Randomised Clinical Trials in Surgery: A Look at the Ethical and Practical Issues

Anjan Kumar Das

Received: 25 October 2010 / Accepted: 31 December 2010 / Published online: 19 May 2011
© Association of Surgeons of India 2011

Abstract An ethically conducted randomised controlled trial (RCT) is the backbone of evidence based medicine. In surgical practice however, RCTs have taken a backseat, drawing much adverse comment. There are several reasons to explain surgeons' disinclination to conduct RCTs. These include many practical difficulties such as the problem of blinding surgical procedures, design and funding issues. There are also many ethical issues which need to be considered including the concept of equipoise as well as the ethical issues associated with sham surgery as a control. While there is no doubt that RCTs are essential and in fact have helped to weed out several unnecessary surgical procedures, it is important not to lose sight of the fact that they may not be always necessary in order to obtain evidence in favour of a procedure. Possible solutions could be to follow guidelines that have been issued by learned bodies and a strict adherence to all ethical norms that have been recommended in the conduct of trials

Keywords Evidence based medicine · Randomised controlled trials · Ethics

The Background

The Nuremberg trials after the Second World War brought into sharp focus the extent to which ethics and humanity could be jettisoned in the pursuit of so called

clinical research. This led to the famous Nuremberg trials where 16 physicians were convicted and sentenced to prison terms and the death sentence in some cases [1] and the enunciation of the Nuremberg Code that articulated a series of principles including voluntary informed consent, non-coercion and so on. Even after these horrific revelations, the entirely unbelievable human experiments done in Guatemala by American researchers proved that only a strict regulatory regime could prevent horrors that caused revulsion against the entire concept of clinical research. A landmark paper published by Beecher [2], the then Dorr Professor of Anesthesiology at the Harvard Medical School in 1966 recorded 22 examples of unethical human experiments, including a bizarre experiment where 26 babies less than 48 h old underwent vesicourethrography followed by multiple spot films to record reflux into the ureters. This was done ostensibly to answer the question whether urethral reflux occurs in the normal bladder!

The World Medical Association (WMA) was forced to take note and the Declaration of Helsinki was adopted by the World Medical Assembly, which for the first time laid out ethical principles for medical research using human subjects [3]. This declaration remains in force today, though it has undergone several modifications, the latest being at the 59th WMA General Assembly in Seoul in 2008. The basic principles that guide clinical research today are four; an informed consent, the primary duty of causing no harm, the principle of beneficence, i.e. of benefit to the patient and justice in treating the patients fairly. While the ethical parameters of medical research using drugs have been laid down and are fairly well understood and implemented, some question arise regarding surgical procedures.

A. K. Das (✉)
Clinical Research Leader, Stempeutics Research Malaysia,
Kuala Lumpur, Malaysia
e-mail: anjan.das@stempeutics.com.my

The Era of Evidence Based Practice

In the nineties of the last century, evidence based medicine (EBM) rapidly became the sine qua non of clinical therapy [4]. The idea was to use the best available evidence to make clinical judgments. Despite several logical objections to an uncritical acceptance of the tenets of EBM, it has today become ubiquitous in medical practice. EBM depends on several levels of evidence (Table 1), the highest of which are the randomized controlled trials (RCTs). The cat was set among the pigeons of surgical research by Horton who in an editorial in *Lancet* heavily criticized the surgical community for its non-acceptance of RCTs. In an article combatively entitled “Surgical Research or comic opera: questions, but few answers” Horton [5] questioned the paucity of RCTs in surgical literature showing that a mere 7% of original papers in a group of major surgical journals reported RCTs. The vast majority of the reports, he showed, were case series. This predictably led to a flurry of articles in the medical press acclaiming and criticizing his claims. However it was generally accepted that surgical research could and should use more rigorous statistical methods to be able to seriously refute these allegations.

The validity of these observations become more apparent when we look into history to consider the many procedures that were accepted as useful until a randomized trial disproved their efficacy. Procedures like the extra cranial-intracranial bypass operation for stroke prevention enjoyed an unmerited popularity for at least 15 years before an RCT showed that it increased the risk of stroke, [6] as did the internal mammary ligation procedure for ischemic heart disease [7]. In more recent years, several RCTs have disproved the notion that arthroscopy is beneficial for osteoarthritis of the knee [8].

