*Review Article*

# Randomized controlled trials in neurosurgery

Radwan Takroni[1], Sunjay Sharma[1], Kesava Reddy[1], Nirmeen Zagzoog[1], Majid Aljoghaiman[1], Mazen Alotaibi[1], Forough Farrokhyar[2]

Departments of [1]Neurosurgery and [2]Health, Evidence and Impact, McMaster University, Hamilton, Ontario, Canada.

E-mail: *Radwan Takroni - radwan.takroni@medportal.ca; Sunjay Sharma - sunjay.sharma@gmail.com; Kesava Reddy - kesh@keshreddy.ca; Nirmeen Zagzoog - nirmeen.zagzoog@medportal.ca; Majid Aljoghaiman - majid.aljoghaiman@medportal.ca; Mazen Alotaibi - mazen.alotaibi@medportal.ca; Forough Farrokhyar - farrokh@mcmaster.ca
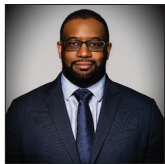
**\*Corresponding author:**
Radwan Ibrahim Takroni,
Department of Neurosurgery,
McMaster University, Hamilton,
Ontario, Canada.

radwan.takroni@medportal.ca

**Quick Response Code:**

## ABSTRACT

Randomized controlled trials (RCTs) have become the standard method of evaluating new interventions (whether medical or surgical), and the best evidence used to inform the development of new practice guidelines. When we review the history of medical versus surgical trials, surgical RCTs usually face more challenges and difficulties when conducted. These challenges can be in blinding, recruiting, funding, and even in certain ethical issues. Moreover, to add to the complexity, the field of neurosurgery has its own unique challenges when it comes to conducting an RCT. This paper aims to provide a comprehensive review of the history of neurosurgical RCTs, focusing on some of the most critical challenges and obstacles that face investigators. The main domains this review will address are: (1) Trial design: equipoise, blinding, sham surgery, expertise-based trials, reporting of outcomes, and pilot trials, (2) trial implementation: funding, recruitment, and retention, and (3) trial analysis: intention-to-treat versus as-treated and learning curve effect.

**Keywords:** Neurosurgery, Randomized controlled trials, Research methodology

## INTRODUCTION

The term randomized controlled trial (RCT) refers to a type of study in which people are allocated randomly to receive one or more clinical interventions. One of these interventions is usually the standard of care, also known as the control. The control may be a standard practice, a placebo (e.g., sugar pill), or no intervention at all.[39] People who take part in an RCT are called participants or subjects. RCTs seek to measure and compare the outcomes after the participants receive the intervention in question. Because the outcomes are measured, RCTs are quantitative studies. The main advantage of the RCT design is that it minimizes selection bias and the effect of the known and unknown confounders by producing balanced intervention and comparison groups. The term "randomized" indicates that the allocation of the participants to study groups is solely by chance, that is, random. The term "controlled" denotes the fact that the new intervention is being compared to a control group. If conducted properly, RCTs can yield powerful evidence.

The history of RCTs in neurosurgery is relatively young compared to other medical and some surgical specialties. Neurosurgery as a specialty did not start until the late 19th century, which could explain why it is one of the specialties that had a late start in exploring new treatments/

interventions through controlled trials.[60] The first identified RCT in neurosurgery was a comparison between surgical and nonsurgical management of posterior communicating artery aneurysms published by McKissock *et al.* in 1960.[37] This RCT created an immense controversy in the field and led to many subsequent publications discussing both clinical and methodological aspects of the trial, criticizing its design and generalizability.

To illustrate the poor quality and low quantity of neurosurgical RCTs during that time, Haines reviewed 51 neurosurgical RCTs after searching the literature published between 1945 and 1981.[24] The quality scores of the reviewed articles ranged between 0.09 and 0.82, with a mean score of 0.47 before 1978 and a mean score of 0.57 after 1978, suggesting a trend of improvement as 50% of the studies were published before 1977. Sample sizes were small, with a median of 66 patients. One of the main reasons neurosurgical RCTs were poor in quality during that period was the lack of biostatisticians' involvement (only 35% of the studies acknowledged their participation). In addition, failure to mention blinding of the data analyst (61%), start and end dates of the studies (45%), providing an incomplete definition of the therapeutic regimen used in the study (35%), and outlining unclear criteria of the selected patients (39%) also hindered the studies' quality. The authors, further, recommended that investigators should take into consideration the issue of small sample size and the statistical power needed to answer the question of interest.

