Ethical and Scientific Features of Cutoff-based Designs of Clinical Trials:

A Simulation Study

JOSEPH C. CAPPELLERI, PhD, MPH, WILLIAM M. K. TROCHIM, PhD

Cutoff-based clinical trial designs are geared towards balancing ethical and scientific concerns when it is deemed unethical or infeasible to randomize all patients to study treatments. In a cutoff-based design with randomization, patients who are the least sick based on a quantitative baseline indicator are assigned the control treatment, patients who are the most sick based on the same indicator are assigned to test treatment, and patients who are moderately sick based on the indicator are randomly assigned. Simulations were conducted to examine statistical efficiency and potential bias for designs with varying amounts of cutoff-based assignment and randomization. All design variations yielded unbiased estimates of a main treatment effect and a linear interaction effect. While randomization tends to lead to greater efficiency (or lower standard errors of treatment effect), the correlation between the binary treatment variable and baseline assignment variable completely determines the efficiency of a design. *Key words:* clinical trials; randomized control trials; randomization; cutoff-based designs; experimental designs; quasi-experimental designs; methodology; statistical efficiency; bias. **(Med Decis Making 1995;15:387–394)**

The randomized clinical trial (RCT) is generally regarded as the best available method for assessing the efficacy and toxicity of treatments. Randomization, in theory, serves at least three important purposes: 1) it avoids known and unknown biases on the average in assigning patients to treatment groups; 2) it helps convince others that the trial was conducted properly; and 3) it is the basis for the statistical theory that underlies hypothesis tests and confidence intervals.[1] There has been ample discussion, however, that the traditional RCT may be unethical when strong a priori evidence suggests that the experimental (test) treatment may be more effective than the standard (control) treatment and when the disease under investigation is potentially life-threatening.[2-5] Examples include the recent controversies about the use of extracorporeal membrane oxygenation (ECMO) in neonatal intensive care, the release of drugs for AIDS, and the availability of drugs for cancer treatment that have questioned the ethics of implementing RCTs.[6-8] For such life-threatening diseases, some patients most in need of the presumably more beneficial test treatment, and who are desperate enough to undertake its risk, are denied the test treatment; some patients who are less in need of the test treatment, and who are currently well enough not to chance its side effects, are offered the test treatment.

The Role of Cutoff-based Designs

This article offers an alternative design strategy—the cutoff-based RCT—intended to balance ethical and scientific concerns when it is deemed unethical or infeasible to randomize all patients to study treatments.[9,10] The cutoff-based RCT incorporates a need-based, clinically-valid baseline assignment covariate, which is at least ordinal in scale, so as to combine random assignment with cutoff-based assignment. In a basic cutoff-based RCT, patients who are the least sick, as measured by their baseline scores falling below a cutoff value on a baseline indicator (e.g., severity of illness), are automatically assigned to control treatment; patients who are the most sick, as measured by their baseline scores being above a higher cutoff value, are automatically assigned to test treatment; and patients who are moderately ill, as measured by their baseline scores falling between the two cutoff values, are randomly assigned. The major advantage of the cutoff-based RCT is that it seeks the "best of both worlds"—by allowing a new treatment to be given to those most in need or most willing to assume risk, while capitalizing on some of the efficiency and sta-

Received June 13, 1994, from the Division of Clinical Care Research, Department of Medicine, New England Medical Center, Tufts University School of Medicine, Boston, Massachusetts (JCC) and the Department of Human Service Studies, College of Human Ecology, Cornell University, Ithaca, New York (WMKT). Revision accepted for publication September 27, 1994. Supported in part by grant number RO1 HS-07782, awarded by the Agency for Health Care Policy and Research.

Address correspondence and reprint requests to Dr. Cappelleri: New England Medical Center—Box 63, 750 Washington Street, Boston, MA 02111.

tistical properties of the conventional RCT.