This is particularly important in modern surgical practice when technology drives surgery. There is pressure from manufacturers, including extremely generous financial ones, to use devices and procedures which may on scrutiny turn out to be useless. In 1997 a trial disproved the then widely prevalent notion that surgical staplers reduced the incidence of intestinal leak after anastomosis. It was shown that the leakage levels were the same and stapler use led to an increased incidence of strictures! [9] The current (over) enthusiasm for minimally

invasive surgery should be tempered by the realization that several trials have failed to demonstrate much benefit for laparoscopic cholecystectomy over small incision surgery [10]. Slim and his colleagues analysed reports of published RCTs in laparoscopic surgery and showed that the majority of these trials were poorly designed, did not address appropriate end points and used inadequate sample sizes [11]. It is not inconceivable that later RCTs may throw up nasty surprises for surgeons who are doing many presently popular procedures without adequate evidence.

What are the Problems of Doing RCTS in Surgery?

Several observers pointed out [12, 13], quite logically, that there were several valid difficulties in doing RCTs for surgical procedures (Table 2). While a pharmaceutical intervention requires no physician skill, and the intervention i.e. the drug administration is easy and the same every time, no two surgical procedures are the same. The operation differs from patient-to-patient, and between surgeon-to-surgeon. It is humanly impossible to do exactly the same procedure every time.

It is also practically impossible to blind the procedure. As Johnson and Dixon have pointed out, [12] it was possible to compare the operations in the famous Leeds-York trial [14, 15] in the 1970s because the same incision was used every time. This landmark trial investigated three commonly performed surgical procedures for chronic duodenal ulcer. This is obviously not possible in, say, a comparison of open with laparoscopic cholecystectomy. It is also very difficult to blind the assessor who may by simply glancing at the incision be able to evaluate the nature of the intervention. Various unwieldy methods have been used in order to achieve this blinding including the use of iodine soaked bandages to conceal the incision and swearing the patient to secrecy! [12] However the fact remains that designing a good RCT for a surgical procedure is an exercise in ingenuity which is perhaps not always apparent to physicians. Surgical procedures also undergo incremental improvements and there is often a steep learning curve vitiating the results of any RCT.

Table 1 Medical therapy: classes of evidence

Class 1	High quality randomized controlled trials (RCTs).
Class 2	Prospective matched group cohort studies or randomized controlled trials lacking adequate randomization concealment or blinding or potentially liable to attrition or outcome ascertainment bias.
Class 3	Other studies such as natural history studies.
Class 4	Uncontrolled studies, case series, or expert opinion

Table 2 Problems in conducting surgical RCTs

1	Surgeon skill levels vary and there is a learning curve that may vitiate results.
2	Blinding of the procedure from the assessors is very difficult.
3	Designing a suitable study may be difficult.
4	Ethical issues including those related to equipoise and sham surgery have to be addressed.
5	Obtaining funding is often difficult.
6	Conflicts of interest have to be guarded against.

What are the Ethical Issues?

The Question of Equipoise

It is critical, when designing any clinical trial to ensure that there is equipoise to conduct it; in other words, is there a genuine uncertainty that needs to be cleared up regarding the merits of each arm of the treatment? Thus, clinical equipoise “refers to uncertainty among groups of experts who honestly disagree which treatment is better” [16]. In other words, any clinical trial should, if ethical, start out to try to test the null hypothesis, i.e. that one particular treatment is better than another. If the investigator feels that one of the treatments being offered is better than the other, he is ethically bound to offer that treatment. If, during the course of the trial, the equipoise is altered, i.e. it becomes obvious that one arm is better than the other, then the trial has to be ended [17].

