

# Clinical and Prophylactic Trials with Assured New Treatment for Those at Greater Risk: I. A Design Proposal

## ABSTRACT

**Objectives.** The accepted *sine qua non* for estimating the difference in efficacy between a new and a standard treatment is a randomized controlled clinical trial. Yet in some situations it is either practically or ethically impossible to conduct such a trial. For example, patients who are desperately ill may decline to participate when they learn they may not receive the new treatment, especially when that treatment is readily available outside the experimental protocol. Likewise, in a prophylactic trial of a promising vaccine, recruitment of persons at greater risk may falter or fail. Our objective is to demonstrate that a rigorous comparison of treatments may still be attainable.

**Methods.** The features of a controlled clinical or prophylactic trial are reviewed from the perspectives of Food and Drug Administration regulations, ethical considerations, and practical problems.

**Results.** An explicit risk-based allocation method of design and analysis is proposed, one guaranteeing that all subjects at greater risk will receive the new treatment.

**Conclusions.** Under certain conditions, a risk-based allocation trial can furnish consistent estimates of both standard and experimental treatment effects for those at greater risk while avoiding certain difficulties caused by randomized treatment allocation. (*Am J Public Health.* 1996;86:691-695)

Michael O. Finkelstein, JD, Bruce Levin, PhD, and Herbert Robbins, PhD

## Introduction

The human immunodeficiency virus/acquired immunodeficiency disease (HIV/AIDS) epidemic has focused attention on the way new drugs are tested and approved in the United States. The basic requirement, mandated by law, is a controlled clinical or prophylactic trial that demonstrates the efficacy and safety of the proposed treatment. In recent years, the protocols for such trials have received new scrutiny, and the acceptability of long-established practices has been questioned. One feature of particular concern is the randomly selected control group, which receives a standard treatment or a placebo. Although there is no assurance that the experimental therapy or prophylaxis will be better than the standard one, a drug or vaccine that reaches the trial stage is usually perceived as promising, while the standard treatment may be largely ineffective. In such cases, random assignment of patients or designated populations to the control group raises important ethical and practical problems. These problems arise from the lack of assurance that participants who are either very sick or at grave risk will receive the new intervention.

To cope with these problems in severe cases, we propose that instead of random allocation to intervention or control status, the trial protocol use a risk-based method of allocation. In this method, all those at greater risk receive the new treatment or prophylaxis while only persons at lesser risk are the controls. Statistical models based on data for the concurrent controls are then used to estimate what the response of those at greater risk would have been if they had received the placebo or standard treatment.

These methods are applicable beyond the realm of HIV/AIDS. They apply to any circumstance in which a control group would normally be used but one is reluctant to deny the possible advantage of the treatment or intervention to those who most need it, or in which recruitment would be compromised when prospective subjects are told (as they must be) that they may be denied the treatment. Thus, risk-based allocation methods have application in trials in both clinical and public health settings. The latter might include prophylactic trials for either primary or secondary prevention, as exemplified by the control of hypercholesterolemia to prevent coronary heart disease. In the United States, trials of vaccines against HIV are quite likely to be compromised because of the possible denial of the vaccine. A long-standing case is the secondary prevention of cervical cancer through Papanicolaou smears to detect cellular dysplasia: controlled trials were hamstrung once the approach became firmly entrenched without benefit of any trials.

Michael O. Finkelstein is with the law firm of Patterson Belknap, Webb & Tyler, and Columbia Law School, New York, NY. Bruce Levin is with the Division of Biostatistics, Columbia University School of Public Health, and the New York State Psychiatric Institute, New York, NY. Herbert Robbins is with the Department of Statistics and the Institute of Biostatistics at Rutgers University, New Brunswick, NJ.

Requests for reprints should be sent to Bruce Levin, PhD, Division of Biostatistics, Columbia University School of Public Health, 600 W 168th St, New York, NY 10032.