Such a cutoff-based RCT has been used in the Cocaine Treatment Study at the University of California at San Francisco[11] in order to resolve ethical issues as well as to maintain scientific validity. This study, which recruited about 500 men and women aged 21 to 58 living in California, has focused on whether persons dependent on cocaine assigned to inpatient treatment will have better outcomes (and how much better) than they would if they had received outpatient treatment instead. Project staff developed a baseline measure, which was based on clinically sensible criteria, that consisted of four subscales: employment and legal status; family relationships and recovery environment; alcohol and drug use history; and psychological status. These four subscales were weighted and added in a clinically meaningful way to form a baseline composite indicator. Patients who had higher scores on this composite measure were assumed to be more severely addicted to cocaine and hence were more in need of the intensive treatment provided by the inpatient program; low scorers were assumed to manage more easily as outpatients than high scorers. An outcome could be the same or different measure.

For historical and fundamental reasons, this paper also considers the single-cutoff design with no randomization—the regression-discontinuity (RD) design—intended to balance ethical and scientific concerns when it is deemed unethical or infeasible to randomize any patient.[12-15] In the classical RD design, patients who are not as sick (as measured by baseline scores falling below the cutoff value on a baseline indicator) are automatically assigned to control treatment, while those who are more sick (as measured by baseline scores above the cutoff value) are automatically assigned to test treatment. For both the cutoff-based RCT and the RD design, treatment assignment could be constructed so that lower baseline scores indicated more illness (e.g., immunodeficiency scores for AIDS patients).

Besides addressing the ethical dilemma, cutoff-based designs (with or without randomization) may in some situations address a few additional criticisms of conventional RCT.[9] Relative to the conventional RCT, the cutoff-based RCT and RD design may be more consistent with program and public policy objectives, as the test treatment is assigned to the clinically-relevant target population. Because clinically appropriate factors motivate equitable assignment of severely ill patients to test treatment and significantly less ill patients to control treatment, cutoff-based designs may reduce the numbers of post-assignment dropouts and post-assignment "treatment switching." Because the patients who are the most sick and the least sick may be included in a cutoff-based study but may be excluded in a fully randomized study, cutoff-based designs may have their results generalized to a broader base of patients, may be more likely to evaluate relevant confounders, and may offer a lower cost per patient enrolled.[16]

Another important design consideration, especially relevant at this time when national priority has been placed on a more efficient and cost-effective health care system, is the ability of a design to allocate resources more efficiently and to contain cost as the trial progresses. Cutoff-based variations can be constructed, as is done in this paper, to allocate resources more effectively.

The RD design, and especially the cutoff-based RCT, which also preserves some of the benefits offered by randomization, can provide a fruitful area of clinical research for developing and evaluating methods that take into account the clinical realities of treatment delivery, provided that scientific validity is not seriously jeopardized. Trochim and Cappelleri[9] provided an overview of cutoff-based assignment, gave an illustrative example, and conducted Monte Carlo simulations of an analysis-of-covariance model based on six cutoff-based RCT variations, the single-cutoff RD design, and the traditional RCT. The cutoff-based designs yielded unbiased estimates of treatment effects, but the estimates differed in efficiency, with the RCT being most efficient and the RD design being least efficient. Based on unpublished work,[10] we provide a complementary presentation of a different set of cutoff-based RCT variations and extend the population model to include a linear treatment-by-baseline interaction term as well as a main-effect term. We also expand upon ethical considerations.

Methods

SIMULATION OF THE DESIGNS

Monte Carlo simulations were conducted for five interval-related RCT variations and the classic RD design that may be useful in different medical and health research settings. Simulations were also conducted for the traditional RCT so that it could be compared and evaluated with the cutoff-based designs with respect to statistical efficiency and potential bias. These simulations were offered as a way to demonstrate potentially useful alternatives illustrating some key principles of combining different amounts of cutoff-based assignment and randomization.

All of the simulation models used the same variable (e.g., severity-of-illness indicator), measured before and after therapy. Higher scores indicated greater illness. All models were two-group, treatment–control designs. Each model gave an overall probability of assignment of 0.50 to each treatment group. In addition, the five cutoff-based RCT models also imposed a 50/50 assignment rule within the interval of randomiza-

tion. In all simulations, an assignment baseline covariate (X) was generated for each subject such that each value of X was independently and identically normally distributed with a mean of 50 and a variance of 25. Random measurement error in the outcome measure (e_Y) was considered and taken to be normally distributed with a mean of 0 and a variance of 4. The treatment group variable was a dummy-coded variable (Z) with a value of 1 for the experimental (test) treatment and 0 for the standard (control) treatment.