Years later, Mansouri *et al.* reviewed 61 neurosurgical RCTs published between 2000 and 2014.[33] This review continued to show issues with reporting blinding (failed to report in 65.8% of studies), sample size (median of 100), and protocol implementation. Conversely, there have been improvements regarding clarity of reporting eligibility criteria, study objectives, and statistical methods. A similar review of 108 RCTs by Vranos *et al.*, with a median sample size of 68 patients, showed that only 28.7% of studies described allocation concealment, 21.3% gave power calculations, and 13.6% were double blinded.[55] A recent review by Azad *et al.* assessed 401 articles published between 2003 and 2016 by applying the same method employed by Vranos *et al.*[3] The median sample size was 73 patients. Only 28.9% of the articles detailed allocation concealment, 35.4% described power calculations, and 13% were double-blinded RCTs. We can conclude from these reviews that the prevalence of neurosurgical RCTs is low, and the quality of RCT design and reporting in neurosurgery also remains suboptimal.

## CHALLENGES IN NEUROSURGICAL RCTs

Despite advancements over the past 50 years, the quality of designing and reporting of neurosurgical RCTs remains suboptimal. In this section, we outline the important aspects commonly encountered during designing, implementing, and analyzing RCTs in neurosurgery. While focusing on some of the most critical challenges and obstacles that face investigators, we will attempt to provide suggestions for potential solutions.

## TRIAL DESIGN

### Equipoise

Equipoise refers to the situation in which there is no clear evidence that one intervention is superior or inferior to another, which constitutes the rationale behind conducting RCTs. This is common when comparing surgical to medical treatment. However, some trialists fail to make the distinction between two related but completely different concepts: Clinical Equipoise and Individual Equipoise. Individual Equipoise exists when the clinician involved in the research study has no preference or is truly uncertain about the overall benefit or harm offered by the treatment.[2] Clinical Equipoise, on the other hand, is a term that relates to the collective opinion of a body of clinicians whose consensus opinion is that none of the various interventions in a clinical trial is clearly superior,[57] which applies to the profession as a whole (as expressed in guidelines or recommendations). This means that physicians with Individual Equipoise will enroll their patients in a clinical trial simply because they do not have a preference as to which treatment is better.

Individual Equipoise, however, introduces inherent bias and, as such, may not equate to Clinical Equipoise, which is typically based on the best available literature at that time. However, when Clinical Equipoise is utilized as a base for designing a clinical trial, the participating neurosurgeons will have to set aside their individual biases and agree a priori on the study design and methodology, especially the inclusion and exclusion criteria for a research study. They will also have to be bound by the consensus opinion, which will enable the physician involved in the trial to make decisions efficiently.

Effective employment of the principle of Clinical Equipoise by the clinicians participating in the trials will result in the recruitment of a larger number of patients with more homogenous baseline characteristics. Consequently, this will deliver more clearly interpretable and valid results that will help in standardizing the practice.

### Sham surgery

In placebo-controlled studies of medical interventions, a double-blinded design is always encouraged. Simply put, this means that neither the subject nor the investigator are aware of which study group a particular subject is placed in. In this

case, masking treatment assignment is generally considered ethically acceptable provided that the "shared ignorance" has been made clear in the consent process. However, when comparing two different surgical interventions, blinding physicians and patients are not possible. In sham surgery trials, only the patient is "blinded" to the treatment he/she receives. The clinician, who can distinguish between active and inactive treatment, may be required to engage in active deception.[38]

A common area in neurosurgery where the concept of sham surgery has been applied is in the management of Parkinson's disease (PD). Two famous randomized double-blinded controlled trials by Gross *et al.*[23] and Freed *et al.*[19] randomly assigned patients with severe PD to receive a transplant of embryonic nerve cells or sham surgery. Both studies showed no significant benefit overall in the control group compared to the sham surgery group. Based on these results, Polgar and Mohamed argued that the use of sham surgery for the evaluation of cellular therapies in PD is unnecessary and should, therefore, be deemed unethical.[44]

Since the aim of sham surgery is to ensure that both the investigator and the patient are blinded to the type of intervention, a possible solution is to consider using the prospective, randomized, open-label, and blinded-endpoint (PROBE) study design.[25] The PROBE design utilizes a strict randomization approach in which both investigators and patients are not blinded, but the results are adjudicated by an independent committee that is unaware of the treatment allocation, thus guaranteeing the unbiased comparison of therapies and evaluation of study results.[51]

### Blinding

The term blinding refers mainly to keeping investigators and trial participants unaware of the assigned intervention, in addition to outcomes assessors and/or analysts.[50] Blinding can be single, double, or triple with different researchers having different definitions to each term.

