# Clinical and Prophylactic Trials with Assured New Treatment for Those at Greater Risk: II. Examples

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


## ABSTRACT

*Objectives.* The preceding article proposed an assured treatment design that would address certain difficulties in recruiting persons who are at greater risk into randomized clinical trials. The purpose of this article is to illustrate the statistical validity of the design in a practical setting.

*Methods.* Three actual randomized clinical trials were considered as case studies; in each, the data that would have been obtained under assured allocation were identified. Then, with only these data, together with a reasonable choice of model describing the response of subjects under standard treatment as a function of initial severity, the treatment effect was estimated for the subjects at greater risk. The estimates were compared with conventional estimates for the sicker patients randomized in the original trials.

*Results.* In each case, the estimates produced in the assured treatment trial were close to those observed in the randomized trial.

*Conclusions.* Risk-based allocation trials deserve serious consideration when randomized clinical trials are difficult or impossible to execute. The proposed designs and analyses would allow physicians to offer persons at greater risk assurance that they would receive the new treatment, while researchers would retain the ability to draw valid statistical conclusions about treatment efficacy. (*Am J Public Health.* 1996;86:696–705)


## *Introduction*

Part I of this article describes the practical recruitment problems faced in some randomized clinical trials and suggested that when those problems are severe, consideration should be given to replacing randomized allocation with risk-based (assured) allocation, in which all of the higher-risk persons receive the new treatment.[1] Here we give three examples of risk-based allocation as applied to studies involving measurements, rates of events, and survival times. The data sets of these studies were the only ones used to test risk-based allocation methods, and the cutpoints used to define the high-risk groups were the only ones tried. We did not dredge data sets for favorable results.

## *A Measurement Study*

In many clinical trials, the allocation variable is a measure of risk of future disease, an estimate of disease severity, or a baseline measurement of a disease marker. For a trial of a cholesterol-reducing drug, the allocation variable may be a serum cholesterol measurement; in an acquired immunodeficiency syndrome (AIDS) drug treatment trial, it may be a CD4 cell count. We use the cholesterol example to illustrate the risk-based allocation design involving a measurement.

High cholesterol (at least the low-density lipoprotein component) is generally regarded as a risk factor in heart disease. Our plan is to give all the high-risk patients (those with high cholesterol measurements) a recommended diet and an experimental drug for reducing cholesterol, and to give the lower-cholesterol patients (the control or standard treatment group) the recommended diet and a placebo. In the usual randomized study, the treatment effect is the difference in (or ratio of) the average change between pretreatment and posttreatment measurements in the treatment and control groups. This measure is not valid under risk-based allocation because by the regression-to-the-mean phenomenon alone one would expect a larger decrease in cholesterol for the treatment (high-cholesterol) group than for the control group.

One method of dealing with this problem is to take a second ("auxiliary") pretreatment measurement of each subject's cholesterol level. The auxiliary measurement need not be independent of the allocation measurement, given the "true" or long-term cholesterol level of the individual. The treatment effect is then redefined (although the redefinition is mathematically equivalent to the original) as the average change from the auxiliary to the endpoint (posttreatment) measurement in the new treatment (high-risk) group less the average change that would have resulted for the same subjects if they had received the standard treatment. The

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.

actual average change for the high-risk subjects is observed directly from the results of the trial. The hypothetical average change for these subjects had they received the standard treatment cannot be directly observed since, under risk-based allocation, they all received the new treatment. As a result, we must use a statistical model to estimate what average change would have occurred had the standard treatment been given to this group.

The simple model that we use in this example assumes that the expected change between the auxiliary and endpoint measurements for a subject receiving the standard treatment is a linear function of the form $E(Y - X' | X) = A + BX$, where $E(Y - X' | X)$ is the expected value of the difference between the endpoint measurement $Y$ and the auxiliary measurement $X'$ for a given value of the allocation measurement $X$. The key assumption is that this model holds for *all* values of $X$, both above and below the cutoff point for allocation to treatment or control. Here $A$ and $B$ are constants to be estimated from the data for those control group subjects actually receiving the standard treatment. Once $A$ and $B$ are estimated by standard linear regression methods, the response to the standard treatment is estimated for the high-risk subjects by using their $X$ values in the equation.

The reader may ask why the auxiliary measurement $X'$ is needed at all; why not simply regress the endpoint measurement $Y$ on the allocation variable $X$, using the model $E(Y|X) = A + BX$ to estimate $A$ and $B$? This is, in fact, sometimes proposed for making adjustments in observational studies. It suffers, however, from the objection that the linear relationship between $Y$ and $X$ posited by the model will hold only in special circumstances. Suppose, for example, that each subject has a true (long-term, or "error-free") level of cholesterol that varies arbitrarily in the patient population. Suppose further that the allocation measurement $X$ varies about the true level with a normally distributed error. Then, even assuming that the mean value of $Y$ is a linear function of the true level and $X$, it can be shown that only a normal distribution for the *true* cholesterol level in the population will produce the linear model $E(Y|X) = A + BX$. While the assumption of normally distributed measurement error for $X$ around the true level is quite reasonable, the assumption that the true level is normally distributed in the population is gratuitous. As a result, we would generally not find $E(Y|X)$ to be linear.

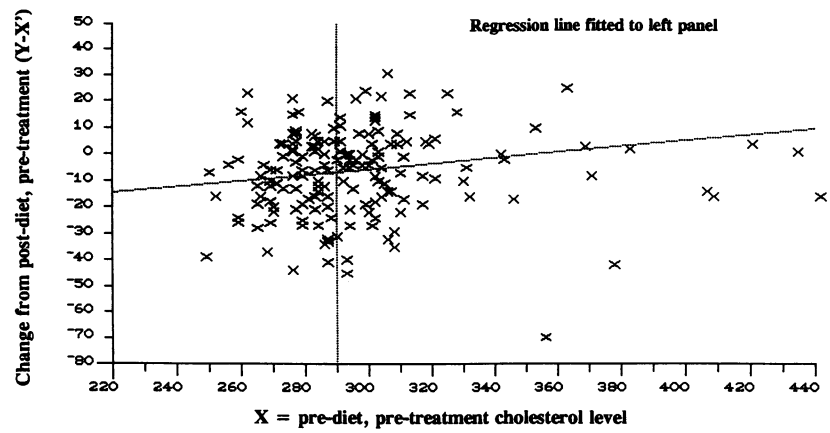
**TABLE 1—The Effect of Cholestyramine vs Placebo on Total Plasma Cholesterol Levels, by Pretrial Cholesterol Level and Treatment vs Control Group Status**

| Group[a] | Treatment[a] | Reduction between Pretrial Auxiliary and Posttreatment, mg/dL[b] | Standard Error | No. Subjects |
|---|---|---|---|---|
| ≥290 | Experimental | 34.42 | 3.295 | 88 |
| ≥290 | Placebo | 5.02 | 1.831 | 89 |
| <290 | Experimental | 27.03 | 3.473 | 77 |
| <290 | Placebo | 8.17 | 1.670 | 83 |

*Source.* Data are from 337 subjects randomly assigned in the Stanford portion of the Lipid Research Clinic's Coronary Primary Prevention Trial.

[a]Grouped by initial total cholesterol level above or below 290 mg/dL (approximately the median), as determined by a baseline (allocation) measurement and by assignment to experimental or placebo treatment.

[b]The mean difference in mg/dL of total cholesterol between a second pretrial (auxiliary) measurement and the posttrial measurement.




*Note.* On the abscissa, $X$ is the prediet, pretreatment total cholesterol level as determined by the allocation measurement. On the ordinate, $Y$-$X'$ is the difference in milligrams per deciliter in total cholesterol between the posttrial measurement $Y$ and pretreatment level as determined by the auxiliary measurement $X'$. The regression line is based on the 83 subjects with initial total cholesterol levels below 290 mg/dL.

**FIGURE 1—Change in total plasma cholesterol level for 172 placebo-treated men in the Stanford portion of the Lipid Research Clinics Coronary Primary Prevention Trial.**

With an auxiliary measurement $X'$, the relationship between $Y - X'$ and $X$ can be linear without any assumed distribution of true cholesterol values. As a result, the linearity assumption is more likely to be valid with the auxiliary measurement than without it. For a discussion of this point, see the appendix. In addition, the model with an auxiliary measurement is advantageous because its estimates of treatment effect will usually have a smaller sampling error.

