014.
31. Selby JV, Lipstein SH. PCORI at 3 years—progress, lessons, and plans. *N Engl J Med*. 2014;370(7):592-595.
32. Tunis SR, Pearson SD. Coverage options for promising technologies: Medicare's "coverage with evidence development." *Health Aff (Millwood)*. 2006;25(5):1218-1230.

33. Aaron HJ, Gelband H, eds; Institute of Medicine. *Extending Medicare Reimbursement in Clinical Trials*. Washington, DC: National Academy Press; 2000.

Brett E. Youngerman, MD, is a resident in neurosurgery at Columbia University Medical Center in New York City. He was a recipient of the Doris Duke Clinical Research Fellowship in 2010-2011. His interests include movement disorders, functional imaging and neurophysiology, and outcomes and quality improvement.

Guy M. McKhann II, MD, is a neurosurgeon and translational neuroscientist at Columbia University Medical Center/New York Presbyterian Hospital in New York City. At Columbia, he is the director of the Adult Hydrocephalus Center and of epilepsy and movement disorder surgery and awake brain mapping for tumors and epilepsy. In addition, he leads the Epilepsy Neurophysiology Laboratory and extensively collaborates with the Brain Tumor and Cognitive Neurophysiology research programs.

**Related in VM**

[Technological Innovation and Ethical Response in Neurosurgery](#), January 2015

[Integrating Ethics into Science Education and Research: Report of the Presidential Commission for the Study of Bioethical Issues](#), January 2015

[A Preparatory Neuroethical Approach to Assessing Developments in Neurotechnology](#), January 2015

[Risks and Benefits of Innovative Off-Label Surgery](#), February 2010

[Deep Brain Stimulation: Calculating the True Costs of Surgical Innovation](#), February 2010

*The viewpoints expressed on this site are those of the authors and do not necessarily reflect the views and policies of the AMA.*

Copyright 2015 American Medical Association. All rights reserved.