

# Clinical trials in cancer

**R. A. Badwe**

Tata Memorial Hospital, Tata Memorial Centre, Dr Ernest Borges Marg, Parel, Mumbai 400 012, India

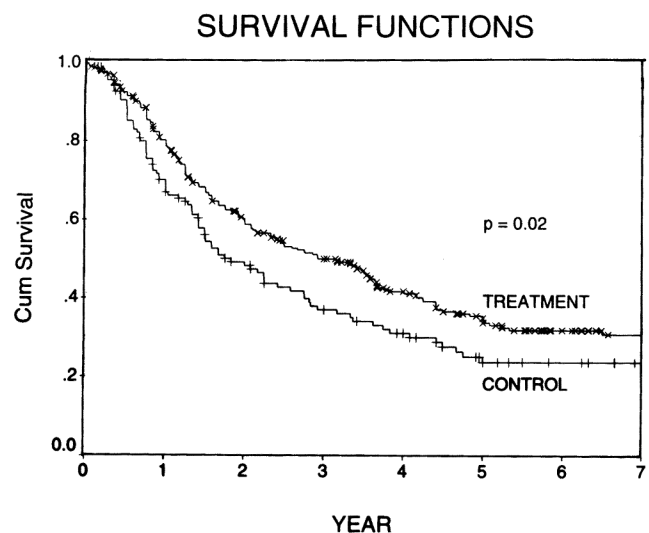
**Clinical trials are clinical experiments in medicine and the randomized clinical trial is a precision instrument that allows generation of the most reliable knowledge. Understanding the technology of clinical trials equips a clinician to run this precision experiment. But equally important is the imagination or intuition required to generate the ideas to be tested in an experiment. Furthermore, even greater imagination is necessary to correlate these multiple facts derived from trials to guess a simple, strange yet beautiful underlying generalization. This guess when tested within the tenets of randomized trials allows us to arrive at the truth that is capable of effecting a paradigm shift or a major leap in understanding disease. The article deals with the achievements of clinical trials, the technology involved in running it and importance of unbiased, uninhibited thought for generation of novel avenues to be addressed in clinical trials.**

THE middle of last century saw dramatic developments in medicine in the form of antibiotics and vaccines, which led to the saving of millions of lives. These achievements raised the hope that similar major developments will be repeated in other areas of medicine in the future. However, experience over the next 50 years has shown that this hope has been belied. The much-awaited breakthrough in major human diseases have not come about and the 'magic bullet' for the cure of cancer or HIV/AIDS has not materialized. In the absence of a major breakthrough, medical progress in preventing deaths from common and chronic diseases is going to be slow and will take place in small increments. Nevertheless, a small incremental progress in such common diseases as cancer, diabetes, coronary heart disease and AIDS may save many thousands of lives worldwide. The only scientific instrument, which has the capability to measure reliably small medical benefits, is the randomized-controlled trial (RCT) (Figure 1).

At an individual level, for patients as well as the treating doctor, the optimism of magic bullet still persists. A patient consulting an oncologist dreads the thought of being cursed with the diagnosis of cancer and expects the doctor to do something clever to put the things right. The doctor on his part strives to do the best based on his/her experience, training and updated evidence. An occasional

scientist hidden in the mind of the doctor wishes to strike with the magic bullet that would work miracles to cure the patient. In no other branch of medicine such optimism is evident as it is in oncology. The problem with such optimism is the ease with which one can accept unproven and/or ineffective interventions as magic bullets, e.g. increasing emphasis on high technology medicine adopted on the wave of enthusiasm and often in the absence of robust evidence of its efficacy. It is this enthusiasm coupled with market forces influencing health care that has sent the costs spiraling where one would expect small gains at huge cost.