As mentioned previously, RCTs in neurosurgery have had issues with blinding with regard to quantity and quality of reporting. A systematic review of 82 neurosurgical RCTs by Martin *et al.* showed that most trials were open label (59.8%), and double-blinded trials were relatively rare (8.5%).[34] Since surgical trials use intervention with physical components, blinding can be complicated. Furthermore, some aspects of blinding are specific to surgical RCTs. Kiehna *et al.*[29] have suggested that improved awareness of the CONSORT guidelines[40] by all neurosurgery stakeholders may lead to improved trial design and reporting. Recommendations for planning, reporting, and assessing blinding in surgical trials were proposed by the Study Center of the German Surgical

Society.[45] This framework can serve as a simple guide for neurosurgeons when planning a blinded RCT.

### Surgeon expertise

The definition of expert and expertise can vary between specialties, even within the same surgical specialty. Expertise can be understood as the ability to consistently reproducible good performance involving a given procedure.[14]

Many neurosurgical procedures are not as stereotypical and as easily classified compared to other surgical specialties, given the significant variability in the surgical approach used to deal with a specific operative target. This lack of surgical intervention standardization among neurosurgeons has led to differences in their preferred approach based on their expertise, which could introduce expertise bias in trials that compare two different surgical approaches/techniques.[15]

Another problem with surgeons' expertise is the generalizability of the results. When a procedure is performed by a highly expert surgeon, the results may differ when comparing them to those of a surgeon with less exposure to a similar procedure (which constitutes the majority of practicing surgeons, especially outside tertiary care centers, where the volume of cases is expected to be less). An example is the Barrow Ruptured Aneurysm Trial which compared the safety and efficacy of microsurgical clipping and endovascular coil embolization for the treatment of acutely ruptured cerebral aneurysms.[36] In this trial, both of the compared procedures were performed by world-renowned, highly expert physicians in the field of cerebrovascular and endovascular neurosurgery. Regardless of the trial's results, its generalizability to the overall practice of the field is questionable.

A clear definition of expertise threshold is needed in neurosurgical RCTs. The North American Symptomatic Carotid Endarterectomy Trial (NASCET trial) has set a good example of defining a surgeon's expertise before participating in a trial.[16] To participate in NASCET, centers were required to demonstrate that their participating surgeons had a perioperative rate of stroke and death of 6% in a minimum of 50 consecutive cases accumulated over 2 years. The effect of expertise bias can also be minimized using expertise-based design when feasible, where participating surgeons provide only the intervention they are experts in.[11]

### Definition and Reporting of Outcomes

Outcomes (also called events or endpoints) are the variables monitored during a study to determine the impact of a given intervention or exposure on the health of a specific population.[17] The primary outcome is the variable that is most relevant to answering the research question. Ideally,

it should be patient centered (i.e., an outcome that matters to patients, such as quality of life and survival). Secondary outcomes are additional outcomes monitored to help interpret the results of the primary outcome.

Patient-reported outcomes (PROs) are defined by the U.S. Food and Drug Administration as "a report that comes directly from the patient about the status of a patient's health condition without amendment or interpretation of the patient's response by a clinician or anyone else."[53] Patient-reported outcome measures (PROMs), in turn, are either generic or disease-specific previously validated instruments used for reporting PROs.[4] PROMs are routinely used in spine surgery trials to assess pain and functional outcomes, as seen in the Spine Patient Outcomes Research Trial (SPORT trial), which compared surgical versus nonoperative treatment for lumbar disk herniation.[59] To measure their primary outcome, the authors used the Medical Outcomes Study 36-item Short-Form Health Survey (for bodily pain and physical function scales)[56] and the modified Oswestry Disability Index (American Academy of Orthopedic Surgeons MODEMS version).[7] PROMs have been also used in other neurosurgical subspecialties, such as endonasal skull base surgery and epilepsy surgery.[21,54] Reponen *et al.* encouraged the use of PROMs as tools in neurosurgical outcome reporting as data collected from such tools can help in developing validated neurosurgery-specific PROMs.[48]

Another method of reporting outcomes is in the form of a composite. The use of composite outcomes increases the event rate and reduces the sample size but harbors the risk that clinically less relevant but typically more frequent outcomes may drive the trial's main results or that the individual components move in different directions, thus generating uncertainty.[20,52] The medical management with or without interventional therapy for unruptured brain arteriovenous malformations trial (ARUBA trial) reported a composite outcome of death or symptomatic stroke.[41] The authors concluded that medical management alone is superior to medical management with interventional therapy for the prevention of death or stroke in patients with unruptured brain arteriovenous malformations followed up for 33 months. The ARUBA trial faced a lot of criticism from the neurosurgical community because the results were driven by the more frequent events of stroke compared to death, which is a more relevant outcome. The use of composite outcome is attractive for trial design as it reduces the number of subjects needed, but this needs to be carefully considered, and the component parts of the composite outcome measure need to be equally weighted and clinically relevant.