For each model, the outcome variable (Y) was constructed using the equation

$$Y = X + (-5)Z + (-2)(X-50)Z + e_Y$$

which incorporated both a five-point main effect and a two-point interaction effect at the mean score of 50. A (negative) five-point main effect means that all participants in the test-treated group (Z = 1) benefitted by having had their outcomes lowered by five points. A (negative) two-point interaction effect means that subjects in the test-treated group were expected to have a further benefit (reduction) by two points for every one-point increase over 50 in their baseline scores.

All seven models in the regression analysis had their treatment effects estimated at a common point—namely, at the baseline mean of 50—to allow for a fair comparison and evaluation of the seven models. Each of the seven models was simulated 100 times, which led to 700 separate simulation runs, where each run contained a sample size of 1,000.

SPECIFICATIONS AND RATIONALES FOR TREATMENT ASSIGNMENT

The clinical rationales (when appropriate) and the simulation specifications for group assignment are provided in this section. Figures 1–5 depict the five cutoff-interval RCTs.

Model 1—The Traditional RCT

The traditional RCT served as the "gold standard" against which the other models were compared and evaluated. It is the only model discussed here that was based entirely on random assignment. A standard normal variable (i.e., normally distributed variable with a mean of 0 and a variance of 1) was generated and then dichotomized at its mean of zero such that patients with values greater than or equal to zero were assigned to the test-treatment group (Z = 1) and patients with values less than zero were assigned to the control-treatment group (Z = 0).

Model 2—The Classic RD Design

This is the single-cutoff-point model with no randomization and cutoff value at 50. Thus, all patients who scored at or above 50 were automatically assigned to test treatment (Z = 1), while all patients who scored below 50 were automatically assigned to control treatment (Z = 0). It is the only model discussed here based strictly on nonrandom assignment.

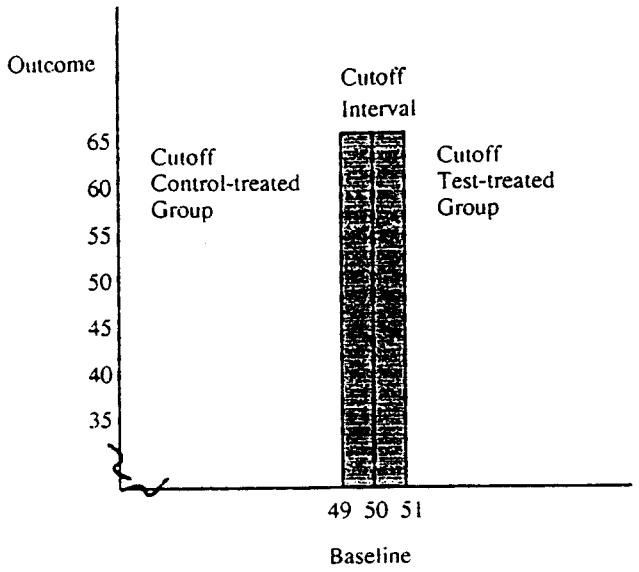

FIGURE 1. Cutoff-based randomized clinical trial with a narrow interval.

Model 3—Cutoff-based RCT with a Narrow Interval (Figure 1)

The assignment criteria in a cutoff-based design must be strictly followed without exception; if they are not, the treatment-effect estimate will be biased.[12,17] Using an interval of randomization instead of a single cutoff value reduces the chance of misassigning some subjects to the other treatment either by mistake or as a result of pressure from physicians, clinical staff, or patients or their families.

Two cutoff values defined the interval of randomization. The cutoff values of 49 and 51 were chosen so that a relatively narrow interval width symmetrically surrounded the cutoff value of 50. About 25% of all patients fell within this interval of randomization. Patients who had baseline scores between 49 and 51 were automatically randomly assigned to either treatment group with equal probability. All subjects who scored 51 or higher were automatically assigned to test treatment, while those who scored 49 or lower were assigned to control treatment.