We now apply this method to a portion of the data from a placebo-controlled, double-blind, randomized clinical trial in which the drug cholestyramine was used to lower blood cholesterol.[2] The data are from the Stanford University portion of the original multicenter study and have been the subject of previous discussion.[3] The purpose of the trial was to test the efficacy of lowering cholesterol in reducing risk of coronary heart disease, but we take as our endpoint for discussion the effect of cholestyramine on total cholesterol levels.

In this study, there were 165 subjects in the treatment group and 172 in the placebo group. A series of pre- and posttreatment measurements were made

**TABLE 2—The Effect of Low-Dose (Experimental) AZT vs High-Dose (Standard) AZT on the Rate of Opportunistic Infections in AIDS Patients, by Pretrial CD4 Cell Count and Treatment**

| CD4 Group[a] | Treatment[a] | No. Opportunistic Infections[b] | Total Follow-Up Time, Years[c] | Rate (Infections Per Year) | No. Subjects[d] |
|---|---|---|---|---|---|
| >60 | Experimental (low dose) | 292 | 217.88 | 1.340 | 126 |
| >60 | Standard (high dose) | 290 | 210.92 | 1.375 | 133 |
| ≤60 | Experimental (low dose) | 262 | 207.10 | 1.265 | 128 |
| ≤60 | Standard (high dose) | 296 | 181.33 | 1.632 | 126 |

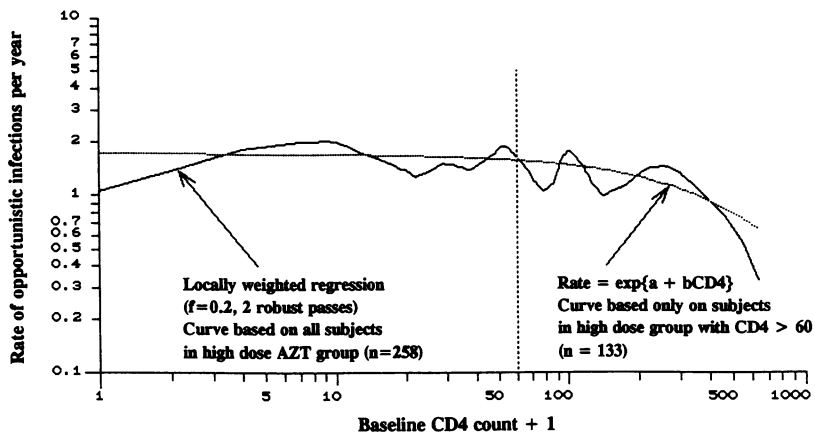
*Source.* Data are from 513 patients randomly assigned in the AIDS Clinical Trials Group Study 002 of low- vs high-dose AZT.

[a]Grouped by CD4 cell count above or below 60/μL of blood (approximately the median), as measured at randomization and by assignment to experimental (low-dose) and standard (high-dose) treatment.

[b]The total number of opportunistic infections in the group during follow-up.

[c]The total number of years of follow-up for the group, with follow-up period from the date of randomization to the date of death or withdrawal from therapy.

[d]Includes one subject in fourth group with zero days of follow-up. Eleven subjects from the original trial with missing CD4 cell counts have been excluded.




*Note.* The abscissa is the log of 1 plus the CD4 cell count at randomization. The ordinate is the log of the rate of opportunistic infections. The wavy solid line is the actual rate of opportunistic infections smoothed by locally weighted regression ($f$ = 0.2, 2 robust passes). The curve is based on all 258 subjects in the high-dose group. The smooth dotted curve is the rate of such infections estimated by the model, Rate = exp ($a$ + $b$CD4). The curve is based only on the 133 subjects in the high-dose group with a CD4 cell count above 60/μL. AZT = zidovudine.

**FIGURE 2—Estimated and observed rates of opportunistic infections for 258 high-dose AZT subjects in AIDS Clinical Trials Group Study 002.**

of total cholesterol and its components; here we focus on the total. One of the pretreatment measurements was made just prior to a baseline period diet prescribed for all trial participants. This we use as our allocation variable $X$. A subsequent pretreatment measurement was taken just prior to randomization about 4 months later. This we use as our auxiliary measurement $X'$. To compare this actual trial with our proposed risk-based allocation trial, we considered the subgroup of individuals whose prediet cholesterol measurement exceeded the median and chose them to receive the new treatment.

In the randomized clinical trial there were 177 subjects whose prediet cholesterol measurement exceeded 290 mg/dL of blood (approximately the median); of these, by random selection, 88 were in the treatment subgroup and 89 were in the control subgroup. The 88 treatment subjects in this high-cholesterol subgroup showed an average reduction of 34.42 mg/dL from the baseline auxiliary (post-diet cholesterol) measurement, while the 89 control subjects showed an average reduction of 5.02 mg/dL from that measurement. The "gold standard" estimate of treatment effect is thus a 29.40 mg/dL reduction (34.42 mg/dL − 5.02 mg/dL) for the high-cholesterol subgroup, while the standard error of this estimate is ±3.77 mg/dL. This result from the randomized trial is for the subgroup with cholesterol levels above 290 mg/dL to make that group correspond to the group receiving treatment under risk-based allocation. For the entire group of subjects in the randomized trial, including those with cholesterol levels below 290 mg/dL, the 165 treatment subjects showed an average reduction of 30.97 mg/dL from the baseline auxiliary measurement, while the 172 placebo subjects showed an average reduction of 6.54 mg/dL from the baseline. The treatment effect was thus measured as a reduction of 24.43 mg/dL from the baseline, with a standard error of ±2.70 mg/dL. This smaller average reduction for the entire group suggests that the drug is less effective (in absolute terms) for subjects with lower initial measurements. Some results from the data of the study are summarized in Table 1.

To test how close risk-based allocation results would be to this conventional randomized clinical trial result, we created a risk-based allocation by discarding those cases in which a placebo was given to subjects with cholesterol readings by the allocation measurement at or above 290 mg/dL and those cases in which the new treatment was given to subjects with cholesterol readings below 290 mg/dL. With these deletions, the data consisted of 88 subjects with $X$ at or above 290 mg/dL, all of whom received the new treatment, and 83 subjects with $X$ below 290 mg, all of whom received the placebo.

Using the linear model $E(Y - X' \mid X) = A + BX$, we estimated the constants $A$ and $B$, by ordinary least squares regression from the control group data, to be $A = -38.54$ and $B = 0.1102$. The treatment effect before adjustment for the placebo effect is, as stated, a reduction of 34.42 mg/dL for the cholestyramine group. Using the regression equation, the estimated standard treatment response for that group is a reduction of 3.66 mg/dL. The treatment effect after adjustment for the standard treatment response is the difference between the two, or a reduction of 30.76 mg/dL. This is close to the randomized trial result of 29.40 mg/dL. The standard error of the 30.76 mg/dL estimate is ±8.02 mg/dL. (The standard

error is calculated by summing the independent contributions to the variance from the treatment and control groups. The variance from the treatment group is the usual variance of a sample mean, which in this case is $[30.91]^2/88 = 10.86$. The mean value of $X$ among those with $X \geq 290$ is 316.56. The estimated variance of the predicted mean from the regression model at $X = 316.56$ is 53.39. The variance of the sum is 10.86 + 53.39 = 64.25. The standard error is the square root of that, or ± 8.02, as stated.) The data and the regression line are shown in Figure 1. Thus, for the high-risk patients, the risk-based allocation results are virtually identical to those of the conventional trial.