The second important determinant of acceptance of intervention is the logic behind it. Many of our present day interventions are based on logic, which is very appealing but not always true unless backed by first class evidence. One of the glaring examples is screening for common cancers. It is logical that if stage I disease has better survival than stage III, early detection should save lives. Such simple logic is thrown out of the window by first class evidence from randomized trials of screening for lung<sup>1</sup>, prostate<sup>2</sup> and breast<sup>3</sup> cancer in young women. It is for this purpose that intervention with any degree of enthusiasm from the market forces and/or solid logical



**Figure 1.** Survival curves for two groups of patients (treatment vs control). Both the curves start at 1 (100%) and drops down, every time there is a death. At the seventh year, the survival rate for the control group is 24% and the same for the treatment group is 31%. The Logrank test is used to compare survival proportions in two groups ( $p = 0.02$ ).

basis should be rigorously tested in clinics before adopting as standard patient care. A randomized clinical trial is the most ruthless technique for testing hypothesis and has become the gold standard for the practice of evidence-based medicine.

*Nothing improves the performance of an intervention more than the lack of controls* and the second best are historical controls. Studies without controls and historical controls boast of fantastic results. Such extraordinary improvement is the effect of selection and has nothing to do with the intervention, e.g. overview of studies with historical controls for chemotherapy in oesophageal cancer<sup>4</sup> showed 68% ± 8 relative reduction (RR) in odds of death (Figure 2). For an improvement of that magnitude prospective randomized trial would be deemed unethical but fortunately 11 randomized trials were carried out over the same time period as the historical control studies. A meta-analysis of these RCTs<sup>4</sup> showed a non-significant 4.2% ± 24 RR in odds of death (Figure 3). If we were to adopt chemotherapy as the standard treatment for oesophageal cancer based on historical control data we would be no different than the Shamans of Kalahari desert offering traditional healing shrouded by mystery but delivered with intricate pageantry. On the other hand, historical controls may just be the only evidence of highly effective interventions but unfortunately such giant leaps are rare and are usually unplanned, e.g. antibiotics in pneumonia, Semmelweis' hand washing and Lister's antisepsis. The confidence in assessment of modest improvements by any controlled experiment is in fact dependent on the strength of its controls, which in the order of increasing reliability are historical, concomitant and randomized.

The tradition of randomized trials is quite old in oncology but was not very popular in early half of the last century. This was probably due to the dominance of surgeons in the field of oncology where their image of 'sometimes wrong but never in doubt' precluded them from testing uncertainties in treatment options. Some of the earliest

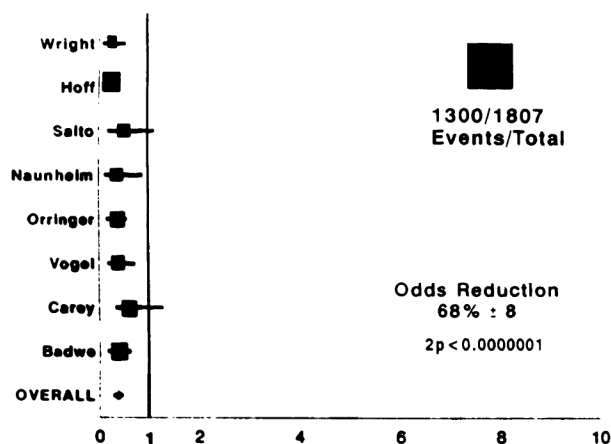
trials although have been surgical in treatment of cancer, e.g. hypophysectomy<sup>5</sup>, oophorectomy<sup>6</sup>. The most important randomized trials that have had a major impact on understanding the biology as well as clinical practice are those testing the magnitude of surgery in breast cancer. These trials have tested the hypothesis of more vs less surgery either at the primary site<sup>7</sup> (modified radical mastectomy vs breast conservation treatment) or in the axillary lymph node dissection<sup>8</sup>. The lack of difference in results of these trials strongly supported the deterministic model propounded by Fisher<sup>9</sup> replacing the century-old Halstedian model<sup>10</sup> of sequential spread. The deterministic model assuming the presence of micro-metastases at distant sites at the time of diagnosis paved the way for research and trials in early detection and adjuvant systemic therapy. Each of these biological insights has further added modest improvements in survival. A similar thought process has pervaded in research and treatment of other cancers where RCTs have shown small-to-moderate benefits of adjuvant systemic therapy, e.g. in colorectal cancer, soft tissue sarcoma.