Finally, a careful description of the outcome measure, which includes key criteria used to adjudicate it, is a crucial factor to ensure the external validity of a successful RCT. Furthermore, the inclusion of PROs through the use of PROMs with clinical outcomes in research and clinical practice provides a more complete understanding of the impact of an intervention, therapy, and/or service on the patient. We advise following the recommendations of the 2010 CONSORT statement, which points out that all outcome measures, whether primary or secondary, should be identified and completely defined.[40]

**Pilot trials**

A pilot study is a small-scale, preliminary study that evaluates the feasibility, duration, cost, sampling strategy, and other research techniques before conducting a large, definitive clinical trial.[26] Pilot studies also provide researchers with preliminary data to gain insight into their proposed experiment's potential results. However, pilot studies should not be used to test hypotheses since the appropriate power and sample size are not calculated. Instead, pilot studies should be used to assess the feasibility of participant recruitment or study design.[30]

Pilot studies are very important to conduct before commencing large-scale surgical trials for many reasons, including testing the feasibility of performing a new procedure or administering an experimental therapy. As mentioned earlier, surgical trials usually face some unique challenges compared to medical trials, making the pilot study an essential stage in any research project to identify potential problems and deficiencies in the research instrument and protocol.

The neurosurgical literature has many examples of pilot studies that tested the feasibility of recruiting patients to conduct larger clinical trials. These examples are distributed among different subspecialties, such as neurotrauma, neuro-oncology, neurovascular, neuropediatric, and perioperative seizure management.[6,8,12,13,18,32,43] A critical question that we should ask is how well-done these neurosurgical pilot studies are and how many – of them – were translated to larger, definitive trials. Desai *et al.* published an interesting systematic review that examined the characteristics of pilot RCTs in the orthopedic surgery literature, looking to answer whether these pilot RCTs have led to definitive RCTs. Based on this systematic review, the authors concluded that the majority of published pilot RCTs did not lead to definitive trials.[10] Having a similar systematic review of pilot RCTs in the neurosurgical literature is vital to have to answer the question mentioned above. While pilot RCTs can provide much valuable information, they do have some limitations that authors should be aware of. One important limitation is that pilot studies are not hypothesis testing, so safety and efficacy cannot be evaluated.[31] Another limitation is the small sample size. Pilot studies are usually not powered to assess treatment effects.

In conclusion, despite their limitations, pilot trials remain extremely instructive and helpful, especially when planning

a large, multicenter trial in a common neurosurgical problem such as managing traumatic brain injury, subarachnoid hemorrhage, or spinal stenosis.

## TRIAL IMPLEMENTATION

### Funding

Funding is one of the key challenges to the success of any RCT. Obtaining adequate funding is more difficult in surgical trials compared to medical trials. A review published by Rangel *et al*.[47] looking at the trend of surgical research funding by the National Institutes of Health showed that proposals for surgical trials received less funding support relative to other nonsurgical proposals. To address this, many scientists have opted to obtain funding from industry. Khan *et al*. reviewed 110 RCTs published from 1981 to 2017 in three leading neurosurgical journals[28] and found that 36.4% (40 articles) stated that industry funding was provided. Of the RCTs in which industry sponsorship was present, 78% (31 out of 40) had a conclusion in favor of the new drug, device/implant, or surgical technique compared to 12.8% (nine out of 70) of RCTs without industry sponsorship. Azad *et al*. had similar findings after reviewing 401 RCTs published from 2003 to 2016 and reported that industry-supported trials (21.9%) were associated with substantial increases in the proportion of statistically significant trial outcomes.[3]

Several types of bias can be associated with industry-funded trials. Radcliff *et al*.[46] identified four types of bias present in cervical arthroplasty trials: publication bias (tendency to publish studies with positive results), external validity (the extent to which the study results can be applied to situations outside of study conditions), confounding bias (a distortion that modifies an association between an exposure and an outcome because a factor is independently associated with the both), and financial conflicts of interest. Similar types of bias can be applied to other subspecialties of neurosurgical research funded by the industry. To mitigate the influence on industry-funded trials, some RCTs opt for a mixed model in which both industry and nonprofit organization participate in funding the trial. In their review of RCTs published in the Lancet and New England Journal of Medicine between 2013 and 2015, Delgado *et al*. showed that for-profit-financed RCTs are associated with a higher odds ratio for a favorable outcome of a new treatment than nonprofit and mixed-funded RCTs.[9]