This paper was accepted October 24, 1995.  
**Editor's Note.** Dr Susser was the editor responsible for this paper. Dr Levin is consulting editor for statistics for the Journal and, as is our practice, had no part in the review and decision process.

See related editorial by Mosteller (p 622) in this issue.

How assured allocation to surmount such difficulties would work in practice is the subject of two articles. In this article, we review briefly the basic elements of randomized controlled trials and the problems that have arisen in some cases. In the second article, we give some examples in which risk-based allocation methods are applied to the data of randomized controlled trials and the results of the two methods are compared. We emphasize at the outset that a randomized controlled trial is the method of choice in testing new therapies and that alternatives should be considered only when practical difficulties pose a substantial threat to the conduct or integrity of a proposed trial.

### *Regulations for Clinical Trials of New Drugs*

The protocols for clinical trials of new drugs are governed by the federal Food, Drug, and Cosmetic Act and Food and Drug Administration (FDA) regulations. With certain exceptions, these require that the FDA must approve a new drug as safe and effective before that drug may be introduced into interstate commerce.<sup>1</sup> The approval standards for safety and efficacy were established at different times and are not identical. Premarketing approval for safety stems from the original 1938 act, which provides broadly that there must be "adequate tests by all means reasonably applicable to show whether such drug is safe."<sup>2</sup> (The history and requirements of the 1938 act are discussed in *Weinberger v Hynson, Westcott & Dunning*, 412 US 609 [1973].) A requirement of approval for efficacy, added in 1962, is more specific; as amended in that year, the act requires the FDA to refuse approval of a new drug and to withdraw any prior approval if "substantial evidence" that the new drug is effective for its intended use is lacking. "Substantial evidence" is defined as "evidence of adequate and *well-controlled* investigations, including clinical investigations, by experts qualified by scientific experience and training" (emphasis added).<sup>3</sup>

FDA regulations implement the statute by emphasizing the importance of controlled studies and specifying their key elements. Under the regulations, the agency reports of "adequate and well-controlled studies" as the "primary basis" for determining whether there is "substantial evidence" to support the claims of effectiveness for new drugs.<sup>4</sup> The regula-

tions emphasize that "uncontrolled or partially controlled studies are not acceptable as the sole basis for the approval of claims of effectiveness," although such studies may provide "corroborative support of well-controlled studies regarding efficacy and valuable data on safety."<sup>5</sup>

The protocols for controlled trials are outlined in the regulations, which begin by invoking the authority of scientific consensus: "The characteristics of [controlled clinical trials] . . . have been developed over a period of years and are recognized by the scientific community as the essentials of an adequate and well-controlled clinical investigation."<sup>4</sup> The study must use "a design that permits a valid comparison with a control to provide a quantitative assessment of drug effect."<sup>6</sup> The core is a proper control group. The permissible controls are described in some detail. There is a "placebo concurrent control," in which the control group receives a placebo; a "dose-comparison concurrent control," in which the control group receives a different dose from that of the treatment group; a "no-treatment concurrent control," to be used only when the placebo effect is anticipated to be negligible; and an "active treatment concurrent control," in which the test drug is compared with a known effective therapy.<sup>7</sup> These concurrent controls stand in contrast to "historical" controls, in which the results of treatment with the test drug are compared with "experience historically derived from the adequately documented natural history of the disease or condition, or from the results of active treatment, in comparable patients or populations."<sup>8</sup>

For a concurrent control study, the method of assigning patients to treatment and control groups must minimize bias and "assure comparability of the groups with respect to pertinent variables such as age, sex, severity of disease, duration of disease, and use of drugs or therapy other than the test drug."<sup>9</sup> According to the regulations, it is because such pertinent variables cannot be well assessed in historical populations that the use of historical controls is limited to "special circumstances," such as diseases with "high and predictable mortality (certain malignancies) and studies in which the effect of the drug is self-evident (general anesthetics, drug metabolism)." The goals of unbiasedness and comparability generally are met by randomizing the allocation of treatment and control status across subjects.