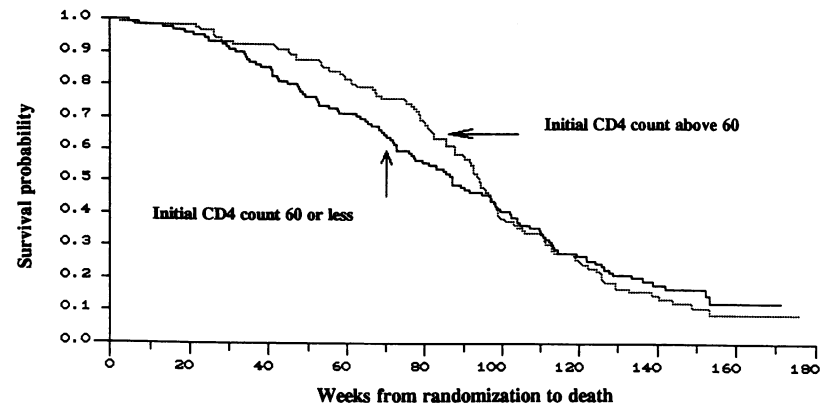
As the calculation of the variance shows, for comparable sample sizes and variances, the size of the regression model component of the variance depends primarily on the degree of extrapolation from the lower observed range of the data for the control group to the higher range of the data for the treatment group. The predominance of that term in the sum indicates that the bigger the average gap in severity of illness between the treatment and control groups, the larger the variance of the estimates because the projection requires a greater extrapolation. On the other hand, if the two groups are close together, the problems encountered in the classical randomized trial that we noted in Part I[1] begin to emerge. One way to decrease the standard error of the estimate when the difference in severity of illness between the two groups is large is to augment the size of the control group.

## *A Rate-of-Event Study*

In many clinical trials, the outcome being measured is the rate of some event (i.e., the expected number of events per unit time per person). In immunosuppressed patients, a therapy may be tried to reduce the rate of opportunistic infections. In risk-based allocation trials of such therapies, the allocation variable may be a CD4 cell count, and one must use a model that relates the value of the count to the rate of opportunistic infections. The key, of course, is the correctness of the model. There should be data from prior studies or from the natural history of the disease against which to test the proposed choices.

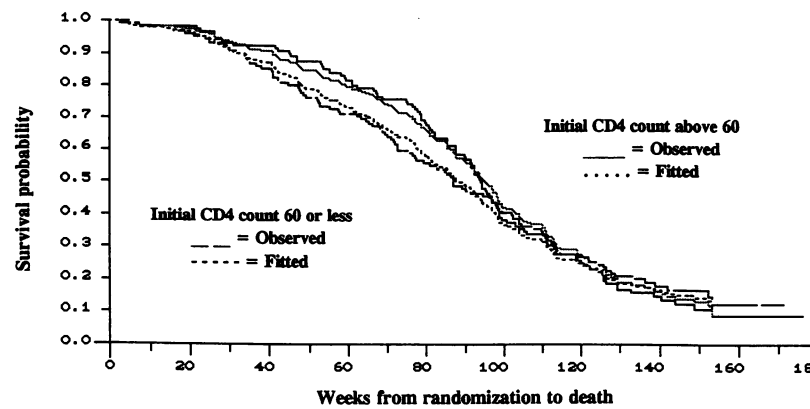
To illustrate this situation, we use data from a clinical trial to test the efficacy of low-dose vs high-dose zidovudine (AZT) on very sick AIDS patients.[4] At the time of this trial (referred to as ACTG [AIDS Clinical Trials Group] 002), low-dose AZT (500 mg/day) was the experimental treatment and high-dose AZT (1500 mg/day) was the standard. Although survival time was the major endpoint of the trial, and although the study has been criticized as biased,[5] we focus here on a secondary endpoint, the number of opportunistic infections, to illustrate a study involving count data. (The data are available on diskette from the National Technical Information Service as ACTG 002, NTIS Order No. PB93-506087.)

In this trial, there were 254 subjects in the low-dose (experimental) group and 258 in the high-dose (standard treatment) group. (There were 524 patients enrolled in the study, but 11 had missing information on CD4 cell counts and 1 had zero days of follow-up time.) Data for those subjects receiving the standard treatment suggest that the opportunistic infection rate can be appropriately modeled by the exponential function $R(X) = A \exp\{BX\}$,



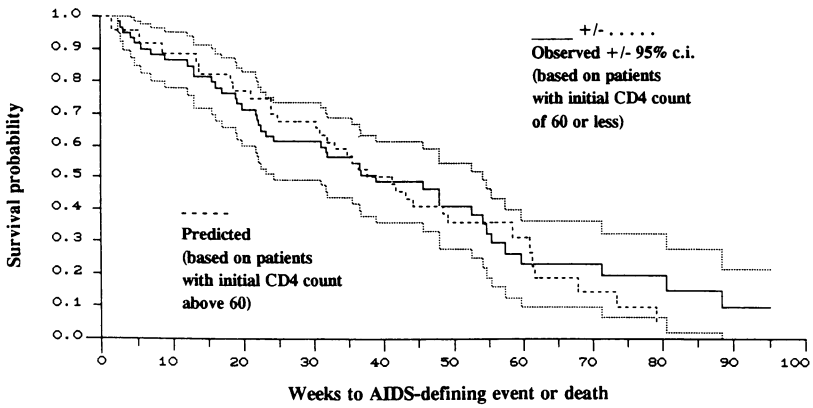

*Note.* Observed survival curves are product-limit estimates for the indicated subgroups from the randomized clinical trial: CD4 cell count at or below 60 (n = 128) and CD4 cell count above 60 (n = 126). AZT = zidovudine.

**FIGURE 3—Observed survival rates for 254 patients on low-dose AZT in AIDS Clinical Trials Group Study 002, by initial CD4 cell count.**




*Note.* The solid and broken lines are the observed survival probabilities for the groups with initial CD4 cell counts above or below 60/μL of blood, respectively. The dotted and dashed lines are averages of the adjusted survival probabilities for patients with initial CD4 cell counts above or below 60/μL, respectively, as estimated by a Cox model with a time-varying relative hazard parameter. AZT = zidovudine.

**FIGURE 4—Estimated and observed survival rates for 254 patients on low-dose AZT in AIDS Clinical Trials Group Study 002, by initial CD4 cell count.**




*Note.* The solid line represents observed survival probabilities. The dotted lines provide 95% confidence intervals for the observed survival probabilities. The dashed line represents probabilities estimated from a Cox model with a time-varying relative hazard parameter based on 149 patients with an initial CD4 cell count above 60/μL. AZT = zidovudine.

**FIGURE 5—Estimated and observed times to AIDS event or death: 60 AIDS-related complex or asymptomatic patients with initial CD4 cell counts at or below 60/μL on AZT in AIDS Clinical Trials Group Study 116B/117.**

where $X$ is the CD4 cell count at the start of the trial and $A$ and $B$ are constants to be estimated from the data. We assume there would have been a basis in prior experience for choosing this form of model. The rate, $R(X)$, is the expected number of opportunistic infections per year of survival per patient for those with initial CD4 count $X$. This is a commonly used form of model for event data, equivalent to the model $lnR(X) = ln\,A + BX$. We take CD4 cell counts at or below 60 μL of blood as a marker for the sicker patients; this is approximately the median for the group.

In the original trial, the number of opportunistic infections in the high-dose (standard) treatment group among the sicker patients was 296, with 66 186 days of follow-up, for an opportunistic infection rate of 1.632 per year. The follow-up period was from the date of randomization to the date of withdrawal from therapy or death. Many subjects withdrew because of the toxicity of AZT. Indeed, this was partly the reason for testing the efficacy of the low-dose therapy. The number of opportunistic infections in the low-dose (experimental) treatment group among the sicker patients was 262, with 75 591 follow-up days and a rate per year of 1.265. The ratio of treatment to control group rates is 1.632/1.265 = 1.290 (standard error ±0.109). Thus, the standard estimate of the low-dose effect on the sickest patients is that it reduces their rate of opportunistic infections by about 22.5% (1–1/1.290 = 0.225). Selected results from the trial are summarized in Table 2.

If the trial had used a risk-based allocation scheme with all of the sicker patients receiving the experimental low dose, the effect of the high dose on the sicker patients would have been estimated instead of being directly observed. Under the modeling approach, this would have been done by fitting the model from the data of the less sick (CD4 cell counts > 60) patients who received the standard treatment. This involves selecting by computer iteration the pair of values for the parameters $A$ and $B$ that best fits the data. (Technically, we used maximum likelihood estimation for the data under a Poisson regression model, which assumes that the number of opportunistic infections occurring in a given time period $t$ has a Poisson distribution with mean $tR[X]$.) The result is $A = 0.541$ and $B = -0.00155$. The curve thus estimated compared with the smoothed actual data is shown in Figure 2.

