DOI:
10.1111/bju.13778

Weighing the Evidence from Surgical Trials

Quoc-Dien Trinh, MD[1] Alexander P. Cole MD[1], Prokar Dasgupta, MD, FRCS(Urol)[2]

[1]Division of Urological Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA
[2]Guy's Hospital, King's College London, England.

Keywords: Surgical Trials, Quality Improvement, Comparative Effectiveness, Outcomes Research, Research Methods

Acknowledgments: Quoc-Dien Trinh is supported by an unrestricted educational grant from the Vattikuti Urology Institute, a Clay Hamlin Young Investigator Award from the Prostate Cancer Foundation and a Genentech BioOncology Career Development Award from the Conquer Cancer Foundation of the American Society of Clinical Oncology. Alexander Cole is supported by a Brigham Research Institute micro-grant. Prokar Dasgupta acknowledges support from the MRC Centre for Transplantation, KCL, NIHR Biomedical Research Centre grant to GSTT-KCL, King's Health Partners, Prostate Cancer Research Centre, Prostate Cancer UK, The Urology Foundation, The Guys and St. Thomas' Charity, The Vattikuti Foundation

Conflicts of Interest: Dr. Trinh, Professor Dasgupta and Dr. Cole have no relevant conflicts of interest to declare.

Corresponding Author:
Quoc-Dien Trinh, M.D.
Division of Urological Surgery
Brigham and Women's Hospital
45 Francis St, ASB II-3; Boston, MA 02115, USA
Tel: +1 617 525-7350; Fax: +1 617 525-6348
Email: qtrinh@bwh.harvard.edu

Word Count: 987/1000 **Reference:** 6/6

With growing calls to improve value in health care, the assessment of surgical outcomes has moved to the spotlight. Public awareness of medical errors has spurred initiatives like the ProPublica "Surgeon Scorecard" to measure and report complications (https://projects.propublica.org/surgeons/). High-tech and expensive innovations such as robot-assisted surgery must be measured against traditional approaches. Together, these factors and others have spurred calls to measure, assess and compare surgical techniques. The past months have seen publication of two high-profile randomized trials evaluating surgical techniques in urology. The first was early results from the first phase-3 randomized trial comparing open and robotic prostatectomy.[1] Second was the publication of 10-year outcomes from the ProtecT trial, comparing surgery, radiation and active monitoring for prostate cancer.[2]

Using trials to evaluate techniques is seemingly obvious. Since James Lind's first trial of treatments for scurvy on eighteenth-century British sailors, prospective randomized trials have been the gold standard for assessing medical interventions. The concept is simple. Patients are assigned to one or more interventions and otherwise treated the same. Properly designed randomized trials can minimalize spurious causality, reduce bias and give as close a picture to "true" causal relationships as possible. **The authors of such clinical trials should be commended given the years of work and substantial hurdles involved in designing and executing these trials.** With that said, we believe that there are some key issues which should be highlighted when clinical trials are used to evaluate surgical techniques and effectiveness.

Among the most evident limitations of surgical trials are practical considerations. While use of placebos, allocation concealment and blinding is *de rigueur* in pharmaceutical trials, these are difficult to achieve in surgery. Some aspects of surgical technique, for example stapled versus sewn anastomosis, or small differences in equipment can be tested without patients or outcomes assessors being aware. But blinding patients to large differences in technique (e.g. open versus robotic approach) or surgical versus non-surgical approaches (e.g. surgery versus radiation or endoscopic approaches) is challenging.

Additionally, there is a large question about confounding. For all but the smallest details of operative techniques, patients' clinical team must know what surgery has been done. This is vital for providing appropriate postoperative care, but also allows for potential differences in care, which could skew outcomes.

The generalizability of surgical trials may be an issue if the surgeons in the study are not representative of those in **non-academic** practice. To use a recent example – both the Asymptomatic Carotid Trial 1 (ACT I) and Carotid Revascularization Endarterectomy versus Stenting Trial [3] published results comparing surgical versus endovascular treatment of asymptomatic carotid artery stenosis.[3] [4] Both used a credentialing process to guarantee that only the best surgeons and interventionists performed the interventions.[5] But even when these steps are taken, if the procedures turn out to have greatly different learning curves or require different levels of skill, results achieved at high-volume academic centers may not be seen **at non-academic or private centers**.