### *Issues Raised by Control Groups*

However, these established methods of conducting trials do not satisfy desperately sick people—most conspicuously, victims of advanced AIDS—who face the prospect of near-term death. These people have leveled barrages of criticism at the way clinical trials are managed, from the entry criteria to the speed of approval of new drugs. In particular, they argue that the use of control groups is scientifically unnecessary (because death is predictable) and ethically unacceptable (because a doctor cannot prescribe less than the best available treatment, and the control group may not receive that treatment). The relative permissiveness of the National Institutes of Health, which regulates cancer trials, has been contrasted favorably with the FDA's practice.<sup>10</sup>

The argument that control groups are scientifically unnecessary has not been accepted and has not resulted in any significant structural change in the conduct of controlled clinical trials. Ethical concerns are another matter. To meet ethical objections to control groups, the principle has been articulated that all branches of a trial must be in clinical equipoise—that is, that none of the treatment strategies has been established as clearly preferable.<sup>11</sup> It follows that if, as the trial progresses, accumulating data indicate strongly that the new drug is either efficacious or toxic, the trial should be discontinued and the new drug either offered to all (if there appears to be efficacy) or withdrawn (if there appears to be toxicity).

This principle rules out egregious abuses (e.g., giving placebos to a control group when therapies are established and available) but is in fact far more problematic and complex than would appear from the deceptive simplicity of its formulation. Views differ as to whether the personal opinion of the patient's physician, unsupported by clinical trial evidence, should create an ethical obligation that would preclude participation in a trial in which the patient may be randomized to a treatment that the physician does not regard as optimal.<sup>12-16</sup> If so, trials would be severely restricted. Moreover, if a trial is terminated early when a new drug seems to be promising, the data of the trial can no longer be analyzed in a simple way to determine the drug's effect.<sup>17</sup> The statistical theory of sequential trials can be used to correct for bias and to calculate significance levels and confidence intervals when the stopping rule is prespeci-

fied. But the precision of information gained about primary and secondary endpoints may not be great, and the theory does not cover those common situations in which trials are stopped because of side effects or information outside the trial (e.g., late-breaking results from other trials). When early termination is suggested by such input, scientific and patient interests may collide.<sup>18</sup>

In addition to ethical issues, control groups create practical problems in recruitment and risks of sabotage. While experts debate the ethics of controlled trials, some AIDS patients have taken matters into their own hands. In significant numbers they have refused to volunteer for trials or have undermined them either by sharing the drug with others, thereby sabotaging dosage design, or by opting out if they believe they are in the standard treatment group. Describing these and other problems, one commentator has concluded that "clearly the magnitude of noncompliance in AIDS research threatens the validity of many important ongoing clinical trials."<sup>19</sup>(p24) Those involved in the design and conduct of clinical trials have moved cautiously in response to activist pressure.<sup>20</sup>

The problem of recruitment is exacerbated when the experimental treatment is readily available outside of a controlled trial. Randomized studies of bone marrow transplant therapy for breast cancer patients have been stymied in New York by the refusal of women to participate in controlled trials since they can be assured of receiving the transplant treatment elsewhere.<sup>21</sup> Recruitment for AIDS trials may have been hampered by a 1988 congressional enactment providing for early access to new AIDS drugs still undergoing trials.<sup>22</sup>

Clinical trial recruitment may be helped by providing for patient crossovers. Some trial protocols permit a patient who does poorly under the originally selected treatment to cross over to another arm of the trial during the follow-up period. Under the conventional intention-to-treat principle, however, patients are counted in the group to which they were originally randomized, despite subsequent changes. While preserving a valid type I error rate for the trial, intention-to-treat analysis degrades the clinical meaningfulness of the measured treatment effect when there are a large number of such crossovers, and it causes a loss of statistical power when the null hypothesis is false.<sup>23</sup> Crossovers, and

patients who drop out or are otherwise noncompliant, thus cast doubt on the statistical integrity of a trial.