Model 4—Cutoff-based RCT with a Wide Interval (Figure 2)

Again, two cutoff values defined the interval of randomization. The cutoff values of 48 and 52 were chosen so that a relatively wide interval symmetrically surrounded the cutoff value of 50. About 47% of all patients fell within this randomization interval and hence all subjects with baseline scores between 48 and 52 were randomly assigned. All subjects who scored above

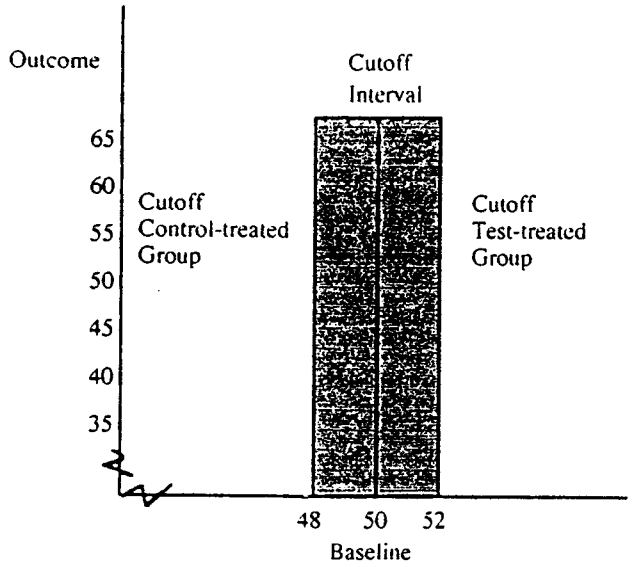

FIGURE 2. Cutoff-based randomized clinical trial with a wide interval.

that interval were assigned to test treatment and those who scored below that interval were assigned to control treatment.

Model 5—Cutoff-based RCT with Five Different Proportions within the Interval (Figure 3)

This model was also a two-cutoff, single-interval RCT. Cases within the interval, however, were randomly divided into five equal subgroups where each subgroup had a different probability of assignment to the new technology (0.25, 0.33, 0.50, 0.66, 0.75), so that about half of all randomized cases were allocated to each treatment. Interval endpoints at 48.5 and 51.5 were selected to bracket a relatively medium-sized cutoff band.

Situations exist where patients accrue over time, the test treatment involves inpatient rehabilitation, and the control treatment involves outpatient rehabilitation for the treatment of, say, cocaine addiction. Hospitals can be put in a difficult financial position if there is either a shortage or a surplus of beds during a prolonged period of time. Hospital administrators can deal with this potential problem by having the flexibility to make adjustments over time in the proportions of incoming patients randomized. When the cutoff strategy places too many patients in hospital beds during a period of time, the proportion of incoming patients randomly assigned to inpatient care can be reduced; conversely, when the cutoff strategy places too few inpatients in beds, the proportion can be increased. Similarly, if the proposed treatment makes more patients ill or makes more patients suffer adverse events than expected, then the proportion of incoming patients randomly assigned to the proposed treatment can be lowered. Another application is when five investigators or five research sites use different probabilities of random assignment, or when there is a single site that over time (e.g., five months) systematically changes (e.g., every month) the probability of random assignment as patients accrue and more treatment information is gathered.

Model 6—Cutoff-based RCT with Two Interval Widths (Figure 4)

This design variation explored the possibility of using *two* random-assignment intervals with different widths. One cutoff interval had its endpoints at 48.5 and 51.5; the other cutoff interval had its endpoints at

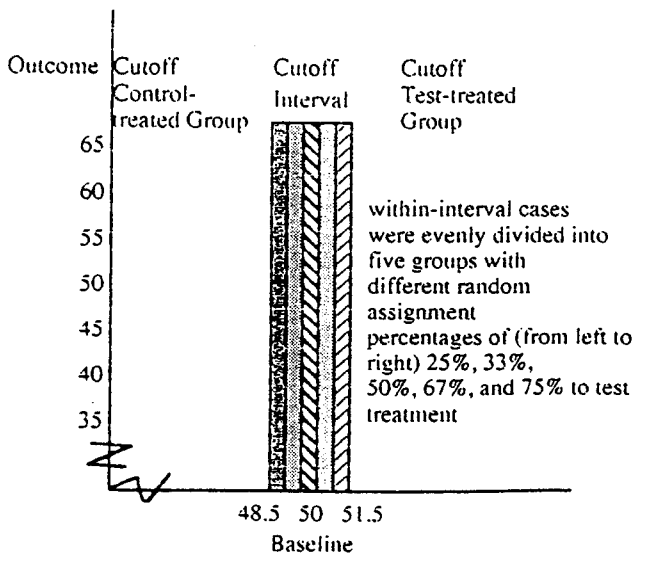

FIGURE 3. Cutoff-based randomized clinical trial with five different proportions within the interval.