One problem that is faced during a clinical trial and which has the potential of disturbing the equipoise is the strong preference that patients have towards a novel intervention, especially for a life-threatening or previously untreatable condition. This is particularly true in oncology trials where patients and their relatives are desperate to obtain experimental therapy and often refuse to continue the trial if placed in the placebo group. Particularly in a small trial, patient crossover to the treatment arm may vitiate the results of the trial [18]. This problem was graphically manifest in the early 1990s when an experimental procedure (bone marrow transplantation and high dose chemotherapy for breast cancer) was demanded by patients and patient advocate bodies before clinical trials were completed and it was only when 4 clinical trials failed to demonstrate benefit, did the craze for this procedure die down [19].

The Question of Sham Operations

Most surgical operations have evolved as contributions from single surgeons who developed a surgical procedure as a therapy for certain diseases. In fact the vast majority of the common general surgical procedures in use today were introduced in the first and second decade of the past century well before the advent of the present era of RCTs [20]. The

procedure was usually tested in a small group of patients by the deviser of the operation who then reported it to the surgical community at large and then this technique was then taken up by larger numbers of surgeons, and then it became a standard procedure. Alternatively, the procedure was found by the peer group to be useless and abandoned. The comparison was usually with their relative efficacy compared to older treatments. This approach was useful and did serve the surgical community well; however there were spectacular failures as some of the examples quoted earlier testify. One of them was, of course the internal mammary artery tie off for ischemic heart disease. It was shown by Cobb in 1959 that the then popular approach of tying off the internal mammary artery was essentially useless [7].

While doing a double blind RCT, it may be necessary to have a sham surgery arm. As an RCT is the gold standard for testing the efficacy of an intervention, it is often necessary to have a sham surgery arm to compare the efficacy of a surgical procedure. One of the most publicized such interventions took place for the treatment of Parkinson's disease with cell replacement [21].

Parkinson's disease is caused by the non-functioning of dopamine secreting neurons of the substantia nigra. It was proposed that this deficiency could be overcome by different cell replacement strategies. To prove this concept, it was necessary to have careful RCTs where the patients had to be assigned to sham operated groups or to the grafted groups. This was concealed from the patients as well as the physicians who followed up these patients. The sham operations raised a storm of ethical issues. One acute commentator, Miller, [22] was of the opinion that there is no ethical reason to prohibit sham surgery in surgical clinical trials. He wanted a careful assessment of the cost and benefit on a case-by-case basis. Polger and Ng [23] however showed by analyzing the evidence that there was no long term clinical improvement in either the sham operated or the treated group in some trials and in fact, this was evident even without analyzing the data from the sham surgery group. They thus suggested that the sham surgical group was unnecessary and that it was unethical to subject the patients to this surgery. One other issue that was often not really considered is that a patient entering a sham surgery group in a trial like this one for Parkinson's disease often commits himself to a follow-up period which may be

as long as 2 years during which time he is deprived of the opportunity to avail of newer medical treatments that may come into the market in the meanwhile.

However it is undeniable that for a surgeon, the concept of doing a sham operation and then allowing someone else to follow-up the patient is a change in culture to which it is difficult to adapt. It is also essential that the risks of the sham surgery be minimal. In the best of circumstances the patients who undergo sham surgery have to willy-nilly be faced with the possible complications of anesthesia and infection at the very least. Miller and Emanuel of the National Institutes of Health, USA proposed a middle ground to determine the ethics of these studies which included a sham surgery [24]. They emphasized that the patients in the placebo group must not, under any circumstances, be exposed to greater risks than the treated group and should not undergo more discomfort or danger. They also pointed out several key factors that could justify the use of sham surgeries, including high placebo response rate, a waxing and waning course of the disease, frequent remissions of the disease and an only partially effective existing treatment. Also as Wolf and Buckwalter [25] pointed out, the societal benefits of a trial using sham surgery may be substantial. They have shown that the Moseley study [8] for arthroscopy in OA (which was shown to be useless) saved a large amount of money at a saving of approximately \$ 5000 a procedure in the United States. The Council for Judicial and Ethical Affairs (CJEA) of the American Medical Association report which went into question of sham surgery made several important recommendations, [26] They recommended among other things, that sham surgery should be used only if there is no other trial design that will generate the required data, the informed consent process should be rigorous, and pointed out that sham operations are unjustified to test minor modifications of existing procedures.