The model estimates, for example, that with a CD4 cell count of 60, the opportunistic infection rate would be 1.452 per year; with a count of 10, it would be 1.526 per year. In the experimental group, the expected number of opportunistic infections for each patient is derived by multiplying the observed follow-up for that patient by the rate per year. Under standard treatment, the total expected number of opportunistic infections for the sicker patients is the sum of these expectations over all 128 such patients who in fact received the low dose. That sum is 340.46, as compared with the actual number of 262. Thus, the estimated rate ratio (the treatment effect in the present case) among the sicker patients is 340.46/262 = 1.2995 (with a standard error of approximately ±0.147 after adjusting for overdispersion). Under risk-based allocation, then, the estimated low-dose effect on the sicker patients is a reduction in the rate of opportunistic infection of 23.0% (1–1/1.2995 = 0.230), close to the standard estimate of 22.5%.

Note that the rate ratio in the original trial among the less sick patients (CD4 cell counts >60) was nearly unity (1.375/1.340), suggesting little treatment efficacy in that subgroup. The rate of opportunistic infections per unit time (our parameter of interest, $R$) was lower in the experimental group than in the standard treatment group because of a reduced number of opportunistic infections and a longer survival time in the experimental group. Both factors are a legitimate benefit of the treatment that is appropriate to reflect in $R$ because the longer survival time probably resulted from enhanced protection against opportunistic infections afforded by the low-dose regimen. In the estimation of $R$, the use of risk-based allocation and an appropriate model generated results that are virtually indistinguishable from those of the randomized clinical trial.

## *A Survival Time Study*

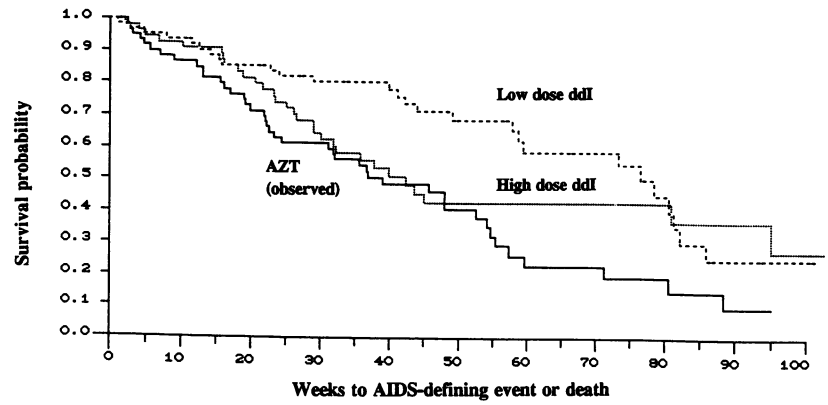
A third subject of study in clinical trials is the effect of the treatment on time to some event. In the AIDS context, an important endpoint is death, so the measurement of interest is survival time. This is commonly appraised by comparing survival curves for the treatment and control groups. In a randomized trial, these curves are calculated from the observed data for those groups. In a risk-based allocation trial, the survival probabilities for the high-risk group receiving the new treatment are calculated in the usual way but those for the high-risk group assuming receipt of the standard treatment are estimated by fitting a model to the data for the low-risk group and then projecting the results to the high-risk group. The treatment effect is the difference between the observed and estimated survival curves for the high-risk group.

To demonstrate the method of estimation, we use data from another trial in the AIDS Clinical Trials Group series (ACTG 116B/117) that compared continued low-dose AZT (which by then had become the standard treatment) with didanosine in patients infected with human immunodeficiency virus (HIV).[6] In this study, HIV-infected patients who had tolerated AZT for at least 16 weeks were randomized to two different doses of didanosine; the randomly selected control group was continued on AZT. The study included patients who had AIDS and those who had AIDS-related complex (ARC) or were asymptomatic but HIV positive. For the latter group (ARC or asymptomatic), the endpoint was death or the occurrence of an AIDS-defining event. We focus on this group. (The data are available from the National Technical Information Service as ACTG 116B/117, NTIS Order No. PB94-504099.)

Our first step is to specify a model for the risk of an AIDS-defining event or death (collectively, the endpoint). The usual choice would be from the family of Cox models of the form $H_x(t) = H_0(t) \exp \cdot \{BX\}$. In this model, $X$ is a function of the initial CD4 cell count, defined as $X = \ln[(\text{CD4} + 1)/61]$. The reference CD4 value is thus arbitrarily set at 60, with 1 added to prevent a zero value for which the logarithm is undefined. The quantity $H_0(t)$ is the hazard odds at time $t$ for a patient with an initial CD4 cell count at the reference point. Hazard odds are used in discrete time models, and time here is measured in discrete days. The hazard odds are defined as the probability of an endpoint event occurring on day $t$ divided by the probability of an endpoint event occurring after day $t$. The parameter $B$ is the approximate percentage of change in the hazard odds for each change of 0.01 in $X$, or each 1% change in CD4 + 1. The product of the adjusting factor, $exp\{BX\}$, and the reference hazard odds, $H_0(t)$, is the hazard odds, $H_x(t)$, at time $t$ for a patient with a CD4 value of $X$. For a patient with an initial CD4 cell count of 60, $X = 0$ and $H_x(t) = H_0(t)$; as $X$ moves away from the reference point, the odds are adjusted accordingly.

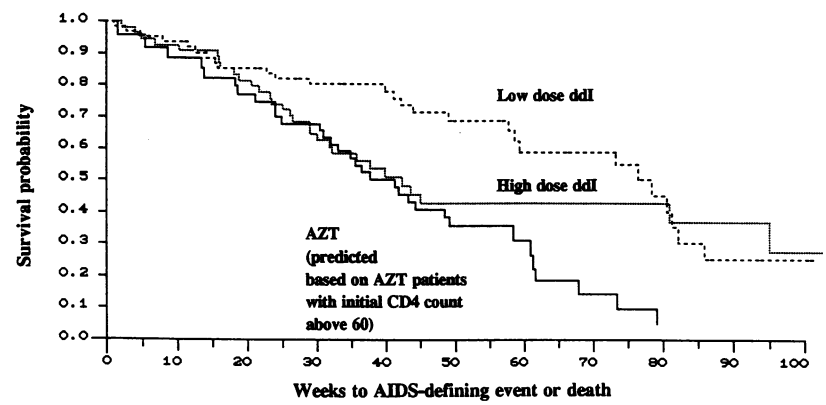
To check the appropriateness of this model, we look at data from the high-dose, low-dose AZT trial previously discussed as our prior experience. The patients in ACTG 002 all had AIDS and the endpoint was death alone, so the experience should be viewed as no more than suggestive for the ACTG 116B/117 patients. Using a CD4 cell count of 60 as the cutoff point defining the high-risk (CD $\leq$ 60) and low-risk (CD > 60) groups, we find that the data are not consistent with the simplest Cox model. The survival curves for the two groups are not separated, as they would be if $B$ were a constant; instead, they converge over time and eventually cross (see Figure 3).

Figure 3 suggests that in ACTG 002, the effect of the different CD4 starting points "wears off" as the population of patients is reduced by death. The same pattern may or may not apply to the data from ACTG 116B/117, but to allow for possible changes in $B$ as time progresses, we make $B$ a time-varying parameter by substituting $B^* = B_0 + B_1 \max[0, (t - 364)/7]$, where $t$ is the number of days since the baseline measurement. This in effect gives $B^*$ two legs: the first year, in which it is constant, and the later period, in which it changes over time. This model




*Note.* Depicted are the observed survival probabilities from the indicated subgroup from the randomized clinical trial: low-dose didanosine (ddI) (n = 63), high-dose didanosine (n = 55), and AZT (n = 60).

**FIGURE 6—Observed times to AIDS event or death comparing AZT and didanosine treatments: 178 AIDS-related complex or asymptomatic patients with initial CD4 cell counts at or below 60/μL in AIDS Clinical Trials Group Study 116B/117.**




*Note.* Depicted is the same information for didanosine (ddI) as shown in Figure 6, with survival probabilities for the AZT patients estimated from a Cox model with a time-varying relative hazard parameter based on 149 AIDS-related complex or asymptomatic patients with an initial CD4 cell count above 60/μL. AZT = zidovudine.

**FIGURE 7—Estimated and observed times to AIDS event or death comparing AZT and didanosine treatments: 178 AIDS-related complex or asymptomatic patients with initial CD4 cell counts at or below 60/μL in AIDS Clinical Trials Group Study 116B/117.**

closely fits the data from ACTG 002. Figure 4 displays the observed survival curves and the average across subjects of the fitted survival curves for the high- and low-risk subgroups.

