

## REVIEW

## Limits to clinical trials in surgical areas

Marco Kawamura Demange,<sup>1</sup> Felipe Fregni<sup>II</sup><sup>I</sup>Orthopaedics and Traumatology, Faculdade de Medicina da Universidade de Sao Paulo, São Paulo, São Paulo, Brasil. <sup>II</sup>Harvard Medical School - Department of Physical Medicine and Rehabilitation Boston, Massachusetts, United States.

Randomized clinical trials are considered to be the gold standard of evidence-based medicine nowadays. However, it is important that we point out some limitations of randomized clinical trials relating to surgical interventions. There are limitations that affect the external and internal validity of many surgical study designs. Some limitations can be bypassed, but can make it more difficult for the study to be carried out. Other limitations cannot be bypassed. When it is intended to extrapolate the result of a randomized clinical trial, the premise is that the performed or to be performed intervention will be similar wherever applied and/or for every doctor using it. However, no matter how standardized the technique may be, the results are not similar for all surgeons, which implies a significant limitation to surgical randomized clinical trials concerning external validity. When considering the various limitations presented for performing surgical trials capable of generating scientific evidence within the patterns currently proposed in the evidence level classifications of medical publications, it is necessary to rethink whether those scientific evidence levels are similarly applicable to surgical works and to nonsurgical trials. We currently live in a time of supposed "inferiority" of surgical scientific works under the optics of the current quality criteria for a "suitable" clinical trial.

**KEYWORDS:** Clinical Trial; Surgery; Randomization; Blinding; Evidence Level.

Demange MK, Fregni F. Limits to clinical trials in surgical areas. Clinics. 2011;66(1):159-161.

E-mail: demange@usp.br

Tel.: +55-11-7754-9607

Randomized clinical trials are nowadays considered the gold standard of evidence-based medicine.<sup>1</sup> Clinical trials aim at creating subsidies to ground the use of one or other therapeutic by means of analysis of different interventions (for instance, medications), or comparing interventions with placebos.<sup>2,3</sup> In their essence, the scope of clinical trials is to create scientific evidence to direct medical conduct.<sup>4</sup>

Displaying internal and external validity is a condition for the clinical trial to have potential to suggest that a given therapeutic is the best indicated one. Internal validity presupposes that the work design be suitable to answer the question posed by the study. External validity, in turn, implies the potential of the trial results to be extrapolated to other populations (being valid not only for the survey participants). Displaying internal validity is a *sine qua non* condition to display external validity, that is methodological problems affecting the study itself do not allow its results to be extrapolated.<sup>5</sup>

Therefore, it is important that we point out some limitations displayed by randomized clinical trials relating to surgical interventions,<sup>6</sup> limitations that affect the external and internal validity of many surgical studies designs.<sup>7</sup>

- Adherence of participants: the random selection of patients in distinct surgical treatment groups implies lower patient adherence than clinical studies. In cases of different surgical interventions in each group or in comparisons with nonsurgical treatment, this adherence is even lower. Randomization itself is not attractive to patients, and is less attractive in surgical trials; for patients it is easier not to know which medication will be prescribed to them than not to know the type of surgery that will be performed on them.
- Feasibility of both groups: the random selection of patients requires that both surgical techniques can be performed in compliance with the selection. Any limitation to the performance of one of the techniques (in a given week, the surgeon's determination, etc.) implies a risk of limiting the selection. For instance, in cases of urgency or trauma the selections can be limited when both techniques are not performed by all participating doctors.
- Randomization-based data analysis: in a clinical trial it is suggested to perform the obtained data analysis based on the group to which the patient was selected, a principle named "intention to treat analysis".<sup>8,9</sup> In surgical trials, following this principle can generate strange or even incongruent responses. For instance, a patient assigned a nonsurgical treatment who due to whatever reason receives the surgical treatment, following the intention-to-treatment principle, should be analyzed as nonsurgical. If this patient displayed an infection at the surgery site, we would have a "surgical site infection occurrence" as a "nonsurgical treatment complication" in the results.

Therefore, the application of this intention-to-treat principle, which aims at ensuring sample randomization (and homogenization among the groups), displays serious limitations in the surgical area.

- Need for a longer interval between randomization and intervention: drug treatment randomization can be carried out beforehand, considered ideal, or upon meeting the patient ("flip of a coin"), considered simpler. In surgical works, depending on specific instruments or preparations (such as orthopedic implants) before the surgery, randomization must necessarily precede the intervention by a significant time interval.