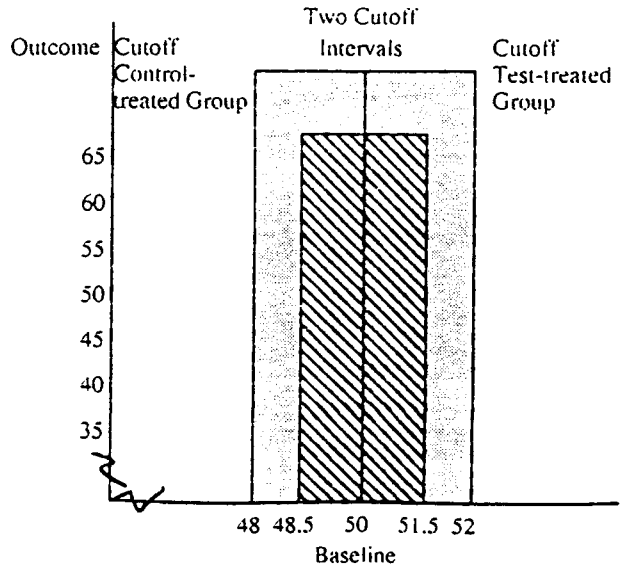

FIGURE 4. Cutoff-based randomized clinical trial with two interval widths.

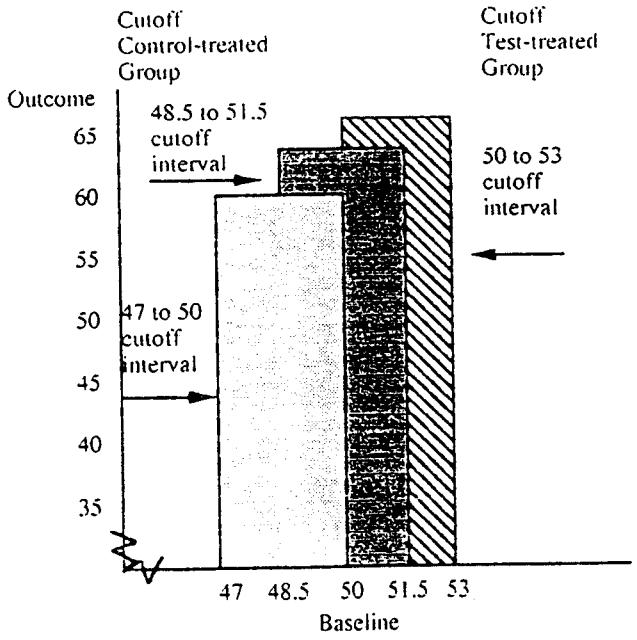

FIGURE 5. Cutoff-based randomized clinical trial with three interval centers (48.5, 50.0, 51.5).

48 and 52. Both intervals were symmetric around the baseline mean of 50. Because the two intervals had different widths, they included different numbers of randomized patients, with the wider interval containing more randomized patients.

One clinical rationale for Model 6 stems from a potential need for the study's investigators to examine the effects of using different interval widths at different treatment sites or at different points in time as patients accrue. This type of design may be useful for different hospital sites with varying amounts of resources to accommodate a novel treatment or an experimental inpatient therapy. Like Model 5, Model 6 may also prove beneficial for a given site where periodic adjustments over time are desired in the number of incoming patients randomized. In the early months of the study, for instance, the interval of randomization could be fairly wide. As evidence for the efficacy of the new treatment becomes stronger, the width of the interval could be periodically compressed in order to place more of the sickest patients who are accruing in the currently more efficacious treatment.

Model 7—Cutoff-based RCT with Three Interval Centers (Figure 5)

This design variation involved moving the entire interval up or down on the baseline scale while preserving the same interval width. Each interval could be asymmetric with respect to the baseline mean. Model 7 incorporated three cutoff windows where each interval covered three units on the baseline scale. The cutoff intervals were taken at [47, 50], [48.5, 51.5], and [50, 53], with their respective interval centers at 48.5, 50, and 51.5. The baseline value at which treatment effects were estimated was 50, the average of the three interval centers.