The next step is to estimate $B_0$ and $B_1$ from the data from the low-risk group in ACTG 116B/117. To do so, we use a standard maximum discrete-time partial likelihood estimation. The results are $B_0 = -0.90689$ (95% confidence interval [CI] = $-1.691, -0.123$) and $B_1 = -0.054357$ (95% CI = $-0.164, 0.055$). Thus, $B_0$ is significant at the usual 5% level, but $B_1$ is not. We keep $B_1$ in the model because ACTG 002 data suggest that it is needed and because leaving it out would involve a further assumption. These estimates imply that a 1% decrease in the initial CD4 count (plus 1) is associated with a 0.9% increase in risk during the first year; thereafter, the risk grows over time although the change is not statistically significant.

Using these values of $B_0$ and $B_1$, we calculate from a logistic regression model the value of $H_0(t)$ at each day on which an endpoint event occurred, with the parameter $H_0(t)$ estimated by maximum likelihood. The hazard odds for the patients in the high-risk group are then obtained from the regression model using their initial CD4 cell counts, and the survival curve for the group is calculated from the hazard probabilities in the usual way. A technical description of the model and the formulas for calculating survival curves are given in the appendix.

Agreement between the model estimates and the observed survival curve for the high-risk group is quite good. As shown in Figure 5, the estimated curve tracks the observed curve quite closely and is well within a 95% confidence interval for the observed curve (calculated by the Greenwood formula[7]).

Both the observed and estimated endpoint curves for AZT, compared with those for didanosine, tell the same story: for the sicker patients, high-dose didanosine confers little if any benefit, whereas low-dose didanosine confers a substantial benefit (compare Figure 6 [randomized trial] with Figure 7 [risk-based allocation trial]). In this example, as in the earlier ones, a trial using risk-based allocation yields essentially the same results for the sicker patients as the randomized trial.

## *Conclusion*

It is clear from our examples that a risk-based allocation scheme can produce estimates of treatment effect that are close to those of a standard randomized clinical trial. Nevertheless, we do not suggest that nonrandom allocation is a preferred way of conducting well-controlled trials. There are distinct disadvantages. The choice of model introduces uncertainty and must be justified; there is the probable loss of double blinding; and it may be difficult to appraise unexpected side effects for which there is no good basis for choosing a model. It is only when poor recruitment, noncompliance, or ethical objections make it difficult or impossible to carry out the conventional randomized design that risk-based allocation is an alternative to be considered. In such cases it should be considered, because randomized trials, if they can be performed at all, may in practice suffer from threats to statistical integrity that are much greater than those that would arise under risk-based allocation designs. □

## 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. We thank David Freedman for his careful reading of and comments on the manuscript, and Brad Efron for kindly supplying the Stanford data from the cholestyramine study.

## References

1. Finkelstein MO, Levin B, Robbins H. Clinical and prophylactic trials with assured new treatment for those at greater risk: I. a design proposal. *Am J Public Health.* 1996;86: 691–695.
2. Lipid Research Clinical Program. The Lipid Research Clinic's Coronary Primary Prevention Trial results, parts I and II. *JAMA.* 1984;251:351–374.
3. Efron B, Feldman D. Compliance as an explanatory variable in clinical trials. *J Am Stat Assoc.* 1991;86:9–26 (with discussion).
4. Fischl MA, Parker CB, Pettinelli C, et al., and the AIDS Clinical Trials Group. A randomized controlled trial of a reduced daily dose of zidovudine in patients with the acquired immunodeficiency syndrome. *N Engl J Med.* 1990;323:1009–1014.
5. Robbins J, Greenland S. Adjusting for differential rates of prophylaxis therapy for PCP in high- vs low-dose AZT treatment arms in an AIDS randomized trial. *J Am Stat Assoc.* 1994;89:737–749.
6. Kahn JO, Lagakos SW, Richman DD, et al. A controlled trial comparing continuing zidovudine with didanosine in human immunodeficiency virus infection. *N Engl J Med.* 1992;327:581–587.
7. Lawless JF. *Statistical Models and Methods for Lifetime Data.* New York, NY: John Wiley & Sons; 1982:74.

## APPENDIX

Let $X$ denote a pretreatment observation used to allocate subjects to either new or standard treatment. Allocation to standard treatment occurs if $X \in A$, where $A$ is a subset of possible $X$ values; allocation to new treatment occurs if $X \notin$ A. In the cholestyramine example, $A = \{X < 290\}$, where $X$ is the first serum cholesterol measurement (in milligrams per milliliter), while in the AIDS trial examples, $A = \{X > 60\}$, where $X$ is the baseline CD4 cell count (in cells per millimeter cubed). Most often, $X$ will be a measure of risk of future disease, an estimate of true disease severity, or a baseline measurement of a disease marker. Let $Y$ denote the posttreatment measurement of the endpoint or response—for example, the final cholesterol level, the number of opportunistic infections in the observed follow-up period, or the time to death or AIDS-defining event (or censoring). Let $X'$ denote an auxiliary pretreatment measurement, which may be a concomitant measure of risk, a replicate measure of $X$ (e.g., a second blood assay), or a baseline version of the endpoint $Y$.

Our goal in an assured allocation trial is to estimate the *treatment effect parameter* $\tau$, defined as the average response for the sicker patients under the new treatment minus the average response for the same patient group under the standard treatment. In symbols,

$$\tau = E_1[Y|X \notin A] - E_0[Y|X \notin A], \quad (1)$$

where $E_1[Y|X \notin A]$ denotes the population average (expected value) of $Y$ for subjects with $X \notin A$ given the new treatment, and $E_0[Y|X \notin A]$ denotes the same for subjects given the standard treatment. Note that any pretreatment variable may be subtracted from $Y$ without changing $\tau$ because the expected value of such a variable is the same under either treatment subsequently delivered. Thus equivalent expressions for $\tau$ are

$$\tau = E_1[Y - X|X \notin A] - E_0[Y - X|X \notin A] \quad \text{or} \quad (1a)$$

$$\tau = E_1[Y - cX'|X \notin A] - E_0[Y - cX'|X \notin A] \quad (1b)$$

for any constant $c$. Henceforth, we omit the subscript on expectations of pretreatment variables.

To fix ideas, suppose there exists a subject-specific parameter, $\theta$, characterizing each subject's true risk or disease severity. For example, $\theta$ might be a subject's long-term average cholesterol level (free of daily fluctuation or measurement error) or a subject's true immune status (of which CD4 cells comprise one component). *A key feature of our analysis is that no parametric assumptions are made concerning how the variable* $\theta$ *is distributed in the population.* However, because patients with $X \notin A$ are not observed under standard treatment in the assured allocation trial, some assumptions about the relations between $X$, $X'$, and $Y$ are necessary to estimate $\tau$. Without the auxiliary variable $X'$, $\tau$ can be estimated under certain parametric assumptions about the distribution of $X$, together with a model for the expectation of $Y$. We refer to this as the semiparametric case (parametric in $X$ and $Y$, nonparametric in $\theta$). With the auxiliary variable $X'$, $\tau$ can be estimated without any parametric assumptions about the distributions of $X$, $X'$, or $Y$ but merely a model for the expected values of $X'$ and $Y$ under standard treatment. We refer to this as the nonparametric case.

To illustrate the semiparametric case, suppose that, given disease severity $\theta$, $X$ follows a Poisson distribution with mean $\theta$, $X|\theta \sim \mathbb{P}(\theta)$, and that standard treatment affects $\theta$ multiplicatively:

$$E_0[Y|X, \theta] = a \cdot \theta \cdot g(X; b), \quad (2)$$

where $g(x; b)$ is a given function of $x$ depending on (at most) an unknown parameter $b$. The case $g(x; b) = 1$ for all $x$ specifies a constant ratio, $a$, of post- to pretreatment means for each subject and was considered in Robbins and Zhang.[1] The more general case of nonconstant $g(x; b)$ allows the magnitude of the standard treatment effect to depend on $X$. (This would apply, e.g., when there is a dose-response relation and dose is based on $X$.) Note that in (2) we do not necessarily assume that $Y$ has a Poisson distribution.