Another issue related to headline-grabbing clinical trials is that media-based dissemination of information on novel surgical techniques leads to alteration in practice-based non-representative surgical practice, these changes may be unwarranted. There are examples of media misinterpretation of clinical studies (http://www.nytimes.com/2016/07/21/health/advanced-prostate-cancer-false-alarm.html) – headlines often fail to convey the clinical complexity and true content of many types of studies.

Ultimately, trials comparing surgical techniques depend on surgeon expertise. The most that a clinical trial of surgery can show is whether surgeons A, B, C performing procedure X are better than surgeons D, E, F performing procedure Y. For example, in the above study by Yaxley, et al, men were randomized to either open surgery with surgeon A, or robotic surgery with surgeon B. The authors should be commended for planning and executing this innovative and challenging study. But does this really answer the question? Patients and the public need to know if robot-assisted prostatectomy is *generally* better, not whether surgeon A is better than surgeon B at their respective techniques.

Preoperative selection, operative technique, surgical equipment, and post-operative care are components of complicated *systems* with multiple inputs and interlocking parts. Finding out how to optimize operative systems and techniques is vital – but may require analytic techniques that go beyond traditional randomized trials.

When large companies try to optimize manufacturing systems or large global supply chains, they experiment, analyze and adapt. Real-time data gathering allows fine-tuning. It wouldn't make sense to design a series of vast "trials" where products are "randomized" to different manufacturing systems altering only one variable at a time in complex supply and manufacturing systems, with results measured for months or years. Analytic tools such as statistical process control and design of experiments allows these companies to determine and measure important variables in complex systems.[6] These techniques may prove vital in the study of complex systems like surgical care.

Randomized controlled trials provide ironclad evidence on the superiority of drug A versus B or versus placebo. They can also provide key insight on some questions in surgery. However, at the same time, it is essential to understand how surgery is different. It is often said that surgery is more "an art than a science." While this cliché does not excuse us from applying scientific rigor, it does highlight that surgery may require novel evaluative approaches compared to prospective trials. We believe that novel; data-driven approaches that encompass the complexity of surgical care will provide key complementary insights for surgical evaluation and quality improvement in years to come.

In fact, there are already examples of pioneering surgical improvement initiatives such as the Michigan Surgical Quality Collaborative **in the United States, or the IDEAL collaborative in the United Kingdom (http://www.ideal-collaboration.net/), which** emphasize real-time, systems-based improvements and analysis within existing hospital systems. **The paradigm of such initiatives is that innovation; evaluation and validation in complex and dynamic systems such as surgery can and should happen *in parallel*, and that while clinical trials provide extremely reliable information, complementary and more nimble approaches are vital as well.**

1. Yaxley JW, Coughlin GD, Chambers SK, Occhipinti S, Samaratunga H, Zajdlewicz L, et al. Robot-assisted laparoscopic prostatectomy versus open radical retropubic prostatectomy: early outcomes from a randomised controlled phase 3 study. The Lancet.
2. Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P, et al. 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. The New England journal of medicine. 2016.
3. Brott TG, Howard G, Roubin GS, Meschia JF, Mackey A, Brooks W, et al. Long-Term Results of Stenting versus Endarterectomy for Carotid-Artery Stenosis. The New England journal of medicine. 2016;374(11):1021-31.
4. Rosenfield K, Matsumura JS, Chaturvedi S, Riles T, Ansel GM, Metzger DC, et al. Randomized Trial of Stent versus Surgery for Asymptomatic Carotid Stenosis. The New England journal of medicine. 2016;374(11):1011-20.
5. Spence JD, Naylor AR. Endarterectomy, Stenting, or Neither for Asymptomatic Carotid-Artery Stenosis. The New England journal of medicine. 2016;374(11):1087-8.

6. O'Brien T, Viney R, Doherty A, Thomas K. Why don't mercedes benz publish randomized trials? BJU international. 2010;105(3):293-5.