Many other trials carried out with the aim of improving survival have sprung new insights into quality of life and organ preservation. The use of primary chemotherapy has allowed organ preservation in breast, larynx, urinary bladder and osteosarcoma without compromising on the long-term survival.

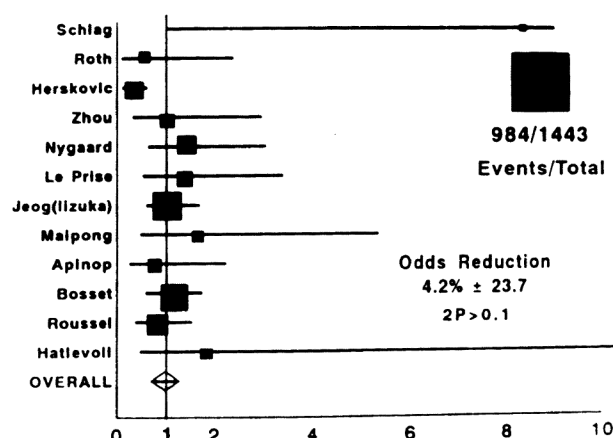
## Components of clinical trial

### Origin of ideas

The beginning of the most logical process of analysis, the RCT is illogical. The origin of idea is the most unknown part of RCT. In the words of Karl Popper<sup>11</sup>: *every discovery contains an 'irrational element' or a 'creative intuition'*; or in the words of Einstein, *There is no logical path*



**Figure 2.** Meta-analysis of randomized clinical trials for cisplatin-based adjuvant/neoadjuvant chemotherapy in esophageal carcinoma.  $\chi^2$  for heterogeneity = 17.80 ( $P = NS$ ).



**Figure 3.** Meta-analysis of historical control studies for cisplatin-based chemotherapy in esophageal carcinoma.  $\chi^2$  for heterogeneity = 8.6 ( $P = NS$ ).

leading to the universal laws from where the world can be deduced. They can be reached by intuition, based upon something like an intellectual love of the objects of experiences. Einstein's statement suggests existence of universal laws that would reduce multiplicity into unity. This is in direct contradiction to the present-day flow of ideas from reductionism. It is a matter of urgency that a paradigm shift be effected to generate new ideas. Suffice to state at this point of time that novel ideas are derived out of pure, unbiased experience, parallels drawn from other evolving sciences and uninhibited intuition.

### Development of technology

The process that deals with justification or validity of idea (Kant's *quid juris?*) revolves around logical analysis of scientific knowledge. Falsification is at the heart of such logical analysis and RCT has been one of the bold steps towards falsification. Over the last century, understanding of RCT technology has evolved to overcome its clinical, ethical and statistical shortcomings and has improved its ability to come closer to certainty.

Sackett<sup>12</sup> has proposed six pre-requisites for conducting a RCT.

- A trial needs to be done (existence of clinical uncertainty)
- The question posed is both appropriate and unambiguous
- Trial design is valid
- The inclusion/exclusion criteria strike a balance between testing efficacy and generalizability
- Trial protocol is clinically feasible
- Trial administration is experienced and effective.

An idea that is novel, has either some clinical or biological basis of efficacy and is addressing a common clinical problem should be first converted into an unambiguous question. This question should be addressed by an easily implementable clinical intervention. Every scientist knows a simple rule called Occam's razor<sup>13</sup>, 'faced with several interventions, prefer the simplest one'. An unspoken corollary to this rule is about which subject to build the hypothesis on and is called Occam's castle<sup>13</sup>: 'faced with several competing places to build new science on, chose the simple one'. I would take the liberty to replace 'simple' with 'common' in this unspoken corollary. A common clinical problem offers adequate number of patients in a short time for completion of accrual. Trials that take long time to complete run the risk of being overtaken by new developments in diagnosis or treatment that make their results redundant or irrelevant.