Taking conditional expectations of $\theta$ given $X$ in (2) yields $E_0[Y|X] = a \cdot g(X; b) \cdot E[\theta|X]$, but the form of $E[\theta|X]$ is not known and generally will *not* be linear (unless $\theta$ is assumed, gratuitously, to follow a gamma distribution). Note, though, that for any function $u(x)$, we have $E_0[Yu(X)|X] = a \cdot E[\theta u(X)g(X; b)|X]$, so that, unconditionally, $E_0[Yu(X)] = a \cdot E[\theta u(X)g(X; b)]$. Robbins and Zhang[1] remark that, under the Poisson semiparametric assumption $X|\theta \sim \mathbb{P}(\theta)$, for any function $v(x)$

$$E[\theta v(X)] = E[Xv(X - 1)]. \quad (3)$$

Applying (3) to $v(x) = u(x)g(x; b)$ yields

$$E_0[Yu(X)] = a \cdot E[Xu(X - 1)g(X - 1; b)], \quad (4)$$

in which the right-hand side involves only observable variables. For example, for $j = 0$ and 1, let $u_j(x) = x^j I(x \in A)$, where $I(\cdot)$ is the indicator function for a specified event. Then, using $u_1$ and $u_0$ in (4) and dividing yields

$$\frac{E_0[XYI(X \in A)]}{E_0[YI(X \in A)]} = \frac{E[X(X - 1)g(X - 1; b)I(X - 1 \in A)]}{E[Xg(X - 1; b)I(X - 1 \in A)]}. \quad (5)$$

Given $n$ pairs $(X_i, Y_i)$ for $i = 1, \ldots, n$, we can estimate $b$ by finding that value of $b$, say $b_n$, such that the sample version of (5) is satisfied—namely:

$$\frac{\Sigma_i X_i Y_i I(X_i \in A)}{\Sigma_i Y_i I(X_i \in A)} = \frac{\Sigma_i X_i (X_i - 1)g(X_i - 1; b)I(X_i - 1 \in A)}{\Sigma_i X_i g(X_i - 1; b)I(X_i - 1 \in A)}. \quad (6)$$

It can be shown that if $g(x; b)$ has monotone ratio—that is, if $g(x; b)/g(x; b')$ is monotone increasing (resp. decreasing) in $X$ for any pair of values $b < b'$, then the right-hand side of (6) is monotone decreasing (resp. increasing) in $b$. Thus, there can be only one value $b_n$ satisfying (6). Given $b$, parameter $a$ may be obtained from (4) using either $u_0$ or $u_1$; for example, using $u_0$ we can estimate $a$ by

$$a_n = \frac{\Sigma_i Y_i I(X_i \in A)}{\Sigma_i X_i g(X_i - 1; b_n)I(X_i - 1 \in A)}. \quad (7)$$

The estimates $a_n$ and $b_n$ defined in (6) and (7) are strongly consistent by the law of large numbers. Using these we may now estimate $E_0[Y|X \notin A]$ as follows. Letting $u(x) = I(x \notin A)$ in (4) gives $E_0[YI(X \notin A)] = a \cdot E[Xg(X - 1; b)I(X - 1 \notin A)]$, so that

$$E_0[Y|X \notin A] = a \cdot E[Xg(X - 1; b)I(X - 1 \notin A)]/P[X \notin A], \quad (8)$$

which we estimate by

$$a_n \cdot \frac{\Sigma_i X_i g(X_i - 1; b_n)I(X_i - 1 \notin A)}{\Sigma_i I(X_i \notin A)}. \quad (9)$$

Then $\tau$ may be estimated by subtracting (9) from the sample mean of $Y$ under the new treatment, $\Sigma_i Y_i I(X_i \notin A)/\Sigma_i I(X_i \notin A)$.

A numerical example is provided by the ACTG 116B/117 data with $X$ equal to the baseline CD4 cell count and $A = \{X > 60\}$. Let $Y$ denote the CD4 cell count for those subjects on low-dose AZT who were alive and had measured values at 24 weeks after baseline. Suppose we wish to estimate the 24-week average cell count $E_0[Y|X \leq 60]$ and the average cell count ratio $E_0[Y|X \leq 60]/E[X|X \leq 60]$ for the patients with lowered CD4 cell counts using only response data from subjects with $X > 60$. From previous experience, we choose the response function $g(x; b) = 1 - \exp(-bx)$ to model a linearly declining rate ratio as CD4 cell counts approach zero and a nearly constant rate ratio for large counts. Using data from 127 subjects on AZT with $X > 60$, we estimate $b_n = 0.0257$ from (6), and from (7) we obtain $a_n = 0.847$ (which is the limiting subject-specific rate ratio for large $X$). Using (9), we estimate $E_0[Y|X \leq 60]$ to be 16.2 and the average cell count ratio to be 16.2/31.4 = 0.52. These results compare nicely with the observed values of 16.7 and 0.53, respectively, based on the 66 subjects with $X \leq 60$ on AZT in the randomized trial. Some characteristics of these data are summarized below.

**TABLE A1—Patients on Low-Dose AZT in ACTG 116B/117**

| Group | No. | Avg X (±SE) | Avg Y (±SE) | Avg Y/Avg X |
|---|---|---|---|---|
| CD4 >60 | 127 | 171.4 ± 8.1 | 141.5 ± 9.2 | 0.83 |
| CD4 ≤60 | 66 | 31.4 ± 2.1 | 16.7 ± 1.8 | 0.53 |

**APPENDIX—*Continued***

Analogous semiparametric results are available for other assumptions about the distribution of $X$ given $\theta$. The normal case is treated in Robbins and Zhang.[2]

For the nonparametric case, suppose that $Y$ and $X'$ have conditional expectations, given $X$ and $\theta$, of the form

$$E_0[Y|X,\theta] = \alpha + \beta X + \gamma\theta \quad \text{and} \tag{10}$$

$$E[X'|X,\theta] = \alpha' + \beta' X + \gamma'\theta. \tag{10'}$$

Assumption (10) includes the simple multiplicative ($\alpha = \beta = 0$) and additive ($\beta = 0$, $\gamma = 1$) cases for standard treatment response and allows for a linear dose-response relation with $X$. Assumption (10') specifies that $X'$ is similar to $Y$ in these respects but with possibly different coefficients. For $X'$ to be informative, we require $\gamma' \neq 0$, and because $\theta$ may be arbitrarily scaled, only the coefficient ratio $c = \gamma/\gamma'$ matters. In fact, upon elimination of $\theta$, we have for each $X$

$$E_0[Y - cX'|X] = a + bX, \tag{11}$$

where $a = \alpha - c\alpha'$, $b = \beta - c\beta'$, and $c = \gamma/\gamma'$. We remark that (10) and (10') may be viewed simply as motivation for equation (11), which is actually all that we require to estimate $\tau$. Previous clinical data will ordinarily be used to validate the model (11) over the entire range of $X$ values, and only (11) requires such validation. The assured allocation trial then calibrates and verifies the model empirically among the concurrent controls in the region $X \in A$. *The key assumption is that (11) holds for both* $X \in A$ *and* $X \notin A$. Then it follows that $E_0[Y - cX'|X \notin A] = a + bE[X|X \notin A]$ and thus that $\tau$ in (1b) may be evaluated as

$$\tau = E_1[Y - cX'|X \notin A] - a - bE[X|X \notin A]. \tag{12}$$

Apart from the coefficients $a$, $b$, and $c$, the right-hand side of (12) involves only means of observable variables.

In the cholestyramine example, we assumed that $c = 1$, equivalent to the assumption $\gamma = \gamma' = 1$ of an additive shift in the expectation of $X'$ from initial true cholesterol level $\theta$ to postdiet, pretreatment level $a' + b'X + \theta$, plus another additive shift in the expectation of $Y$ from prerandomization level $a' + b'X + \theta$ to posttreatment level $a + bX + \theta$. The linear dependence of $X'$ and $Y$ on $X$ could reflect a number of causes: the effect of a time lapse between measurements, including the dietary effect between pre- and postdiet cholesterol measurements during the baseline period; a placebo or other psychological effect due to subjects' knowledge of their $X$ values; a regression to the mean effect; or the previously mentioned clinical dose-response relation in cases where the dietary recommendations or standard treatment dosage is based on $X$.