### *An Alternative: Risk-Based Allocation*

What can be done to close the gap between those who insist on randomized controlled trials for the soundness of scientific investigation and those who object to them for ethical, compassionate, or practical reasons? The rules governing controlled trials are sanctified in science and have been written into law. Nevertheless, we suggest that in cases with severe practical problems, we can still achieve the goals of controlled trials by adopting a *risk-based allocation* design, in which the treatment group is not selected at random but consists of persons who are at greater risk while the control group consists of those who are at lower risk. Statistical models are then used to adjust for the differences between the two groups. The models grow out of theoretical work in empirical Bayes estimation by one of us (H.R.) and Cun-Hui Zhang.<sup>24-27</sup> These models are in the modern spirit of stochastic personal risk parameter models for causal inference.<sup>28</sup> In this approach, the effect of the standard treatment on the persons at greater risk, who in fact receive the new treatment, is estimated from a mathematical model instead of being observed directly. The comparison among persons at greater risk of their directly observed response to new treatment with their estimated response to the standard treatment is the measure of the differential treatment effect. This departs from the conventional measure of treatment effect in that the differential effect is measured only for the persons at greater risk whereas the conventional measure reflects the average difference over all subjects. This measure can be justified if the benefit to the persons at greater risk is a point of central interest. If the response of the lower-risk persons is also of interest, the trial protocol could provide for random allocation of those subjects to new or standard treatment while guaranteeing the new treatment for all the higher-risk subjects. This elaboration offers no analytic difficulties and may, in certain cases, facilitate approval of the protocol.

For example, in the bone marrow transplant randomized trial referred to above, only women with 10 or more involved nodes are eligible. Thus, the trial is designed to estimate treatment effect

for this group of patients. With a risk-based allocation design, these women would be guaranteed bone marrow transplant while women with fewer than 10 involved nodes would serve as controls and be given standard therapy. Our estimate of treatment effect would apply to the originally targeted group of patients.

Under this protocol, the subjects at greater risk, who are in most urgent need of treatment, would *all* get the new treatment and so should be more willing to cooperate in the study. The subjects at lower risk would not have that assurance but would know that, if the drug is effective, they would have an opportunity to be treated with it later. If informed consent is obtained prior to the determination of disease severity (i.e., measurement[s] determining treatment allocation), subjects may be told that they will definitely receive the experimental treatment if their condition or disease severity is found to warrant it.

One might view this risk-based allocation as an intermediate protocol between historical controls and the standard randomized concurrent controls. The investigators would rely on prior experience to choose the mathematical *form* of the model, but on current data to estimate the parameters of that model. Since they would look to history only for generalized guidance in model building, that more limited use should help to meet the objection noted in the regulations that historical control groups cannot usually be relied on because "pertinent variables" cannot be well assessed in historical control populations.

In a similar vein, the use of risk-based allocation addresses one criticism of observational studies, which is that in such studies the allocation to treatment or control groups follows an unknown mechanism involving self-selection and other potentially biasing factors. As a result, unless much effort is exerted in the design of the observational study to ensure that the two groups are comparable, the groups will tend to differ with respect to one or more potentially confounding factors. In such cases, simple comparisons of averages will need to be adjusted for the effect of the confounders. In other words, in observational studies, unadjusted treatment effects and adjusted treatment effects generally differ, and adjustment is required to guard against spurious effects. In randomized studies, by contrast, adjusted and unadjusted treatment effects are equal in theory

because of the orthogonality (uncorrelatedness) of treatment allocation with confounders built in by the randomization. (It may be shown that in randomized trials, adjusted treatment effect estimates are more efficient than unadjusted effect estimates for typical analyses based on linear models or linear logistic models.<sup>29</sup> However, the gain in precision is usually small, especially in large trials. In smaller trials, it may even be deemed inadvisable to adjust for covariates when there are many to select from and they yield different conclusions, such that the trial may be called into question for a kind of analytic selection bias.)