### Recruitment and trial discontinuation

RCTs require a sufficient number of participants to be adequately powered. This is necessary for the trial to answer the particular research question.[5] In general, surgical RCTs seem to have more difficulties recruiting patients compared to medical RCTs. Rosenthal *et al*. reviewed 863 RCTs and found that surgical trials were significantly more likely to be discontinued due to poor recruitment compared with medical trials.[49] Mouw *et al*. performed a more exhaustive review of 88,498 US trials and concluded that surgical trials are more likely to discontinue prematurely than nonsurgical trials.[42] The review also showed that poor recruitment is a major cause of early trial discontinuation for all trials and is more pronounced in surgical trials. Neurosurgical RCTs suffer from a similar waste of resources and ethical concerns surrounding trial discontinuation due to poor recruitment. Jamjoom *et al*. reviewed 64 neurosurgical RCTs registered on the website clinicaltrials.gov and reported that 17 (26.6%) trials were discontinued early, with the main cause being slow or insufficient patient recruitment (57%).[27]

## TRIAL ANALYSIS

### Intention-to-treat (ITT) versus per-protocol analysis

Randomization allows us to compare the trial arms for all measured and unmeasured characteristics if the sample size of the randomized patients is large enough. It also allows for proper causal inference. Over the course of a trial, many factors, including crossover or withdrawal/loss to follow-up, disrupt the randomization process, ultimately producing groups that are imbalanced compared to the original randomization. This form of noncompliance causes a loss of statistical power.[58] To minimize biases and preserve prognostically balanced groups, it is highly recommended that trial data be analyzed according to the ITT principle, in which participants are analyzed according to the group that they were allocated to regardless of the intervention that was eventually received. The counterpart of the ITT, the per-protocol analyses, limits the analysis to patients who completed the study in accordance with the protocol. However, in as-treated analyses, treatment allocation may be determined by patient characteristics or surgeon experience and be substantially more prone to bias. In a superiority trial, the ITT analysis is usually more conservative (biased toward the null) as the treatment effect is attenuated by crossover and protocol violations. In noninferiority or equivalence trials, the ITT is more likely to favor a positive result.[1] The practice of presenting the as-treated analyses, although in many cases serving to corroborate the ITT, may result in an uncomfortable situation of discordant findings. An example is the SPORT trial,[59] in which 39.7% of patients assigned to surgery crossed over to the conservative group, and 44.6% of patients assigned to conservative treatment ended up having surgery.[59] This high rate of crossover caused the results of the ITT and as-treated analyses to be inconsistent, leaving the main trial question unanswered.

Statistical methodology to mitigate the bias introduced by crossover and loss to follow-up includes instrumental

variable analysis, imputation techniques that account for potentially differential drop-out across study arms, marginal structural models, and other models to account for informative censoring. Applying the ITT principles yield an unbiased estimate of the efficacy of the intervention on the primary study outcome at the level of adherence observed in the trial. For instance, when the treatment under study is effective, but there is substantial nonadherence, the ITT analysis will underestimate the magnitude of the treatment effect that will occur in adherent patients. Although an underestimate of effective therapy, it will be unbiased.[35] This method of analysis results in a more accurate, unbiased estimate than that obtained from a per-protocol analysis.

### Learning curve effect

A learning curve (or experience curve) is a graphical representation that provides a visual assessment of the surgeon's skills with time and experience.[61] Through the learning curve for a new technique, the trial results may be affected by the degree of expertise acquired by the operator. During this time, a time-dependent form of confounding (the operator skills) may obscure the true results of the trial. An example of the learning curve effect can be evidently illustrated in the use of new devices in the endovascular treatment of intracranial aneurysms (IAs). The authors of the Analysis of Recanalization after Endovascular Treatment of Intracranial Aneurysm Study published an analysis of aneurysm characteristics, study population, and endovascular techniques used for the treatment of IAs.[22] One of the devices used in the study was the Woven EndoBridge (WEB) system which is considered a novel device. The authors reported that intrasaccular flow disruption (using the WEB device) was used in 6.9% of the unruptured IAs and only in 0.6% of ruptured IAs, and one of the reasons to explain this limited use is that "the learning curve in WEB application might be a limiting factor as well as a possible reluctance to use this novel device in a ruptured aneurysm."