The analysis proceeds by linear regression of $Y - X'$ on $X$, estimating $a$ and $b$ by the ordinary least squares estimators $a_n$ and $b_n$, respectively, based on data $(X_i, X_i', Y_i)$ for $i = 1, \ldots, n$ with $X_i \in A$. The estimate of $\tau$ is

$$\tau_n = \frac{\Sigma_i(Y_i - X_i')I(X_i \notin A)}{\Sigma_i I(X_i \notin A)} - a_n - b_n \mathbb{X}_1, \tag{13}$$

where $\mathbb{X}_1 = \Sigma_i X_i I(X_i \notin A)/\Sigma_i I(X_i \notin A)$ is the sample mean of $X$ among new treatment subjects. The standard error of $\tau_n$, conditional on given $X_i$, is obtained as the square root of the sum of the squared standard error of the mean $Y - X'$ in the new treatment arm and the squared standard error of the estimated regression mean at $\mathbb{X}_1$. The latter term is

$$s_0^2[n_0^{-1} + (\mathbb{X}_1 - \mathbb{X}_0)^2/\Sigma_i(X_i - \mathbb{X}_0)^2], \tag{14}$$

where

$$\mathbb{X}_0 = \Sigma_i X_i I(X_i \in A)/n_0,$$

$$n_0 = \Sigma_i I(X_i \in A),$$

and

$$s_0^2 = (n_0 - 2)^{-1}\Sigma_i(Y_i - X_i' - a_n - b_n X_i)^2 I(X_i \in A).$$

A similar analysis applies for any other assumed value of the constant $c$ When $c$ is unknown, an additional estimation is required; this case will be discussed elsewhere. When the regression of $X'$ on $X$ happens itself to be linear, or if $Y$ is independent of $\theta$ given $X$, then (11) is equivalent to the simple linear regression model of $Y$ on $X$ (set $c = 0$ and omit $X'$). From (10'), though, $E[X'|X]$ would be linear only when $\theta$ follows a conjugate prior, so that $E[\theta|X]$ is linear in $X$.

The two HIV examples in the text are elaborations on conventional models for $Y$ given $X$, like the case $c = 0$ above. For $Y =$ the number of opportunistic infections, we assume a multiplicative rate process such that, given $X$ and follow-up interval of length $t$,

$$E[Y|X,t] = R(X)\cdot t \text{ with } \log R(X) = a + bX. \tag{15}$$

This model would arise, for example, from a Poisson process $Y$ with mean $E[Y|X,t,\theta] = \exp(\alpha + \beta X)\cdot\theta\cdot t$ under the parametric assumption $E[\theta|X,t] = \exp(\alpha' + \beta' X)$. Our goal in this example is to estimate the rate ratio (RR) of the two treatments among the sicker group of patients, conditional on their observed follow-up times. The rate ratio is the total expected number of infections for patients on new treatment to the total expected number of infections on standard treatment:

$$RR = \frac{E_1[\Sigma_i Y_i|X_i \notin A, t_i; i = 1, \ldots, n]}{E_0[\Sigma_i Y_i|X_i \notin A, t_i; i = 1, \ldots, n]}. \tag{16}$$

The numerator of (16) may be estimated directly by the sample mean $M_1 = \Sigma_i Y_i I(X_i \notin A)/\Sigma_i I(X_i \notin A)$. For the denominator, one may estimate $a$ and $b$ in (15) using maximum likelihood estimates $a_n$ and $b_n$ in a Poisson regression model. (One should allow for overdispersion in the Poisson regression model for appropriate standard errors of $a_n$ and $b_n$, although that refinement does not typically affect the point estimates substantially.) Then, a consistent estimate for the denominator of (16) is $M_0 = \Sigma_i\{t_i \exp(a_n + b_n X_i)\}I(X_i \notin A)/\Sigma_i I(X_i \notin A)$, so the rate ratio may be estimated consistently by $RR_n = M_1/M_0$, or, equivalently,

$$RR_n = \frac{\Sigma_i Y_i I(X_i \notin A)}{\Sigma_i\{t_i \exp(a_n + b_n X_i)\}I(X_i \notin A)}. \tag{17}$$

$RR_n$ is in the familiar form of an observed-to-expected ratio, $O_n/E_n$, with $O_n = \Sigma_i Y_i I(X_i \notin A)$ and $E_n = \Sigma_i\{t_i \exp(a_n + b_n X_i)\}I(X_i \notin A)$.

The standard error of $\log RR_n$ is the square root of the sum of the squared standard error for $\log O_n$ and the squared standard error for $\log E_n$. The squared standard error of $\log O_n$ is the squared standard error of the sample sum $O_n$ divided by $O_n^2$. (This term would reduce to $O_n^{-1}$ in a Poisson model for the number of opportunistic infections under the new treatment. To adjust for overdispersion, multiply $O_n^{-1}$ by the variance inflation factor $\chi^2/df$, where $\chi^2$ is the usual chi-squared goodness-of-fit statistic for Poisson data on $df$ degrees of freedom). The squared standard error of $\log E_n$ is given by $V_{aa} + 2\mathbb{X}_1 V_{ab} + \mathbb{X}_1^2 V_{bb}$, where $V_{aa}$, $V_{ab}$, and $V_{bb}$ are, respectively, var $(a_n)$, cov $(a_n, b_n)$, and var $(b_n)$ estimated from the Poisson regression model, and where $\mathbb{X}_1 = \Sigma_i\{X_i t_i \exp(a_n + b_n X_i)\}I(X_i \notin A)/E_n$. The adjustment for overdispersion, if required, is $\chi^2/df$, where $\chi^2$ is the model goodness-of-fit chi-squared statistic, $\chi^2 = \Sigma_i[\{Y_i - t_i \exp(a_n + b_n X_i)\}^2 I(X_i \notin A)]/t_i \exp(a_n + b_n X_i)$ on $df = \{\Sigma_i I(X_i \notin A)\} - 2$ degrees of freedom.

In the survival analysis example, we use a discrete time Cox model for $T$, the time to a primary event (in days from baseline). In ACTG 116B/117, a primary event among the asymptomatic and AIDS-related complex patients was the occurrence of a first AIDS-defining event or death. For $j = 1, 2, \ldots$, the discrete time hazard function is the conditional probability of a primary event at time $T = j$ given survival to time $j - 1$. The hazard function for those with covariate $X$ will be written $P_X[T = j|T \geq j] = P_X[T = j]/P_X[T \geq j]$. We take $X = \log\{(1 + \text{CD4})/61\}$. The reference value $X = 0$ corresponds to those with baseline CD4 cell counts of 60/μL. The discrete time Cox model allows an arbitrary unknown reference *hazard odds*, $P_0[T = j|T \geq j]/P_0[T > j|T \geq j] = P_0[T = j]/P_0[T > j]$, but relates the hazard odds at nonzero values of $X$ to the reference hazard odds via a model for the

**APPENDIX—*Continued***

*log (hazard) odds ratio* at time $j$, $LOR(j)$, where

$$LOR(j) = \log \frac{P_X[T=j \mid T \geq j]/P_X[T>j \mid T \geq j]}{P_0[T=j \mid T \geq j]/P_0[T>j \mid T \geq j]}$$

$$= \log \frac{P_X[T=j]/P_X[T>j]}{P_0[T=j]/P_0[T>j]}. \quad (18)$$

Based on findings of the ACTG 002 trial, we adopt a two-parameter model with a time-dependent log hazard odds ratio:

$$LOR(j) = \{\beta_0 + \beta_1 \max(0, j-364)\} \cdot X = \beta(j) \cdot X. \quad (19)$$

Model (19) specifies a constant hazard odds ratio, $\exp(\beta_0)$, during the first 52 weeks of follow-up, while $\beta_1$ provides for a changing relative hazard odds, $\exp\{\beta(j)\}$, after 52 weeks.

The first step is to estimate $\beta = (\beta_0, \beta_1)'$ using only the data from ACTG 116B/117 patients with baseline CD4 cell counts above 60 on standard treatment (low-dose AZT)—that is, the set $A = \{X > \log 61\}$. We use the exact conditional maximum likelihood estimate, $b = (b_0, b_1)'$, that maximizes the discrete logistic partial likelihood function (accounting for tied event times).