If there are too few inpatients occupying hospital beds during one period or if evidence accumulates that even more strongly favors the test treatment, the entire interval could be moved down on the baseline assignment scale, thereby allowing for a greater area under the normal curve at the right of the interval for assigning incoming patients to inpatient treatment or test treatment. Conversely, if there are too many patients occupying beds during a period or if evidence accumulates that favors the control treatment, the entire interval could be windowed up. In a multi-site study, a situation may arise when different sites with different resources can handle different proportions or "mixes" of treatment and control cases. A particular facility, for instance, may have more surgeons who specialize in the novel test treatment and thus may be able to have a higher proportion of cases assigned to that condition.

Results

Table 1 shows the simulation results. The results that appear are average-treatment-effect estimates and their average standard errors, averaged over 100 runs. The simulations empirically demonstrate that all the models yielded unbiased estimates of both a (minus) five-point main treatment effect and a (minus) two-point interaction effect.

Efficiency, as measured by the average standard errors, was more of an issue than bias for these generated models. Table 2 shows the ranks of the average standard errors, where a rank of 1 signifies the lowest standard error and 7 signifies the highest, along with the percentage of cases randomly assigned and the average correlation coefficient between baseline and treatment group variable, denoted by R(X, Z). This correlation is one way to determine the nature and degree of multicollinearity, which causes the standard error of treatment effect to become inflated, inherent in all of the proposed models except the RCT model (Model 1). The ranks for the standard errors of a main effect

Table 1 • Simulation Results of the Models

	Average Treatment Effect		Average Standard Error	
	Main	Interaction	Main	Interaction
Model 1	−5.010	−2.000	0.063	0.020
Model 2	−5.000	−2.000	0.105	0.033
Model 3	−5.011	−2.003	0.097	0.030
Model 4	−5.002	−2.001	0.083	0.025
Model 5	−5.002	−2.002	0.090	0.028
Model 6	−4.996	−2.002	0.086	0.027
Model 7	−4.990	−2.002	0.085	0.026

Table 2 • Ranks of Standard Errors, Percentages Randomized, and Correlations for the Models

	Rank of Standard Error	Percentage Randomized	R(X, Z)*
Model 1, traditional RCT	1	100.00	.00
Model 4, wide interval	2	47.14	.65
Model 7, three different centers	3	33.98	.67
Model 6, two widths	4	41.65	.68
Model 5, five different proportions	5	36.16	.71
Model 3, narrow interval	6	24.72	.75
Model 2, classic regression-discontinuity	7	0.00	.79

* Average correlation coefficient between baseline (X) and treatment group (Z) variables.

were in full agreement with the ranks for the standard errors of an interaction effect.

Table 2 shows that those models that had more individuals randomly assigned tended to achieve greater efficiency (or lower standard errors of treatment effect). But what completely determines efficiency is the correlation coefficient, R(X, Z), of each model. The design structure of Model 7 reduces the linear association between X and Z, as measured by R(X, Z), slightly more than do the design structures of Models 5 and 6—even though the latter two models had more patients randomly assigned. It should be noted, though, that Model 7 (with three interval centers) was keenly sensitive to the point where the treatment effect was estimated. If each of the three intervals used its own interval center (48.5, 50.0, 51.5) to estimate the treatment effects, with the three separate results then pooled into one overall analysis, then the same design (although still unbiased) would have actually rendered the largest standard error.

Discussion

This simulation study suggests that the interrelationship between the baseline measure and the treatment group variable in cutoff-based variations reflects the precision of their treatment-effect estimates. The intrinsic way that a cutoff-based RCT reduces this interrelationship takes precedence over the percentage of randomized cases. It is worthwhile to note that more randomization of one design than another design does not necessarily imply that the design with more randomization is more efficient. A reduced correlation, however, is usually associated with an increased percentage of randomized cases; everything else being the same, more randomization necessarily leads to greater efficiency.