As to methodology, surgical trials display greater difficulty instituting "blinding" of the patients and examiners.<sup>10-12</sup> Some limitations can be bypassed<sup>13</sup> but make it more difficult for the study to be carried out. Examples of surgical interventions that limit study "blinding" are provided below:

- the type of scar displayed (and the absence of scar in nonsurgical groups)
- other detectable differences, such as radiographs in orthopedic implant cases
- "clues" that indicate the performed treatment, such as the day of surgery or the surgeon's name, when a particular technique is performed only by one doctor.
- different types of rehabilitation or aspects of postoperative care.
- the assessment or part of the data collection depends on the surgeon themselves (operator bleeding data, surgery time, diagnosis of complications)
- diverse medical procedures to be followed in the postoperative period depend on the doctor knowing a patient's intervention. This limits the blinding of the doctor assisting the patient, and therefore makes it more troublesome to blind the study.

In the study design, the choice of a surgical trial control group also displays limitations. In surgical trials the intervention is permanent, it is not possible to use the patient as their own control for two consecutive interventions in cases involving single organs. In cases of two organs (surgery in both eyes or in both knees), at the time of the second intervention the patient is not the same as at the time of the first, because of progress due to the first intervention. In other words, it is not possible to perform a crossover-type design. Besides, it is convenient to highlight that the use of placebo groups in surgical trials frequently presents important ethical limitations.<sup>14-16</sup>

When it is intended to extrapolate a clinical trial result, the premise is that the performed or to be performed intervention (for which the existence of scientific evidence will be sustained) will be similar for the entire institution or for all doctors. Thus, we consider that a given medication will have similar effects if prescribed by different doctors at different places. This is because the drug and the drug action mechanism are the same independently of who prescribes it.

Can this be valid in surgical trials?

Often in surgical trials the intervention performed by a surgeon is not necessarily identical to that performed by another surgeon. That is, no matter how reproducible the technique is, it is not identical. Besides, in general, new

surgical techniques depend on a learning curve, a curve that can vary for each technique and for each surgeon. This means that even for the same surgeon, surgeries performed at different points of the learning curve can substantially differ.

In this sense, the results of a surgical technique obtained by particular authors can be very different from those obtained by other authors, without implying methodological failures or a chance effect, but instead because the interventions are not the same.

Using a culinary example, no matter how familiar a recipe is, the taste will not necessarily be the same when prepared by several different chefs. The recipe is the same, but the result is quite different. This is also true in surgical activity. No matter how standardized the technique may be, the result is not similar for all surgeons. This phenomenon implies a significant limitation to the external validity of surgical works, particularly for those performed by small groups of surgeons.

Another important limitation frequently presented by surgical trials consists in the restricted number of cases studied and the consequent implication on the statistical power of the work. The power of a study depends on the sample size, sample homogeneity, homogeneity of results (standard deviation of the results) and on the difference among the result averages for each group. Frequently in surgical studies the number of patients (sample size) is lower than the minimum required to achieve an 80% statistical power, a percentage considered as the minimum acceptable in most clinical trials.<sup>17</sup>

In view of that, multicenter studies are often proposed, to expand the number of individuals to be included. Again, intervention standardization bias is encountered. It is difficult to state that the surgical intervention performed in various centers is the same as there is surgeon variability and inclusion of innumerable other variables that influence this. For example, the type of anesthesia, different anesthetic and surgical equipment, the method of sterilizing surgical materials, the nursing team's expertise in postoperative care, organization of the surgical center team (influencing surgery times), the availability of different technologies (such as the quality of equipment, e.g. radioscope, microscope) and the different hospital bacterial flora are among several other factors (detectable and undetectable) that influence the surgical procedure. It can even be argued that these biases influence both arms of the work and their effect will be diluted with increasing numbers of individuals. This may be true in some cases; however, it can generate a contrary effect in others, when a given factor causes a bias in only one of the arms (radioscopy use, when necessary in only one of the surgical techniques, for instance). Besides, in the cases in which the increase in the number of individuals can dilute this variance the study will then require a huge sample, rendering a study impracticable from financial and medical viewpoints.

Thus, it becomes necessary to rethink whether those scientific evidence levels are similarly applicable to surgical works and to nonsurgical trials, when considering the various limitations of surgical trials<sup>7,18</sup> capable of generating scientific evidence within the evidence level classifications currently proposed in medical publications.