The next step is to estimate the *reference survival function*, $P_0[T > j]$, for $j = 1, 2, \ldots$. This entails estimating the reference hazard log odds parameter, $\alpha_j = \log\{P_0[T=j]/P_0[T>j]\}$, for each observed event time $j$, based on data $(T_i, X_i)$ from those individuals at risk at time $j$. For these we obtain restricted maximum likelihood estimates, $a_j$, in the logistic regression model $\log\{P_x[T=j]/P_x[T>j]\} = \alpha_j + \beta(j)X$, where the coefficients of $\beta(j)$ are held fixed at their conditional maximum likelihood estimates; that is, the values of $\beta(j)$ are fixed at $b(j) = b_0 + b_1 \max(0, j-364)$. Once the $a_j$ are in hand, the baseline survival function is estimated as $S_0(j) = \Pi_{1 \leq k \leq j}[1 - P_0[T=k]/P_0[T \geq k]] = \Pi_{1 \leq k \leq j} 1/\{1 + \exp(a_k)\}$. Similarly, for any nonzero value of $X$, the *adjusted survival function* is estimated by

$$S_X(j) = \prod_{1 \leq k \leq j} 1/\{1 + \exp(a_k + b(k)X)\}. \quad (20)$$

Our goal is now to estimate the average survival function difference

$$\tau(j) = P_1[T > j \mid X \notin A] - P_0[T > j \mid X \notin A] \quad \text{for } j = 1, 2, \ldots. \quad (21)$$

The first term in (21) may be estimated by the usual product-limit estimate of the survival function using observed data from either of the didanosine treatment groups. The second term in (21) may be estimated by the average adjusted survival function under standard treatment for patients with $X \notin A$, to wit:

$$S_*(j) = \sum_i S_{X_i}(j) I(X_i \notin A) \Big/ \sum_i I(X_i \notin A), \quad (22)$$

with $S_{X_i}(j)$ given by (20) with $X = X_i$.

We state a formula for the approximate squared standard error for $S_*(j)$. Let

$$f_j(\underline{\alpha}, \underline{\beta} \mid X) = \exp\left\{-\sum_{1 \leq k \leq j} \log[1 + \exp\{\alpha_k + \beta(k)X\}]\right\}, \quad (23)$$

where $\beta(k) = \beta_0 + \beta_1 c_k$ and $c_k = \max(0, k-364)$. For any covariate value $X$, $f_j$ is a function of the $j+2$ parameters $\underline{\alpha} = (\alpha_1, \ldots, \alpha_j)'$ and $\underline{\beta} = (\beta_0, \beta_1)'$. Let $n = \Sigma_i I(X_i \notin A)$, and let $\ell = 1, \ldots, n$ index the observations with $X_i \notin A$. Then, with maximum likelihood estimates $(\underline{a}, \underline{b})$ for $(\underline{\alpha}, \underline{\beta})$, we have $S_*(j) = n^{-1} \Sigma_{1 \leq \ell \leq n} f_j(\underline{a}, \underline{b} \mid X_\ell)$, and the approximate squared standard error of $S_*(j)$ is given by

$$\text{Var } S_*(j) = \left\{\frac{\partial S_*(j)}{\partial(\underline{\alpha}, \underline{\beta})}\right\}' \text{Cov}(\underline{a}, \underline{b}) \left\{\frac{\partial S_*(j)}{\partial(\underline{\alpha}, \underline{\beta})}\right\}. \quad (24)$$

In (24), $\partial S_*(j)/\partial(\underline{\alpha}, \underline{\beta})$ is the $(j+2) \times 1$ vector of derivatives $(\partial S_*(j)/\partial\alpha_1, \ldots, \partial S_*(j)/\partial\alpha_j, \partial S_*(j)/\partial\beta_0, \partial S_*(j)/\partial\beta_1)'$ evaluated at $(\underline{a}, \underline{b})$ and $\text{Cov}(\underline{a}, \underline{b})$ is the $(j+2) \times (j+2)$ asymptotic variance-covariance matrix of the maximum likelihood estimates based on the data from subjects with $X_i \in A$, also evaluated at $(\underline{a}, \underline{b})$. The derivatives are

$$-\left\{\frac{\partial S_*(j)}{\partial(\underline{\alpha}, \underline{\beta})}\right\} = n^{-1} \sum_{1 \leq \ell \leq n} f_j(\underline{a}, \underline{b} \mid X_\ell) \frac{\partial}{\partial(\underline{\alpha}, \underline{\beta})} \{-\log f_j(\underline{a}, \underline{b} \mid X_\ell)\}$$

$$= n^{-1} \sum_{1 \leq \ell \leq n} f_j(\underline{a}, \underline{b} \mid X_\ell) \Big[p_1(X_\ell), \ldots, p_j(X_\ell), \Big\{\sum_{1 \leq k \leq j} p_k(X_\ell)\Big\} X_\ell, \Big\{\sum_{1 \leq k \leq j} p_k(X_\ell) c_k\Big\} X_\ell\Big]', \quad (25)$$

where for any $X$, $p_k(X) = \exp\{a_k + b(k)X\}/[1 + \exp\{a_k + b(k)X\}]$. For an explicit expression for $\text{Cov}(\underline{a}, \underline{b})$, let $m = 1, \ldots, n(k)$ index the subjects with $X_i \in A$ in the $k$th risk set, and let $w_{km} = p_k(X_m)\{1 - p_k(X_m)\}$. Let $\mathbb{X}_k = \Sigma_{1 \leq m \leq n(k)} w_{km} X_m / \Sigma_{1 \leq m \leq n(k)} w_{km}$, and let $\mathbb{X}$ be the $2 \times j$ matrix

$$\mathbb{X} = \left\{\mathbb{X}_1 \begin{bmatrix}1\\ c_1\end{bmatrix} \cdots \mathbb{X}_j \begin{bmatrix}1\\ c_j\end{bmatrix}\right\}. \quad (26)$$

Then the inverse information matrix is

$$\text{Cov}(\underline{a}, \underline{b}) = \begin{Bmatrix} \text{Diag}\left\{\left(\sum_{1 \leq m \leq n(k)} w_{km}\right)^{-1}; k = 1, \ldots, j\right\} + \mathbb{X}' V(\underline{b}) \mathbb{X} & -\mathbb{X}' V(\underline{b}) \\ -V(\underline{b})\mathbb{X} & V(\underline{b}) \end{Bmatrix}, \quad (27)$$

where for $V(\underline{b})$ we use the inverse information matrix from the discrete logistic partial likelihood function evaluated at $\underline{\beta} = \underline{b}$. $V(\underline{b})$ is the asymptotic variance-covariance matrix for $b$ that is ordinarily used for drawing inferences about $\underline{\beta}$.

To aid interpretation of (24) to (27), consider the case $n = 1$ corresponding to the adjusted survival function estimate (20), given a single covariate value $X$. Then (24) reduces to

$$\text{Var } S_X(j) = S_X(j)^2 \Big[\Big\{\sum_{1 \leq k \leq j} p_k(X)^2 \Big/ \sum_{1 \leq m \leq n(k)} w_{km}\Big\} + r_j(X)' V(\underline{b}) r_j(X)\Big], \quad (28)$$

where $r_j(X)$ is the $2 \times 1$ vector

$$r_j(X) = \sum_{1 \leq k \leq j} p_k(X)(X - \mathbb{X}_k) \cdot (1, c_k)'. \quad (29)$$

The first term in (28) generalizes Greenwood's formula, which covers the case of no covariates. In that case, the maximum likelihood estimates of $p_k$ are $p_k = d(k)/n(k)$, where $d(k)$ denotes the number of failures at the $k$th failure time out of $n(k)$ subjects at risk; and the first term in (28) reduces to the familiar expression for Greenwood's formula for the asymptotic variance of the product-limit estimate of the survival probability $P[T > j]$, viz $S(j)^2 \Sigma_{1 \leq k \leq j} d(k)/[n(k)\{n(k) - d(k)\}]$. When the Cox model is estimated with covariate $X$, the additional term $r_j(X)' V(\underline{b}) r_j(X)$ increases the variance in a manner analogous to the second term of (14) in the context of linear regression. When $X = 0$, (28) provides the squared standard error for the reference survival function $P_0[T > j]$. □

## References


1. Robbins H, Zhang C-H. Estimating a multiplicative treatment effect under biased allocation. *Biometrika.* 1991;78:349–354.
2. 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.