Cappelleri and colleagues[18] found that in order to achieve the same power as the conventional RCT, approximately 2.75 times more patients are needed in the RD design; 2.50 times more patients are needed in the (one-interval) cutoff-based RCT with 20% randomization; 2.10 times more patients are needed in the cutoff-based RCT with 35% randomization; and 1.70 times more patients are needed in the cutoff-based RCT with 50% randomization. Therefore, the lower power and efficiency of cutoff-based designs could increase rather than decrease the complexity, duration, or expense of controlled clinical trials.[19]

The most critical step in the statistical analysis of cutoff-based designs is correctly modeling the true functional relationship between the baseline covariate and the outcome variable,[12] which makes the statistical analysis more challenging. Because data for both groups are not available across the entire range of the baseline–outcome distribution, cutoff-based designs require regression lines to be extrapolated into the region where there are no observed data. Cappelleri and Trochim[20] mentioned ways to help arrive at the correct functional form. One recommended strategy is to use a polynomial, backward-elimination regression approach that tests interaction effects before main effects. An incorrect functional form will lead to a biased treatment effect.[12,13,17] A biased estimate will undermine scientific validity and will itself inflict a serious ethical consequence that may outweigh other ethical considerations. In the set of simulations, the true functional form between baseline and outcome was correctly modelled with linear terms.

The same set of regression assumptions (and remedies for their violations) used in the traditional RCT also applies to the cutoff-based design. Like the traditional RCT model, the cutoff-based regression model can have a binary outcome.[21,22]

If this set of simulations incorporated random measurement error in the observed baseline covariate (X), the linear interaction effect in both the cutoff-based designs and the traditional RCT would be attenuated by a factor of the reliability coefficient, but the main-effect estimates with both strategies would remain unbiased.[23,24] Random measurement error in the baseline indicator (and the regression to the mean that results) is not a problem in the cutoff-based design because, like the conventional RCT, it incorporates a perfectly known assignment rule. And this rule is fully modeled and hence accounted for in the cutoff-based analysis. The simulation results on unbiased estimates of main effect confirm the theoretical work in Goldberger[25,26] and Rubin[27,28] that analytically proved that the main-effect estimate is unbiased in cutoff-based designs, whether or not a true linear interaction term is included and whether or not the baseline is measured with random error.

In the United States, there is an ethical need as well as a legal requirement to inform all participants about the clinical trial and to obtain their written consent before randomization or treatment allocation. The

randomized, cutoff-assigned treatment, and cutoff-assigned control groups in a cutoff-based RCT should have their own separate informed consent forms, with each form going into specific details. This would make administrative and logistical matters simpler and clearer.

Each of the three consent forms, however, should cover the general elements of informed consent,[29] including full disclosure of the method of treatment assignment. All prospective participants should be told that the patients who are sickest will receive the experimental treatment, those who are the least sick will receive the standard treatment, and those who are moderately ill will be randomized to either treatment. They should also be informed that the benefit of the experimental treatment appears promising but its true relative benefit remains inconclusive, which is why the study is being undertaken. Patients receiving the standard treatment should be assured that they are receiving the best proven therapy available to them in any clinical setting where the intervention is still viewed as experimental. These patients should therefore be likely to give their informed consent.

Each patient (and physician) will know which one of the three risk profiles he or she belongs to before being given a consent form. Only randomized patients and their physicians will be blinded, so a cutoff-based RCT would be a partially-blinded design. It is the life-saving potential of the intervention that may make the traditional view of the patient–physician relationship powerful enough to override the benefit of the most rigorous and efficient properties afforded by a double-blinded RCT. The resulting ethical obligation placed on physicians to their critically ill patients, who are extraordinarily dependent on their physicians, will then take precedence over societal benefits.

Singer[30] discussed problems with using stopping rules as a way to optimize or balance the scientific value to society. In the context of using warfarin to prevent stroke in nonrheumatic atrial fibrillation, Singer criticized stopping a RCT early because the observed treatment effect may be made more imprecise and may result in possible bias towards a larger benefit than if the RCT had continued and more events had accumulated. A cutoff-based design (especially a cutoff-based RCT) without stopping rules can, in principle, achieve unbiasedness and provide more precise estimates than a traditional RCT with stopping rules.