Neurosurgical interventions are complex, which complicates their rigorous assessment through randomized clinical trials. Hierarchical models, time-dependent covariates, splines, and subgroup analyses by cumulative surgeon volume or period of recruitment can be applied to account for the learning curve. The use of concurrent controls is vital to account for the time effect.

## CONCLUSION

Evidence-based medicine is informed by hierarchical evidence, and this hierarchy informs clinical decision-making. RCTs form the basis of today's evidence-based approach to medicine and play an important role in guidelines development, as well as novel drugs and device

approval processes. RCTs in neurosurgery have dramatically improved over the past 60 years since the McKissock's group published their first trial in 1960. However, several issues led to the quality of neurosurgical RCT being suboptimal. Quality of blinding and reporting, bias, funding, recruitment, and learning curve are among the challenges investigators usually face when planning an RCT. In this contribution, we reviewed the history of neurosurgical trials, presented the most critical issues commonly encountered in designing, implementing, and analyzing RCTs in neurosurgery, and provided potential solutions.

### Declaration of patient consent

Patient's consent not required as there are no patients in this study.

### Conflicts of interest

There are no conflicts of interest.

## REFERENCES

1. Abraha I, Montedori A. Modified intention to treat reporting in randomised controlled trials: Systematic review. BMJ 2010;340:c2697.
2. Alderson P. Equipoise as a means of managing uncertainty: Personal, communal and proxy. J Med Ethics 1996;22:135-9.
3. Azad TD, Veeravagu A, Mittal V, Esparza R, Johnson E, Ioannidis JP, *et al.* Neurosurgical randomized controlled trials-distance travelled. Neurosurgery 2018;82:604-12.
4. Black N. Patient reported outcome measures could help transform healthcare. BMJ 2013;346:f167.
5. Chalmers I, Glasziou P, Godlee F. All trials must be registered and the results published. BMJ 2013;346:f105.
6. Cooper DJ, Rosenfeld JV, Murray L, Wolfe R, Ponsford J, Davies A, *et al.* Early decompressive craniectomy for patients with severe traumatic brain injury and refractory intracranial hypertension a pilot randomized trial. J Crit Care 2008;23:387-93.
7. Daltroy LH, Cats-Baril WL, Katz JN, Fossel AH, Liang MH. The North American spine society lumbar spine outcome assessment instrument: Reliability and validity tests. Spine (Phila Pa 1976) 1996;21:741-9.
8. Dayyani M, Mohammadi EM, Ashoorion V, Sadeghirad B, Yekta MJ, Grotta JC, *et al.* Aneurysmal subarachnoid haemorrhage-cerebral vasospasm and prophylactic ibuprofen: A randomised controlled pilot trial protocol. BMJ Open

2022;12:e058895.

9.  Delgado AF, Delgado AF. The association of funding source on effect size in randomized controlled trials: 2013-2015 a cross-sectional survey and meta-analysis. Trials 2017;18:125.

10. Desai B, Desai V, Shah S, Srinath A, Saleh A, Simunovic N, *et al.* Pilot randomized controlled trials in the orthopaedic surgery literature: A systematic review. BMC Musculoskelet Disord 2018;19:412.

11. Devereaux PJ, Bhandari M, Clarke M, Montori VM, Cook DJ, Yusuf S, *et al.* Need for expertise based randomised controlled trials. BMJ 2005;330:88.

12. Donovan EK, Greenspoon J, Schnarr KL, Whelan TJ, Wright JR, Hann C, *et al.* A pilot study of stereotactic boost for malignant epidural spinal cord compression: Clinical significance and initial dosimetric evaluation. Radiat Oncol 2020;15:267.

13. English SW, Fergusson D, Chassé M, Turgeon AF, Lauzier F, Griesdale D, *et al.* Aneurysmal subarachnoid hemorrhage-red blood cell transfusion and outcome (SAHaRA): A pilot randomised controlled trial protocol. BMJ Open 2016;6:e012623.

14. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Acad Med 2004;79 Suppl 10:S70-81.

15. Esene IN, Baeesa SS, Ammar A. Evidence-based neurosurgery. Basic concepts for the appraisal and application of scientific information to patient care (Part II). Neurosciences (Riyadh) 2016;21:197-206.

16. Ferguson GG, Eliasziw M, Barr HW, Clagett GP, Barnes RW, Wallace MC, *et al.* The North American symptomatic carotid endarterectomy trial: Surgical results in 1415 patients. Stroke 1999;30:1751-8.

