The Future of Clinical Research Beyond Phase III Trials

Michael A. Vogelbaum, M.D., Ph.D., F.A.C.S.

It is well accepted that prospectively designed, randomized, and appropriately powered clinical trials serve as the gold standard for generating medical evidence in support of the efficacy of a treatment. Successful completion of these so-called phase III trials, however, often can be challenging. Two major issues that often are impediments to implementation and/or completion of phase III trials are the large patient accrual requirements and the need for a control treatment arm. Alternative trial designs may be available to help overcome these hurdles and yet still provide high-quality clinical evidence for evaluation of the efficacy of a therapeutic approach.

Phase III trials often require hundreds, if not thousands, of patients to provide valid statistical estimates of how effective a treatment will be in the more general population of patients who may subsequently receive treatment outside of a clinical trial. The requirement for large patient numbers is based on the frequency of events (defined as clinical changes that are counted as part of the primary endpoint of the trial, e.g., death) that is expected to occur over the course of the trial and the magnitude of change that is expected from the experimental treatment. A low frequency of events combined with a relatively modest expected change in outcome will result in a much higher patient accrual requirement than would an expectation for a larger magnitude of benefit. Of course, if the expectation for benefit is set too high, a smaller but clinically significant benefit may be missed because of underpowering of the trial. In some cases, the target population may be so small as to create a large hurdle for adequately powering any trial (e.g., the population of patients with 1p/19q codeleted anaplastic oligodendrogliomas or those with amyotrophic lateral sclerosis). In surgical trials, the high cost of therapy associated with treatment of hundreds of patients may become prohibitive.

A second issue is that conventional phase III trials are designed to compare a novel treatment or treatment strategy with a control arm, which may involve active treatment or administration of a placebo. Often, there may be ethical concerns regarding the use of a placebo control, particularly in cancer trials. Active control arms can also be problematic because standards of therapy and even expected outcomes can change over the life of a trial. In the case of rapidly changing technology (e.g., intravascular stents and coils), the available technology may change so rapidly as to outdate the treatment in question before the trial has ended.

Even when there is an adequate population to study and a lack of ethical concerns about the control treatment, there may be a lack of equipoise that prevents successful trial accrual. Equipoise describes the relative lack of bias that exists regarding multiple treatment options and may be used to describe either individuals or communities of physicians. In other words, when equipoise is not present, there are strong biases in favor of one treatment over another, and patients may not be successfully accrued into a trial that randomizes them to one or the other treatment.[6] For example, a lack of equipoise regarding the use of surgery or radiosurgery for the primary treatment of a single brain metastasis is a factor that has prevented the successful completion of a fully randomized clinical trial evaluating these two options. Availability of therapies off-label or off-study also creates hurdles to the successful accrual of trials evaluating their efficacy in a different disease.

The difficulties associated with the design and completion of conventional phase III trials are illustrated by an article by Smith and Pell[8] that examined the evidence supporting the use of parachutes to prevent death and injury associated with skydiving. The authors performed a systematic review of the literature and found that there were no published results from randomized clinical trials supporting the use of parachutes. They further concluded, tongue in cheek, that, given the strong arguments in favor of evidence-based medicine, the most ardent supporters of the need for randomized trials should participate in a "double blind, randomized, placebo controlled crossover trial of the parachute." Are there alternative strategies available for generating high-quality medical evidence when a prospective, randomized phase III trial cannot be performed? There are, in fact, alternative trial design strategies that have been developed specifically to overcome the obstacles to high-quality prospective clinical research discussed above.

One alternative is the so-called Zelen design. This approach involves the use of randomization before patient consent; the clinician presents only the randomly designated treatment to the patient instead of presenting both treatments and the concept

0148-703/09/5601-0037

of randomization. The use of prerandomization raises serious ethical concerns regarding informed consent because the patient consents only for the selected treatment and not for the randomization process. Although this approach may be useful for addressing issues relating to equipoise, it is weakened by high crossover rates.[1] One example of the utility of the Zelen design comes from a pediatric traumatic brain injury trial that randomized patients to medical versus surgical management of their head injuries. Informed consent was obtained only for the surgical procedures. This approach eliminated the time interval to treatment that would be associated with a full discussion of the randomization process and the risks/benefits of both medical and surgical approaches, time that in itself could have worsened patient outcomes.[9]

Another alternative is the noninferiority design. This design may also help to overcome clinicians' lack of equipoise because the goal is to show that two treatment options are sufficiently similar as opposed to one being better than the other. Unfortunately, noninferiority designs typically require even larger patient accrual than conventional phase III trials. Additionally, entry into a noninferiority trial may be difficult to sell to patients, particularly when the outcomes are poor regardless of treatment. A critical point to understand is that unplanned noninferiority analyses, in the setting of a conventional phase III trial, are rarely considered statistically valid unless prospectively specified in the trial design.[5]

A third alternative trial design covered here is referred to as an adaptive, or bayesian, trial design. A key feature of adaptive trials is that accumulating data are routinely accessed during the trial to alter randomization patterns, reduce or expand accrual, drop or add treatment arms, or change the eligibility criteria. This approach is distinctly different from conventional designs, which strictly provide very limited access to accumulating data. To understand the differences between conventional and adaptive trial designs, it is necessary to understand the differences between the frequentist approach and statistical inference.