*Trial design* is the next vital step in RCT. A trial with a pragmatic design allows one to draw inference on efficacy as well as effectiveness. It is conventional to ask a single question with one or two primary endpoints in a RCT. A three-arm study is usually not an ideal design since it is

actually running two trials with same controls. The sample size calculation and interpretation of results becomes difficult in such trials. But a powerful design that has four arms is capable of handling two disparate ideas in one trial. A RCT with a factorial design can answer two distinct interventions, e.g. one anatomical and the second biological and may also suggest a possible additive interaction of the two interventions (Figure 4). In such trial design one can test the effect of intervention A (vertical comparison) and intervention B (horizontal comparison). The interaction of A + B, additive or otherwise, can be gauged by comparing combination cell with others albeit with lesser power. One should be careful in avoiding combination of two interventions that may have adverse interaction.

*Sample size* is the most important statistical consultation after trial design has been selected. The calculation of sample size involves assessment of baseline event rate for the endpoint of the trial. The next step is to make a conservative guess of expected benefit. This guess should be as conservative as possible to avoid a false negative result (type II error), however the magnitude of benefit should be perceived to be clinically important. The probability of detecting given size of benefit is called the *power* of study. Power is directly proportional to sample size and expected benefit. In addition, it is also influenced by study design, technique of analysis and significance level chosen by the investigator. Power could vary from 80 to 90%. Prior to this concept of trial size, the majority of the trials were small and inadequate to gauge the magnitude of benefit. Zelen<sup>14</sup> reported that the mean sample size of trials reported in *Cancer* was 50. Mosteller<sup>15</sup> *et al.* examined 285 trials in cancer and found the median trial size to be 9 for myeloma and chronic myeloid leukemia, 7 for gastrointestinal cancer and 25 per treatment group for breast cancer. Only 7% of trials had a sample size more than 200 patients. Two glaring examples of inadequate sample size leading to type II error in breast cancer are ovarian ablation trials and trials addressing axillary lymph node dissection. None of these trials singly showed a significant result but a meta-analysis of both groups of trials showed a significant beneficial impact of ovarian ablation<sup>16</sup> and axillary dissection on survival in breast can-

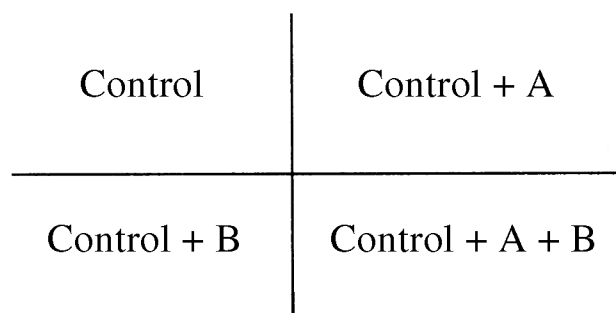


Figure 4. Factorial design of a clinical trial.

cer<sup>17</sup>. NSABP B-04 (ref. 8) addressing the issue of treating axilla or not had a large sample size compared to all earlier trials. But it was still insufficient to pick up benefits less than 10% in women with clinically negative axilla. If the technology as understood today were applied in seventies, Fisher's deterministic paradigm could not have replaced the time-tested Halstedian hypothesis. Trials of ovarian ablation being small were incapable of detecting modest benefit in adjuvant setting but a meta-analysis resurrected its efficacy and found it to be of the same magnitude as adjuvant chemotherapy<sup>16</sup>. Trials of medical castration further lend support to the efficacy of this intervention. A third and less rigorously tested question with inadequate sample size is the duration of post-operative adjuvant chemotherapy (3 or 6 or more cycles).

The other error that needs attention in planning RCT is type I error which is also termed as the level of *confidence* or possibility of results being false positive. This is usually fixed at 0.05, i.e. one in 20 chance of result being falsely positive.

### Randomization

This is the pivotal step in the success of a controlled experiment/trial. Randomization does not mean haphazard or alternate allocation to treatment arms but allocation outside the influence of treating surgeon/physician. Large sample size and stratification by known dependent variables further improves unbiased allocation to competing arms and increase the reliability of results. A central randomization with random number tables almost guarantees unbiased treatment allocation and avoids selection bias in endpoint assessment (ascertainment bias). But such procedures may not be possible in surgical trials and are of lesser importance if the endpoints are objective or clear-cut like pathologically proven recurrence or death due to disease.