Perhaps we should rethink the valorization currently granted to knowledge derived from case series and operator techniques and opinions from specialists in surgical areas.<sup>19</sup> Or even ask the question: Are randomized clinical

trials actually the design for bringing a higher level of evidence to surgical trials? Along this same line, it should be discussed: Are the publication acceptance criteria for a surgical scientific work the same as for clinical works?

At the current scientific moment, it is urgent to reflect on the direction that medical research in surgical areas should take. We currently live at a time of supposed “inferiority” of surgical scientific works under the optics of the current quality criteria for a “suitable” clinical trial.<sup>18,20,21</sup>

Besides, this reflection is fundamental, as it reflects not only on how the clinical trials already published or to be published must be viewed, but also on other scientific articles derived therefrom, such as systematic reviews, currently acclaimed as a high level of scientific evidence.

## REFERENCES

- Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised consort statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med.* 2001;134:663–94.
- Petrelli NJ. Clinical trials are mandatory for improving surgical cancer care. *Jama.* 2002;287:377–8.
- Nelson H, Boller AM. The role of clinical trials in the development of novel surgical techniques. *J Surg Oncol.* 2007;96:704–9.
- Gray M, Meakins JL. Evidence-based surgical practice and patient-centered care: inevitable. *Surg Clin North Am.* 2006;86:217–20.
- Viera AJ, Bangdiwala Si. eliminating bias in randomized controlled trials: importance of allocation concealment and masking. *Fam Med.* 2007;39:132–7.
- Singh JA, Murphy S, Bhandari M. Assessment of the methodologic quality of medical and surgical clinical trials in patients with arthroplasty. *J Rheumatol.* 2009;36:2642–54.
- Mcculloch P, Taylor I, Sasako M, Lovett B, Griffin D. Randomised trials in surgery: problems and possible solutions. *BMJ.* 2002;324:1448–51.
- Diener MK, Blumle A, Szakallas V, Antes G, Seiler CM. [Randomized and nonrandomized controlled clinical trials in a German surgical journal]. *Chirurg.* 2006 Sep;77(9):837–43.
- Schoenfeld PS. Evidence-based medicine in practice: applying intention-to-treat analysis and perprotocol analysis. *Am J Gastroenterol.* 2005;100:3–4.
- Bang H, NI L, Davis CE. Assessment of blinding in clinical trials. *control clin trials.* 2004;25:143–56.
- Schulz Kf, Grimes Da. Blinding in randomised trials: hiding who got what. *Lancet.* 2002 23;359:696–700.
- Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. *BMJ.* 2000;321:504.
- Karanicolas PJ, Bhandari M, Taromi B, Akl EA, Bassler D, Alonso-Coello P, et al. Blinding of outcomes in trials of orthopaedic trauma: an opportunity to enhance the validity of clinical trials. *J Bone Joint Surg Am.* 2008;90:1026–33.
- Birch S. A review and analysis of placebo treatments, placebo effects, and placebo controls in trials of medical procedures when sham is not inert. *J Altern Complement Med.* 2006;12:303–10.
- London AJ, Kadane JB. Placebos that harm: sham surgery controls in clinical trials. *Stat Methods Med Res.* 2002;11:413–27.
- Mehta S, Myers TG, Lonner JH, Huffman GR, Sennett BJ. The ethics of sham surgery in clinical orthopaedic research. *J Bone Joint Surg Am.* 2007;89:1650–3.
- Cadeddu M, Farrokhyar F, Thoma A, Haines T, Garnett A, Goldsmith Ch. Users’ guide to the surgical literature: how to assess power and sample size. Laparoscopic vs open appendectomy. *Can J Surg.* 2008;51:476–82.
- Slim K. Limits of evidence-based surgery. *World J Surg.* 2005;29:606–9.
- Young JM, Hollands MJ, Ward J, Holman CD. Role for opinion leaders in promoting evidence-based surgery. *Arch Surg.* 2003;138:785–91.
- Horton R. surgical research or comic opera: questions, but few answers. *Lancet.* 1996;347:984–5.
- Kitto S, Villanueva EV, Chesters J, Petrovic A, Waxman BP, Smith JA. Surgeons’ attitudes towards and usage of evidence-based medicine in surgical practice: a pilot study. *Anz J Surg.* 2007;77:231–6.