Informed Consent

The question of informed consent is very important in surgical clinical trials, especially when enrolling patients for sham surgery. It has been reported that patients are often under a misconception that their intervention will be therapeutic, despite repeated counseling regarding the possibility that they may be getting a sham surgery. This was true in the Parkinson's disease trial mentioned earlier as well as in several other trials of this nature. The CJEA report laid much emphasis on the need to carefully explain the risks along with a detailed description of the differences between the trial and placebo arms. It also emphasized the need for additional safeguards in the informed consent process including the use of a neutral third party to get the consent or having a consent monitor to oversee the entire process.

Design of the Study

The most important factor that determines the size, duration, expense and outcome of the study is the choice of the clinical endpoints. It is essential to choose meaningful endpoints which will impact clinical decision making. In a recent paper, Philippe Menasche, [27] a pioneer in the field of stem cell use in cardiology pointed out that it is essential to have clinically meaningful endpoints which can help to change decisions regarding clinical management. In the case of surgical trials, mortality is usually the critical endpoint. However with more and more surgical interventions for chronic diseases, other parameters like the quality of life (QOL) indices are assuming a greater importance.

It is also a characteristic of surgical trials as opposed to pharmaceutical ones, that the target population is usually much smaller which often makes it difficult for researchers to recruit sufficient subjects in a given time frame [28]. Thus it is necessary to have smaller sample sizes which dictate the necessity to have composite endpoints. Composite endpoints combine several endpoints (e.g. survival and an important adverse event) thus making it possible to achieve statistical significance with a smaller sample number. There are problems, however with the use of composite endpoints. It might be difficult to interpret the data meaningfully, as the data for each separate endpoint may not tally with that for the composite endpoint. This has been illustrated by the SYNTAX trial, where the interpretation of the data has led to prolonged wrangling between cardiologists and cardiac surgeons! [29].

However sometimes even these methods are not sufficient to enroll sufficient participants. In such cases it may be necessary to accept a higher P value to establish significance thereby trading statistical precision to ensure a randomized design for the trial [30].

Other Problems with Surgical Trials

There are at least two other important issues that have to be faced in designing RCTs for surgical procedures. One is the learning curve phenomenon. It is obvious that any new surgical procedure will have a learning curve and often this can be extremely steep. This can cause a sea change in results when early results are compared with later ones [31]. Another important challenge is the incremental changes that take place in the surgical procedure during the course of the trial. This may be due to changes in the technique of surgery or due to changes in external factors such as the use of different antimicrobials or the use of laminar flow operating rooms.

It is also undeniable that all surgical teams are not equally skilled. If there is a large variation in the skill levels of the groups doing the surgery, the results may be vitiated

and difficult to interpret. The results of a trial done in large volume centre may not be relevant to a small volume centre with a substantial difference in complication rates.

Conflicts of Interest

A conflict of interest (COI) is defined as “a set of conditions in which the professional judgment concerning a primary interest tends to be influenced by secondary interests” [32]. These secondary interests are usually a pecuniary interest in the device, or the technology that is being tested. As more and more industry funding comes into surgical research, the potential for COI also increases substantially. It has been shown that in one case more than half the doctors who participated in a trial evaluating artificial intervertebral discs had financial interests in the product [33]. The industry can influence trials in many ways; these include huge consultant fees, royalties, stock options and so on. It is essential that COIs should not influence the results of research. In the developed world several safeguards have been put in place, though their effectiveness may be a matter of debate. In India, unfortunately, there is practically no restraint on Industry funding for researchers and little control on unethical incentives.

Funding

Funding for clinical research is also a problem. While industry funding has obvious disadvantages, it is very difficult to obtain funding from grant committees to fund surgical procedures and follow-up clinics. The cost of setting up a well designed randomised clinical trial for surgical procedures may be a major constraint that would deter prospective investigators [12].

What are the Possible Solutions for These Problems?

It has been proposed [13] that a framework be set up to mitigate these issues. It has been suggested that the data collection from surgical trials should be audited carefully by quality control measures. It is possible to also ensure a continuous performance evaluation using instruments that have been developed like CUSUM, CRAM or VLAD plots [34–36]. It is also necessary to check whether RCTs are really required to answer the question being asked or whether meta analyses of non-randomized evidence would suffice.