### Other biases

Inappropriate handling of withdrawals, missing data and protocol violations may introduce bias in favour of either arms of randomization. It is mandatory that all trials are reported with intention to treat and trialists take adequate precautions to reduce the missing data to minimum.

### Ethics of clinical trials

There are three important aspects of ethics in clinical research. The first, *offering information* to the subject participating in the study. This information should include the aim of the study, background information, procedures (invasive or non-invasive) over and above the standard care that the patients in intervention arm are expected to go through, the known side effects, how they would be

treated and whom to contact when side effects are experienced. The participation should be entirely voluntary. The second is *obtaining informed consent* in language that patient can read and write about participating in study or use of tissue or blood. This should also include explaining to patient about random allocation to treatment in the best interest of the patient. The third and the most important aspect is the *secrecy of patient information* and its use only for the purpose put forth in the *assessment of endpoints* in relation to intervention and other dependent variables. A constant collaboration between institutional review board and clinician investigator would keep the patient's interest at the forefront. Such co-operation with pre-decided timings of interim analysis would allow both the parties to apply rules of premature termination of trials<sup>18</sup> with ease without jeopardizing the interest of the patients as well as the purpose of running the trial.

### Data collection and analysis

Precise and complete collection of data is vital in the success of any RCT. Assessment of all designated endpoints should be precise. The time, place and individuals collecting data should be identified and well informed about the logistics of data collection in relation to execution of protocol. A monthly meeting to assess the process of data collection and quality of data collected avoids unfortunate surprises of lack of data later on in the trial and inculcates discipline in the team running the study. The type and timing of statistical analysis is decided well in advance at the time of designing protocol. Interim analysis is carried out at pre-specified time of trial and decision on stopping trial<sup>18</sup> is based on higher level of confidence ( $p < 0.01$  or 0.001) depending upon the number of events at which the analysis is performed.

### Reporting a clinical trial

It is mandatory that all randomized trials be reported irrespective of the end result. The trialists owe this responsibility to the patients and society at large to report all experiments. There is an urgent need for a controlled trial registry that maintains the record of all trials running world over. In 1996 a group of epidemiologists, statisticians and journal editors drew a statement called CONSORT<sup>19</sup> (Consolidation of Standards of Reporting Trials) which forms an excellent guideline for reporting randomized trials. A detailed attention to all 21 items on the checklist of CONSORT would guarantee a comprehensive and factual report on any clinical experiment.

### Benefits and future of clinical trials

Over and above saving lives of patients suffering from common cancers, benefits of the randomized trial spawn

all those involved in providing health care. This includes the doctor who can have a meaningful dialogue with the patient based on scientific evidence from a randomized controlled trial rather than subjugating him to his own prejudices. The health care authorities can now decide upon standards of care based on scientific evidence and its cost and benefit in order to introduce appropriate and tested technologies and public health measures.

Future of clinical trials is in the hands of patients and doctors alike. For clinicians, it is as important to train clinical mind for the process of falsification, as it is to keep it open for grasping novel avenues. These two are disparate processes, one disciplined, legal almost stifling whereas the other intuitive, artistic. A scientific mind needs to be tempered to be almost schizophrenic, intuitive on one hand and at the same time detached, legal while testing the intuition. On the other hand, patients also should be made aware of participation in clinical trials. The best standard care that patients get today is the result of trials where patients have participated in yesteryears. Willingness of clinicians and patients would see an increase in the number of patients participating in trials and answering important questions at a quick pace.