17. Ferreira JC, Patino CM. Types of outcomes in clinical research. J Bras Pneumol 2017;43:5-5.

18. Frank JI, Schumm LP, Wroblewski K, Chyatte D, Rosengart AJ, Kordeck C, *et al.* Hemicraniectomy and durotomy upon deterioration from infarction-related swelling trial: Randomized pilot clinical trial. Stroke 2014;45:781-7.

19. Freed CR, Greene PE, Breeze RE, Tsai WY, DuMouchel W, Kao R, *et al.* Transplantation of embryonic dopamine neurons for severe Parkinson's disease. N Engl J Med 2001;344:710-9.

20. Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C. Composite outcomes in randomized trials: Greater precision but with greater uncertainty? JAMA 2003;289:2554-9.

21. Gallagher MJ, Durnford AJ, Wahab SS, Nair S, Rokade A, Mathad N. Patient-reported nasal morbidity following endoscopic endonasal skull base surgery. Br J Neurosurg 2014;28:622-5.

22. Gawlitza M, Soize S, Barbe C, le Clainche A, White P, Spelle L, *et al.* Aneurysm characteristics, study population, and endovascular techniques for the treatment of intracranial aneurysms in a large, prospective, multicenter cohort: Results of the analysis of recanalization after endovascular treatment of intracranial aneurysm Study. AJNR Am J Neuroradiol 2019;40:517-23.

23. Gross RE, Watts RL, Hauser RA, Bakay RA, Reichmann H, von Kummer R, *et al.* Intrastriatal transplantation of microcarrier-bound human retinal pigment epithelial cells versus sham

24. surgery in patients with advanced Parkinson's disease: A double-blind, randomised, controlled trial. Lancet Neurol 2011;10:509-19.

24. Haines SJ. Randomied clinical trials in neurosurgery. Neurosurgery 1983;12:259-64.

25. Hansson L, Hedner T, Dahlöf B. Prospective randomized open blinded end-point (PROBE) study. A novel design for intervention trials. Prospective randomized open blinded end-point. Blood Press 1992;1:113-9.

26. Hassan ZA, Schattner P, Mazza D. Doing a pilot study: Why is it essential? Malays Fam Physician 2006;1:70-3.

27. Jamjoom AA, Gane AB, Demetriades AK. Randomized controlled trials in neurosurgery: An observational analysis of trial discontinuation and publication outcome. J Neurosurg 2017;127:857-66.

28. Khan NR, Saad H, Oravec CS, Rossi N, Nguyen V, Venable GT, *et al.* A review of industry funding in randomized controlled trials published in the neurosurgical literature-the elephant in the room. Neurosurgery 2018;83:890-7.

29. Kiehna EN, Starke RM, Pouratian N, Dumont AS. Standards for reporting randomized controlled trials in neurosurgery. J Neurosurg 2011;114:280-5.

30. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: Recommendations for good practice. J Eval Clin Pract 2004;10:307-12.

31. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res 2011;45:626-9.

32. Lim DA, Tarapore P, Chang E, Burt M, Chakalian L, Barbaro N, *et al.* Safety and feasibility of switching from phenytoin to levetiracetam monotherapy for glioma-related seizure control following craniotomy: A randomized phase II pilot study. J Neurooncol 2009;93:349-54.

33. Mansouri A, Cooper B, Shin SM, Kondziolka D. Randomized controlled trials and neurosurgery: The ideal fit or should alternative methodologies be considered? J Neurosurg 2016;124:558-68.

34. Martin E, Muskens IS, Senders JT, DiRisio AC, Karhade AV, Zaidi HA, *et al.* Randomized controlled trials comparing surgery to non-operative management in neurosurgery: A systematic review. Acta Neurochir (Wien) 2019;161:627-34.

35. McCoy CE. Understanding the intention-to-treat principle in randomized controlled trials. West J Emerg Med 2017;18:1075-8.

36. McDougall CG, Spetzler RF, Zabramski JM, Partovi S, Hills NK, Nakaji P, *et al.* The barrow ruptured aneurysm trial. J Neurosurg 2012;116:135-44.

37. Mckissock W, Richardson A, Walsh L. Posterior-communicating aneurysms. A controlled trial of the conservative and surgical treatment of ruptured aneurysms of the internal carotid artery at or near the point of origin of the posterior communicating artery. Lancet 1960;275:1203-6.

38. Miller FG, Kaptchuk TJ. Sham procedures and the ethics of clinical trials. J R Soc Med 2004;97:576-8.

39. Misra S. Randomized double blind placebo control studies, the "gold standard" in intervention based studies. Indian J Sex Transm Dis AIDS 2012;33:131-4.

40. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, *et al.* CONSORT 2010 explanation and

elaboration: Updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c869.

41. Mohr JP, Parides MK, Stapf C, Moquete E, Moy CS, Overbey JR, *et al*. Medical management with or without interventional therapy for unruptured brain arteriovenous malformations (ARUBA): A multicentre, non-blinded, randomised trial. Lancet 2014;383:614-21.

42. Mouw TJ, Hong SW, Sarwar S, Fondaw AE, Walling AD, Al-Kasspooles M, *et al*. Discontinuation of surgical versus nonsurgical clinical trials: An analysis of 88,498 trials. J Surg Res 2018;227:151-7.

43. Patel SK, Kashyrina O, Duru S, Miyabe M, Lim FY, Peiro JL, *et al*. Comparison of two- and three-dimensional endoscopic visualization for fetal myelomeningocele repair: A pilot study using a fetoscopic surgical simulator. Childs Nerv Syst 2021;37:1613-21.

44. Polgar S, Mohamed S. Evidence-based evaluation of the ethics of sham surgery for Parkinson's disease. J Parkinsons Dis 2019;9:565-74.

45. Probst P, Zaschke S, Heger P, Harnoss JC, Hüttner FJ, Mihaljevic AL, *et al*. Evidence-based recommendations for blinding in surgical trials. Langenbecks Arch Surg 2019;404:273-84.

46. Radcliff K, Siburn S, Murphy H, Woods B, Qureshi S. Bias in cervical total disc replacement trials. Curr Rev Musculoskelet Med 2017;10:170-6.

47. Rangel SJ, Efron B, Moss RL. Recent trends in national institutes of health funding of surgical research. Ann Surg 2002;236:277-86.

48. Reponen E, Tuominen H, Hernesniemi J, Korja M. Patient-reported outcomes in elective cranial neurosurgery. World Neurosurg 2015;84:1845-51.

49. Rosenthal R, Kasenda B, Dell-Kuster S, von Elm E, You J, Blümle A, *et al*. Completion and publication rates of randomized controlled trials in surgery: An empirical study. Ann Surg 2015;262:68-73.

50. Schulz KF, Grimes DA. Blinding in randomised trials: Hiding who got what. Lancet 2002;359:696-700.

51. Smith DH, Neutel JM, Lacourcière Y, Kempthorne-Rawson J. Prospective, randomized, open-label, blinded-endpoint (PROBE) designed trials yield the same results as double-blind, placebo-controlled trials with respect to ABPM measurements. J Hypertens 2003;21:1291-8.

52. Stolker JM, Spertus JA, Cohen DJ, Jones PG, Jain KK, Bamberger E, *et al*. Rethinking composite end points in clinical trials: Insights from patients and trialists. Circulation 2014;130:1254-61.

53. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: Patient-reported outcome measures: Use in medical product development to support labeling claims: Draft guidance. Health Qual Life Outcomes 2006;4:79.

54. Van Gompel JJ, Marsh WR, Meyer FB, Worrell GA. Patient-assessed satisfaction and outcome after microsurgical resection of cavernomas causing epilepsy. Neurosurg Focus 2010;29:E16.

55. Vranos G, Tatsioni A, Polyzoidis K, Ioannidis JP. Randomized trials of neurosurgical interventions: A systematic appraisal. Neurosurgery 2004;55:18-25.

56. Ware JE Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473-83.

57. Weijer C, Shapiro SH, Glass KC, Enkin MW. For and against: Clinical equipoise and not the uncertainty principle is the moral underpinning of the randomised controlled trial. Br Med J 2000;321:756-8.

58. Weinstein GS, Levin B. Effect of crossover on the statistical power of randomized studies. Ann Thorac Surg 1989;48:490-5.

59. Weinstein JN, Lurie JD, Tosteson TD, Skinner JS, Hanscom B, Tosteson AN, *et al*. Surgical vs nonoperative treatment for lumbar disk herniation. JAMA 2006;296:2451.

60. Wickens AP. A History of the Brain: From Stone Age Surgery to Modern Neuroscience. Google Books. Available from: https://books.google.ca/books?id=piKcB QAAQBAJ&pg=PT59&lpg=PT59&dq=history+of+ neurosurgery+incas&redir_esc=y&hl=en#v= onepage&q&f= false [Last accessed on 2020 May 08].

61. Yelle LE. The learning curve: Historical review and comprehensive survey. Decis Sci 1979;10:302-28.