Conventional phase III trials are based on a frequentist design. This design assumes that what is being measured has a fixed value and that the experimental variability lies only in the observations made in the clinical trial. For example, in a simple coin toss trial, one knows that there is a fixed probability of 0.5 that the coin will land as a "heads" with any given toss. Any sample-to-sample deviation from 0.5 is a result of random variation, and this variability is overcome by a sufficient sample size (i.e., number of tosses). Based on the frequentist design, with an unlimited number of observations (e.g., patients), one can reliably determine the effect (e.g., survival rate) associated with a new therapy. The power of a frequentist trial depends on how reliable each observation is and hence how many observations need to be made to come up with an acceptable measure of the efficacy of treatment. Shortcomings of the frequentist design include the requirement for the number of observations to be defined at the trial outset, the fact that early evaluation of results requires an increase in the total number of observations (i.e., patients), and that one must explicitly plan for increases or reductions in accrual at the trial outset to maintain statistical validity.[3,4]

In contrast, the principle of statistical inference is based on the assumption that what is being measured does not have a fixed probability. Instead, all unknowns are thought to have distributions of probability, and the goal of a bayesian design is to continuously reevaluate probability distributions over the life of a trial. Hence, as patient accrual occurs, the need for additional observations is continually reassessed and may be increased or decreased as needed. A bayesian design can be used prospectively or retrospectively and is considered to be an ideal design for the use of biomarker results to refine arms as a trial is progressing. One example of a retrospectively applied bayesian analysis involved the reanalysis of data from five completed trials of pravastatin with or without aspirin that did not randomize aspirin use. Application of statistical inference showed that the combination was synergistic in terms of preventing cardiac events.[7]

Although attractive in concept, bayesian trial designs face their own challenges. The concept of statistical inference can be challenging for most clinical investigators, and in most cases, these trials require significant computational support. Furthermore, although the use of statistical inference has led to drug approval in a limited number of cases, broader US

TABLE 6.1. Alternative designs for prospective phase III clinical trials

Design	Advantages	Disadvantages
Zelen (prerandomization)	Can overcome lack of equipoise	Ethical concerns
	Simpler consenting process	Weakened by high crossover rate
Noninferiority	Can overcome lack of equipoise	Higher accrual requirements reduce feasibility
Bayesian	May require fewer patients to answer clinical question	Conceptually challenging
	Can be performed prospectively or retrospectively	Requires significant computational support
	Can add arms or new analyses to an existing trial	Limited track record with U.S. Food and Drug Administration approval process

Food and Drug Administration acceptance of this approach will be required before more widespread use of this design.[2]

In summary, results from conventional phase III prospective, randomized designs remain the gold standard of clinical evidence. In circumstances in which these trials are not feasible, alternative trial design strategies should be considered (see *Table 6.1* for a summary).

Disclosure

The author has received honoraria and/or consulting fess from: Schering-Plough, Genentech, Eisai, Neopharm and BrainLAB. He is also co-Chair of the Brain Committee of the Radiation Therapy Oncology Group, which receives its funding from the National Cancer Institute of the National Institutes of Health.

REFERENCES

1. Adamson J, Cockayne S, Puffer S, Torgerson DJ: Review of randomized trials using the post-randomised consent (Zelen's) design. **Contemp Clin Trials** 27:305–319, 2006.
2. Bauer P, Brannath W: The advantages and disadvantages of adaptive designs for clinical trials. **Drug Discov Today** 9:351–357, 2004.
3. Berry DA: Introduction to Bayesian methods III: Use and interpretation of Bayesian tools in design and analysis. **Clin Trials** 2:295–300, 2005.
4. Berry DA: Bayesian clinical trials. **Nat Rev Drug Discov** 5:27–36, 2006.
5. Dignam JJ: Early viewing of noninferiority trials in progress. **J Clin Oncol** 23:5461–5463, 2005.
6. Ghogawala Z, Barker FG 2nd, Carter BS: Clinical equipoise and the surgical randomized trial. **Neurosurgery** 62:N9–N10, 2008.
7. Hennekens CH, Sacks FM, Tonkin A, Jukema JW, Byington RP, Pitt B, Berry DA, Berry SM, Ford NF, Walker AJ, Natarajan K, Sheng-Lin C, Fiedorek FT, Belder R: Additive benefits of pravastatin and aspirin to decrease risks of cardiovascular disease: Randomized and observational comparisons of secondary prevention trials and their meta-analyses. **Arch Intern Med** 164:40–44, 2004.
8. Smith GCS, Pell JP: Parachute use to prevent death and major trauma related to gravitational challenge: Systematic review of randomized controlled trials. **Br Med J** 327:1459–1461, 2003.
9. Taylor A, Butt W, Rosenfeld J, Shann F, Ditchfield M, Lewis E, Klug G, Wallace D, Henning R, Tibballs J: A randomized trial of very early decompressive craniectomy in children with traumatic brain injury and sustained intracranial hypertension. **Childs Nerv Syst** 17:154–162, 2001.