In a risk-based allocation trial, the allocation mechanism is known and objective. As a result, all that is required to estimate treatment effect among the subjects at greater risk is a valid model for standard treatment response given true disease severity and/or the allocation variable, but that model need not further adjust for any confounders (see Part II of this paper and the appendix<sup>30</sup> for a precise formulation of such models). Thus, our procedure is exactly parallel to that used in randomized trials where treatment effects are estimated without adjustment for confounders, and, indeed, whatever treatment effect parameters can be consistently estimated for primary endpoints in the randomized trial can also be consistently estimated in the assured allocation design. To be sure, the assumption of a valid model for standard treatment response is critical to the success of the method, but once the form of that response model is determined for the subjects to be recruited, the treatment effect parameter can be estimated with no more bias or need for adjustment for confounders than in the randomized trial. Of course, if it is desired to estimate treatment effects adjusted for a set of confounding factors, one would need to incorporate those factors as conditioning variables into the model, just as in the case of randomized trials.

Apart from issues of statistical estimation, which we discuss in Part II,<sup>30</sup> risk-based allocation would generally be inconsistent with the double-blind feature of many clinical trials. This is a loss that has to be weighed on a case-by-case basis against the benefits of risk-based allocation. Double-blind procedure is not always practically or ethically possible in standard randomized trials, such as the bone marrow transplant trials. Even

double-blind trials have the potential for concealed biases if subjects analyze the drug or otherwise determine what they are receiving. Moreover, the decision to terminate early, if made on a judgmental basis, may reflect the biases of the monitors, who have full knowledge of the results.

The methods we propose are speculative in that they assume specific models for the effect of the standard treatment. To minimize the speculative element inherent in model selection, risk-based allocation should be used only when prior experience with comparable subject populations provides a reasonable basis for specifying the mean standard treatment response as a function of the measure of severity—that is, the allocation variable. Since the appropriateness of the model turns on the effect of a standard treatment, the natural history of the disease or condition, or experience from other studies, must be consulted to provide a basis for model selection. In the bone marrow transplant example, oncologists today can quantify the average decrement in 5-year survival for each additional involved node, which would provide a basis for specifying a model in that case. The model parameters can be estimated from the data for the concurrent control group, and the appropriateness of the model can also be checked against those data. It is an advantage of our method that it requires no model for the effect of the experimental treatment itself, which would usually be quite problematic.

## Conclusion

What has to be considered in each case is the loss of double-blinding and the risk of being seriously wrong about the model for the effect of standard treatment, weighed against the difficulty of finding volunteers, the risks of sabotage, the pressure for premature termination, or even the abandonment of the control group concept. Balancing everything, the safer and ultimately more humane course in some cases may well be a risk-based allocation protocol of the type we propose. We are following here the suggestion that “statisticians should be more sensitive to the physician’s responsibility to the individual patient and should, besides promoting randomized trials when they are ethically and practically feasible, work to improve the planning, execution, and analysis of nonrandomized clinical studies.”<sup>14(p52)</sup>

In Part II, we show how risk-based allocation could have been used in three types of studies, and we compare the results of the actual trials with those obtained under the risk-based methods we propose.<sup>30</sup> □

## Acknowledgments

The authors wish to acknowledge colleagues in the Statistics, Epidemiology, and Data Management Core Group of the Columbia University HIV Center for Clinical and Behavioral Studies (contract #P50-MH43520) for helpful discussions during this work, and to thank David Freedman for his careful reading of and comments on the manuscript.