Developments in RCT technology have had few, not so good fallouts. Large breakthroughs being rare to come, we are so geared up for looking for small or modest benefits that a single institution trial of 400–500 patients is today looked down upon. Mega-trials for small benefits are slowly becoming a norm and running such trials is gradually becoming institutionalized. There are corporate groups who would run trials for industry. This has reduced the nuance and beauty of scientific discovery. It may be ideal if centres of excellence that see large number of patient population in a given disease should run appropriate size trials for innovative ideas with anticipated benefits from 15–20%. If proven to be ineffective for that large magnitude of benefit then the trial can be carried over by large multi-centre groups for smaller benefits. Such dynamic co-operation between individual institutes and co-operative groups will allow rapid testing of novel ideas.

Large trials cost lot money and that could limit the size of study. It is necessary that health budget or insurance premium, or government departments dealing with research in science and technology should allocate funds to appropriate institutes running clinical/laboratory research. One more drawback of these large trials proving modest benefits of statistical significance is their impact on standards of practice. For example, should a hormone sensitive, operable

breast cancer patient receive adjuvant tamoxifen, chemotherapy and ovarian suppression (if pre-menopausal)? If in few more years, two other biological interventions prove to be of modest benefit, should those be added over and above today's standard care as per evidence? It will not be surprising then that women with less than 10% risk be on a treatment of a cocktail of interventions over a protracted period of time (10 years or more) at an exorbitant cost in years to come.

The clinical investigator should understand that the bottom line of running any controlled experiment is to find the truth and the process of experimentation is only a means of reaching this goal. The process should not be influenced by personal hunches, financial gains or else the very purpose of running a controlled experiment will be veiled by the mist of misconduct<sup>20</sup>.

1. Marcus, P. M. *et al.*, *J. Natl. Cancer Inst.*, vol. 92, pp. 1308–1316 (editorial, pp. 1280–1282).
2. Labrie, F. *et al.*, *Prostate*, 1999, **38**, 83–91.
3. Elwood, J. M., Cox, B. and Richardson, A. K., *Online J. Curr. Clin. Trials*, 1993, document 32.
4. Bhansali, M. S., Patil, P. K., Badwe, R. A., Havaldar, R., Desai, P. B., *Dis. Esophagus*, 1997, **10**, 51–54.
5. Harvey, H. A. *et al.*, *Cancer*, 1979, **43**, 2207–2214.
6. Nissen–Meyer Casteration as part of the primary treatment for operable female breast cancer, *Acta Radiologica Suppl.*, 1965, p. 249
7. Fisher, E. R., Anderson, S., Redmond, C. and Fisher, B., *Cancer*, 1993, **71**, 2507–2514.
8. Fisher, B., Wolmark, N., Redmond, C., Deutsch, M. and Fisher, E. R., *Cancer*, 1981, **48**, 1863–1872.
9. Fisher, B., *Cancer*, 1977, **40**, 574–587.
10. Halsted, W. S., *Ann. Surg.*, 1907, **46**, 1–19.
11. Popper, K. R., *Logic of Scientific Discovery*, Unwin Himan, London, 1980.
12. Sacket, D. L., in *Clinical Trials* (eds Shapiro, S. H. and Louis, T. A.), Marcel Dekker, New York, 1983, p. 65.
13. Weiner J., *Time, Love and Memory*, Random House Inc., NY, 1999, p. 68.
14. Zelen, M., Gehan, E. and Gibwell, O., in *Importance of Cooperative Groups* (ed. Hoogstralen, B.), Mason, New York, 1980, pp. 291–312.
15. Mosteller, F., Gilbert, J. P. and McPeck, B., *Controlled Clin. Trials*, 1980, **1**, 37–58
16. Early Breast Cancer Trialist's Collaborative Group (EBCTCG), *Results of the Ovarian Ablation, Treatment of Early Breast Cancer*, Oxford University Press, 1990, vol. 1, pp. 82–83.
17. Orr, R. K., *Ann. Surg. Oncology*, 1999, **6**, 109–116.
18. Pocock, S. J., *Biometrics*, 1982, **38**, 153–162.
19. Begg, C. *et al.*, *JAMA*, 1996, **276**, 637–639.
20. ASCO News Release: Statement of the ASCO on potential misconduct in South African Trial of High Dose Chemotherapy, 4 February 2000.