An independent group of statisticians, epidemiologists, clinical trial experts and biomedical editors published the consolidated standards of reporting trials (CONSORT) statement in 1996 [37]. This statement was immediately supported by a large number of organizations and has now

become standard in reporting RCTs. Unfortunately surgical journals have been slow in adopting these recommendations. The idea of this exercise, which includes a checklist and a flow diagram to show the progress of the subjects in the trial, is to increase the transparency and expose flaws and inadequacies in the reporting process. It is also necessary to register a trial in a national clinical trial registry in order to be able to publish in major biomedical journals.

Flow Diagram and Checklist not Provided

Conclusion

The surgical community as a whole has been negligent in implementing the best evidence practices for therapeutic interventions. This is not entirely purposeful. There are various constraints that make it difficult and sometimes impossible to conduct RCTs to validate many surgical treatments. However, it is essential that surgeons use the tenets of EBMs as much as possible in order to validate their treatment protocols. It is also essential to utilize the services of statisticians, epidemiologists and other experts in order to design proper studies. Whenever practical, RCTs should also be conducted in order to provide the highest level of evidence keeping in mind the constraints and adhering to the highest standards of ethical and scientific values. This is essential to weed out unscientific interventions that can undermine the confidence of the patient community as well as our peers in surgical practice.

References

1. <http://www.law.umkc.edu/faculty/projects/ftrials/nuremberg/NurembergDoctorTrial.html> (accessed 22 October, 2010)
2. Beecher HK (1966) Ethics and clinical research. *New Eng J Med* 274:1354–1360
3. Rickham PP (1964) Human experimentation. code of ethics of the world medical association, declaration of Helsinki. *Br Med J* 2 (5402):177
4. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS (2007) Evidence based medicine: what it is and what it isn't. *Clin Orthop Relat Res* 455:3–5
5. Horton R (1996) Surgical research or comic opera: questions, but few answers. *Lancet* 347(9007):984–985
6. The EC/IC Bypass Study Group (1985) Failure of extracranial-intracranial arterial bypass to reduce the risk of ischemic stroke. Results of an international randomized trial. *N Engl J Med* 313:1191–1200
7. Cobb LA, Thomas GI, Dillard DH, Merndino KA, Bruce RA (1959) An evaluation of internal mammary-artery ligation by a double-blind technic. *N Engl J Med* 260:1115–1118
8. Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH et al (2002) A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 347:81–88

