

## Randomization in Phase II Clinical Trials

### PRO

---

Greg Yothers, PhD, and H. Samuel Wieand, PhD  
NSABP Biostatistical Center  
and University of Pittsburgh  
Department of Biostatistics  
Pittsburgh, Penn.

Is there a place for randomization in phase II clinical trials? Randomized phase II clinical trials in cancer research can generally be separated into two categories.<sup>1</sup> The first category randomizes patients between a control arm and (usually) one experimental arm and the second category randomizes patients between two or more experimental arms. The first category can be subdivided into trials where there is only an informal comparison between the control and experimental arms and trials where there is a formal statistical comparison. Rubinstein and colleagues<sup>2</sup> review design issues with randomized phase II trials, comment on the merits of trials including an informal comparison, and propose a family of randomized phase II designs including a formal statistical comparison with the randomized control. We weigh the risks, costs, and benefits of each category of trial as compared to the more traditional approach of a single-arm phase II clinical trial<sup>3</sup> and find that, in certain instances, the benefits outweigh the costs and risks.

#### Randomization Between Experimental and Control Arms

##### *Informal Comparison*

When no formal statistical comparison is made between the experimental and control arms, the required sample size for this design is roughly double what would be required for a traditional single-arm trial. In this design, the experimental arm is compared to a historical control using the same criteria as in a traditional single-arm phase II clinical trial. The experimental arm is not formally compared to the randomized control arm. However, the randomization offers several potential benefits. It

may reduce selection bias as patients would not be entered unless the investigator considered the control arm to be a reasonable choice. When the trial is over, the investigator can see if the response rate for the randomized control arm is similar to that of historical controls. This may help keep the results for the experimental arm in perspective. For example, if the experimental arm has a much higher response rate than historical controls, but the randomized control has a response rate similar to that of the experimental arm, the investigator may be less enthusiastic about moving the experimental treatment forward without further phase II testing. Similarly, if the experimental arm is not significantly different from historical controls, but the randomized controls have a much lower response rate than anticipated, an investigator may be less likely to abandon the experimental treatment without further phase II testing. We achieve these gains for the price of doubling our sample size. Note that even though the sample size has doubled, the same number of patients is exposed to the experimental treatment as would be in a traditional single-arm trial. Thus, there is no increased risk to patients. The benefits of this design often outweigh the cost of the increased sample size, particularly if the historical control data are of limited reliability.

##### *Formal Comparison*

When the control arm is used for direct statistical comparison with the experimental arm, the required sample size for the trial is roughly quadruple what would be required for a traditional single-arm trial. Additionally, the number of patients exposed to the experimental treatment is twice what would be necessary for a traditional single-arm trial. Along with ruling out selection bias, this design has the additional benefit of being able to identify statistically significant differences between the experimental and control arms. However, this design is easily misused to perform underpowered trials and look at multiple endpoints without adjustment. We believe the additional cost of the increased sample size and the risk of exposing more patients to experimental therapy frequently outweigh the benefits for this design, although the quantity and quality of the historical data must enter in to this decision.

## Randomization Between Multiple Experimental Arms

Randomization between multiple experimental arms is useful when multiple promising experimental therapies are available, but cost or resource issues dictate that only one can be carried forward into phase III trials. In this design, frequently referred to as a selection trial or a “pick the winner” design, each experimental therapy is first compared to a historical control using the same criteria as in a traditional single-arm phase II clinical trial. The “pick the winner” aspect is employed only to choose the best of the experimental therapies that were statistically superior to the historical control for the primary endpoint. If no experimental therapy is statistically superior to the historical control, no experimental therapy is carried forward. There is no increase in sample size or number of patients exposed to experimental therapies for this design as compared to independent single-arm trials. The randomization allows for a greater degree of unbiased comparisons of toxicities and efficacy between regimens than would be provided by independent single-arm trials. Clearly this approach accrues many benefits at little or no additional cost and without risk of exposing additional patients to experimental therapy when there is a need to screen multiple promising experimental therapies.

## Conclusions

Randomized selection trials provide a rational basis for choosing among multiple experimental therapies and can be done at little or no additional cost and with no additional risk to patients as compared to independent single-arm trials. Randomizing between experimental and control therapies, and comparing each arm to a historical control allows for a reduction in selection bias and provides a validation of the choice of historical control with a doubling of the sample size but no increased risk to patients. Direct statistical comparison in a trial that randomizes between experimental and control arms doubles the number of patients exposed to experimental therapy and quadruples the required sample size as compared to a traditional single-arm phase II trial. However, this direct randomization eliminates reliance on what may be highly variable or unreliable historical control data. We conclude that there most certainly is a place for randomization in phase II clinical trials.