## References

1. Food, Drug, and Cosmetic Act. 21 USC §355(a),(b) (1994).
2. 52 Stat 1051, §505(d) (1938); 21 USC 355(d) (1994).
3. 21 USC 355(d) (1994).
4. 21 CFR §314.126(a) (1993).
5. 21 CFR §314.126(e) (1993).
6. 21 CFR §314.126(b)(2) (1993).
7. 21 CFR §314.126(b)(2)(i-iv) (1993).
8. 21 CFR §314.126(b)(2)(v) (1993).
9. 21 CFR §314.126(b)(4) (1993).
10. Rothman DJ, Edgar H. Scientific rigor and medical realities: placebo trials in medical research. In: Fee E, Fox DM, eds. *AIDS, The Making of a Chronic Disease*. Berkeley, Calif: University of California Press; 1992: 194–206.
11. Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med*. 1987;317: 141–145.
12. Hellman S, Hellman DS. Of mice but not men: problems of the randomized clinical trial. *N Engl J Med*. 1991;324:1585–1589.
13. Passamani E. Clinical trials—are they ethical? *N Engl J Med*. 1991;324:1589–1592.
14. Royall RM. Ethics and statistics in randomized clinical trials. *Stat Sci*. 1991;6:52–88.
15. Markman M. Ethical difficulties with randomized clinical trials involving cancer patients: examples from the field of gynecologic oncology. *J Clin Ethics*. 1992;3:193–195.
16. Freedman B. A response to a purported ethical dilemma with randomized clinical trials involving cancer patients. *J Clin Ethics*. 1992;3:231–234.
17. Pocock SJ, Hughes MD. Practical problems in interim analyses, with particular regard to estimation. *Controlled Clin Trials*. 1989;10:209S–221S.
18. On stopping a trial before its time. *Lancet*. 1993;342:1311–1312. Editorial.
19. Arras JD. Noncompliance in AIDS research. *Hastings Cent Rep*. 1990;20:24–32.
20. Byar DP, Schoenfeld DA, Green SB, et al. Design considerations for AIDS trials. *N Engl J Med*. 1990;323:1343–1348.
21. Kolata G. Women resist trials to test marrow transplants. *New York Times*. February 15, 1995:C8.
22. 42 USC §300cc-12 (1994).
23. Weinstein GS, Levin B. The effect of crossover on the statistical power of randomized studies. *Ann Thorac Surg*. 1989;48: 490–495.

24. Robbins H, Zhang C-H. Estimating the superiority of a drug to a placebo when all and only those patients at risk are treated with the drug. *Proc Natl Acad Sci USA*. 1989;86:3003-3005.
25. Robbins H, Zhang C-H. Estimating a treatment effect under biased sampling. *Proc Natl Acad Sci USA*. 1988;85:3670-3672.
26. Robbins H, Zhang C-H. Estimating a multiplicative treatment effect under biased allocation. *Biometrika*. 1991;78:349-354.
27. Robbins H. Comparing two treatments under biased allocation. *La Gazette des Sciences Math du Québec*. 1993;15:35-41.
28. Greenland S. Randomization, statistics, and causal inference. *Epidemiology*. 1990;1:421-429.
29. Robinson LD, Jewell NP. Some surprising results about covariate adjustment in logistic regression models. *Int Stat Rev*. 1991;58:227-240.
30. Finkelstein MO, Levin B, Robbins H. Clinical and prophylactic trials with assured new treatment for those at greater risk: II. examples. *Am J Public Health*. 1996;86:696-705.

## Access Journals from the Developing World through ExtraMED

To improve international awareness of and access to journals from developing countries, the World Health Organization has established ExtraMED, a monthly service that provides a CD-ROM of the complete contents of over 200 biomedical journals from around the world. The journals are chiefly from developing countries, and most are not listed on MEDLINE or other major Western indexes.

Each month, as many as 5000 pages are scanned onto a CD-ROM, with all the bibliographic information (including abstracts) rekeyed to permit full text searching; extensive indexing is also supplied. Print-outs look like copies of the

original journal pages, including titles, page numbers, and illustrations.

The service costs \$1800 per year (10 disks). To order and for technical information, please contact DV Trimmer, Informania Ltd., PO Box 1359, London W3 0ZU, England; e-mail 100060.172@compuserve.com; fax (44) 1730-265398. Further information regarding the policies and philosophy behind the ExtraMED project can be obtained from Christopher Zielinski, Manager, Health and Biomedical Information, WHO/EMRO, PO Box 1517, Alexandria 21511, Egypt; tel (203) 4830090; e-mail zielinski@who.sci.eg; fax (203) 4838916.