9. Law S, Fok M, Chu KM, Wong J (1997) Comparison of hand-sewn and stapled esophagogastric anastomosis after esophageal resection for cancer. A prospective randomized controlled trial. *Ann Surg* 226:169–173
10. Majeed AW, Troy G, Nicholls P, Smythe A, Reed MW, Stoddard CJ, Peacock J, Johnson AG (1996) Randomized prospective single blind comparison of laparoscopic versus small incision cholecystectomy. *Lancet* 347:989–994
11. Slim K, Bousquet J, Kwiatkowski F, Pezet D, Chipponi J (1997) Analysis of randomized controlled trials in laparoscopic surgery. *BrJ Surg* 84:610–614
12. Johnson AJ, Dixon JM (1997) Removing bias in surgical trials. *BMJ* 314(7085):916–917
13. McCulloch P, Taylor I, Sasako M, Lovett B, Griffin D (2002) Randomised trials in surgery: problems and possible solutions. *BMJ* 324:1448–1451
14. Goligher JC, Pulvertaft CN, De Bombal FT, Conyers JH, Duthie HL, Feather DB et al (1968) Five to eight year results of Leeds/York controlled trial of elective surgery for duodenal ulcer. *BMJ* 2:781–787
15. Goligher JC, Feather DB, Hall R, Hal RA, Hopton D, Kenny TE et al (1979) Several standard elective operations for duodenal ulcer: ten to sixteen year clinical results. *Ann Surg* 189:18–24
16. Djulbegovic B (2009) The paradox of equipoise: the principle that drives and limits therapeutic discoveries in clinical research. *Cancer Control* 16(4):342–347
17. Freedman B (2006) Equipoise and the ethics of clinical research. In: Kuhse H, Singer P (eds) *Bioethics: an anthology*, 2nd edn. Blackwell Publishers, Malden
18. Gelijns AC, Ascheim DD, Parides MK, Kent KC, Moskowitz AJ (2009) Randomized trials in surgery. *Surgery* 145(6):581–587
19. Rettig RA, Jacobsen PD, Farquhar CM, Aubry WM (2007) *False hope: bone marrow transplantation for breast cancer*. Oxford University Press, New York
20. McCulloch P (1999) Evidence based surgery: do we need it, and can we get it? <http://ktclearinghouse.ca/cebmsyllabi/gensurgery/intro> (accessed 22/10/10)
21. Polgar S (2006) Evidence based methodology for advancing neural reconstruction. In: Sanberg CD, Sanberg PR (eds) *Cell therapy, stem cells and brain repair*. Humana Press, Totowa, pp 325–339
22. Miller FG (2003) Sham surgery: an ethical analysis. *Sci Eng Ethics* 10(1):157–166
23. Polgar S, Ng J (2005) Ethics, methodology and the use of placebo controls in surgical trials. *Brain Res Bull* 67:290–297
24. Emmanuel EJ, Miller FG (2001) The ethics of placebo-controlled trials—a middle ground. *New Engl J Med* 345:915–919
25. Wolf BR, Buckwalter JA (2006) Randomised surgical trials and sham surgery: relevance to modern orthopedic and minimally invasive surgery. *Iowa Orthop J* 26:107–111
26. Tenery R, Rakatansky H, Riddick Frank A, Goldrich M, Morse LJ, O'Bannon JM et al (2002) Surgical placebo controls. *Ann Surg* 235(2):303–307
27. Menasche P (2010) Cardiac cell therapy: Lessons from clinical trials: *J Mol Cell Cardiol* [Epub ahead of print]
28. Gejlsins AC, Aschleim DD, Parides MK, Kent KG, Moskowitz AJ (2009) Randomised trials in Surgery. *Surgery* 145(6):581–587
29. Serruys PW, Morice MC, Kappetein AP, Colombo A, Holmes DR, Mack MJ, Stähle E, Feldman TE, van den Brand M, Bass EJ, Van Dyck N, Leadley K, Dawkins KD, Mohr FW, Investigators SYNTAX (2009) Percutaneous coronary intervention versus coronary artery bypass grafting for severe coronary artery disease. *N Engl J Med* 360:961–972
30. Parides MK, Moskowitz AJ, Ascheim DD, Rose EA, Gelijns AC (2006) Progress versus precision: challenges in clinical trial design for left ventricular assist devices. *Ann Thorac Surg* 82:1140–1146
31. Park SJ, Gelijns AC, Moskowitz AJ, Frazier OH, Piccioni W, Raines E et al (2005) LVADs as destination therapy: a new look at survival. *J Thorac Cardiovasc Surg* 129:9–17
32. Thompson DF (1993) Understanding financial conflicts of interest. *N Engl J Med* 329:573–576
33. McDonald PJ, Kulkarni AV, Farrokhyar F, Bhandari M (2010) Ethical issues in surgical research. *Can J Surg* 53:133–136
34. Van Rij AM, McDonald JR, Pettigrew RA, Putterill MJ, Reddy CK, Wright JJ (1995) CUSUM as an aid to early assessment of the surgical trainee. *Br J Surg* 82:1500–1503
35. Poloniecki J, Valencia O, Littlejohns P (1998) Cumulative risk adjusted mortality chart for detecting changes in death rate: observational study of heart surgery. *BMJ* 316:1697–1700
36. Lovegrove J, Valencia O, Treasure T, Sherlaw-Johnson C, Gallivan S (1997) Monitoring the results of cardiac surgery by variable life-adjusted display. *Lancet* 350:1128–1130
37. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gøtzsche PC, Lang T. CONSORT GROUP (2001) The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 134(8):663–694