## In Memoriam

It is with great sadness that *Clinical Advances in Hematology & Oncology* reports that H. Samuel Wieand, PhD, passed away in June 2006, after the submission of this article for

publication. On behalf of the editors and publishers, this article is published in memory of Dr. Wieand's contributions to the study of biostatistics and cancer research.

## References

1. Simon R, Wittes RE, Ellenberg SS. Randomized phase II clinical trials. *Cancer Treat Rep.* 1985;69:1375-1381.
2. Rubinstein LV, Korn EL, Freidlin B, et al. Design issues of randomized phase II trials and a proposal for phase II screening trials. *J Clin Oncol.* 2005;23:7199-7206.
3. Simon R. Optimal two-stage designs for phase II clinical trials. *Control Clin Trials.* 1989;10:1-10.

## CON

Emil J Freireich, MD

Director, Adult Leukemia Research Program

and Ruth Harriet Ainsworth Professor

of Developmental Therapeutics

University of Texas M. D. Anderson Cancer Center

Houston, Tex.

The study of a new agent in humans is usually begun at a dose that is substantially below the dose most likely to be biologically effective.<sup>1</sup> This dose level, of course, is intended to reduce the risk of serious toxicity with a drug that has not been given to patients. Regardless of the design of the early clinical trial, there is a progressive escalation in dose, modification of schedule, measurement of pharmacodynamics and pharmacokinetics, and the gathering of information that will be useful in the ultimate evaluation of the efficacy of the new agents. The study of new agents was divided into two phases primarily because many agents were abandoned before an adequate number of patients were studied at biologically effective or maximally tolerated dose. Therefore, phase I trials are designed to determine the biologically effective dose of an agent and phase II trials are used to determine whether the drug has sufficient effectiveness to merit more complete study in what is now called phase III.

Thus, a phase II trial is defined as a study of a new agent, which is conducted after completing phase I evaluation, and studies a sufficient number of subjects at the biologically effective or maximum tolerated dose to obtain an estimate of the agent's effectiveness. Of course, there is a comparative component to the phase II trial. One can compare it to expectations or one can use an arbitrary

standard, such as a minimum of 30% objective response, or whatever criteria one chooses to select the drug for further investigation in phase III trials, where quantitative measurements of effectiveness and quantitative comparison to other treatment strategies is undertaken.

So what is the usefulness of randomization in a phase II study? Clearly, the main reason for randomization in any clinical investigation is to eliminate bias.<sup>2</sup> It is hard to imagine how bias can play any role in a phase II study. So one would have to seek other reasons to randomize. One reason might be to compare to either no treatment or to a standard treatment, or to some known standard in order to get a quantitative assessment on the comparative effectiveness of the new drug to other drugs. This is a potentially dangerous undertaking and can lead to misinformation because the design of the study is not that of a phase III or comparative study but a phase II or estimate study. The consequence of misinterpreting the results of a phase II study can be serious, particularly because it might lead to both unnecessary phase III studies for falsely positive results or rejection of potentially active drugs for falsely negative trials.<sup>3</sup>

There is one circumstance where randomization is useful, and that is the situation where investigators are faced with a large number of potentially effective therapies, all of which have completed phase I trials, and they are competing for patient accrual. In that case, investigators may actually introduce bias by selecting the phase II drug, which they predict to be the best choice. In such a

case, randomization would eliminate that bias and would give each of the drugs a fair chance to be evaluated. In this circumstance, a particularly useful experimental design is adaptive randomization where the randomization is progressively modified to favor the compound that is showing the best results. This allows an objective evaluation of the best choice of an agent for a more detailed phase III investigation.<sup>4,5</sup>

In summary, this is a moment in history when we have a large number new agents to investigate in the clinic. The consequence is that we have to use the techniques that are most efficient in terms of identifying effective new therapies and, most importantly, provide each patient with the best opportunity to receive effective therapy. These goals are best served by adequate, efficient phase I studies, well-designed phase II studies, and phase III studies that identify treatments that provide the greatest benefit for our patients.

## References

1. Freireich EJ, Gehan EA, Rall DP, et al. Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey, and man. *Cancer Chemother Rep.* 1966;50:219-244.
2. Freireich EJ, Gehan EA. The limitations of the randomized clinical trial. In: *Methods in Cancer Research*. DeVita VT Jr, Busch H, eds. Academic Press; New York: 1979;17:177-210.
3. Liu PY, LeBlanc M, Desai M. False positive rates of randomized phase II designs. *Control Clin Trials.* 1999;20:343-352.
4. Estey EH, Thall PF. New designs for phase 2 clinical trials. *Blood.* 2003;102:442-448.
5. Berry DA. Bayesian clinical trials. *Nat Rev Drug Discov.* 2006;5:27-36.