Truog[31] critiqued the ethical dilemmas raised by the Harvard Neonatal ECMO Trial. Troug concluded that the ideal study design for the proposed pediatric ECMO trial should both maintain the essence of the traditional patient–physician relationship and provide patients in the trial with the opportunity to be offered other potentially life-saving and rapidly developing technologies. He proposed a prospective observational study with physicians completely determining the treatment(s) for their patients. By offering potential advantages in terms of ethics, timeliness, attrition, and flexibility of assignment to study treatments as patients accrue, as well as the ability to offer an unbiased estimate of treatment effect, a cutoff-based design could be a suitable alternative in a pediatric ECMO trial.

Cutoff-based designs should not replace the traditional RCT in the great majority of circumstances when there are no objections from either research subjects or clinicians stemming from ethical conflicts regarding the treatments under investigation. Creative approaches that are completely randomized and ethically sensitive should be employed whenever possible, as was done in a study of the chronic left ventricular assist system.[32] But some believe[31,33] that the balance shifts when there is strong (though still inconclusive) evidence mounting that the experimental therapy is potentially life-saving. It is in these situations where ethical issues may prevent complete randomization from being employed that a cutoff-based design, especially a cutoff-based RCT, should be a very attractive alternative. As such, although we compared cutoff-based designs with the traditional RCT—the "gold standard"—they should be more fairly compared and evaluated against nonrandomized designs. When compared with nonrandomized designs, cutoff-based designs (especially the cutoff-based RCT) are among the most scientifically valid and ethically sensitive options available.

The authors are grateful to Harry P. Selker, MD, MSPH, for encouraging them in this endeavor and for his editorial comments. They also appreciate the valuable input from the reviewer, Judy Tesnow, and Joni R. Beshansky, RN, MPH.

References

1. Green S. Patient heterogeneity and the need for randomized clinical trials. Controlled Clin Trials. 1982;3:189-98.
2. Beecher H. Ethics and clinical research. N Engl J Med. 1966;274:1354-60.
3. Veatch R, Sollitto S. Human experimentation: the ethical questions persist. Hastings Center Per. 1973;3:1-3.
4. Schaffner KF. Ethical issues in the use of clinical controls. J Med Philos. 1986;11:297-404.
5. Miké V. Suspended judgment: ethics, evidence, and uncertainty. Controlled Clin Trials. 1990;11:153-6.
6. Ware JH. Investigating therapies of potentially great benefit: ECMO (with discussion). Stat Sci. 1989;4:298-340.
7. Marshall E. Quick release of AIDS drugs. Science. 1989;245:346-7.
8. Marx JL. Drug availability is an issue for cancer patients, too. Science. 1989;245:345-6.
9. Trochim WMK, Cappelleri JC. Cutoff assignment strategies for enhancing randomized clinical trials. Controlled Clin Trials. 1992;13:190-212.
10. Cappelleri JC. Cutoff-based designs in comparison and combination with randomized clinical trials. Doctoral dissertation, Cornell University, Ithaca, New York, May 1991.
11. Havassy BE, Wesson BR, Tschann JM, Hall SM, Henke CJ. Efficacy of cocaine treatments: a collaborative study. NIDA Grant Number

DA05582 awarded to UCSF; 1988.
12. Trochim WMK. The regression–discontinuity design. In: Sechrest L, Perrin P, Bunker J, eds. Research Methodology: Strengthening Causal Interpretations of Nonexperimental Data. Washington, DC: Agency for Health Care Policy and Research, U.S. Public Health Service, 1990.
13. Cook TD, Campbell DT. Quasi-experimentation: Design and Analysis Issues for Field Settings. Boston, MA: Houghton–Mifflin, 1979.
14. Campbell DT. In: Overman ES, ed. Methodology and Epistemology for Social Science. Chicago, IL: University of Chicago Press, 1988.
15. Mohr LB. Impact Analysis for Program Evaluation. Chicago, IL: Dorsey Press, 1988.
16. Luft HS. The applicability of the regression discontinuity design in health services research. In: Sechrest L, Perrin P, Bunker J, eds. Research Methodology: Strengthening Causal Interpretations of Nonexperimental Data. Washington, DC: Agency for Health Care Policy and Research, U.S. Public Health Service, 1990.
17. Trochim WMK. Research Design for Program Evaluation: The Regression-discontinuity Approach. Beverly Hills, CA: Sage Publications, 1984.
18. Cappelleri JC, Darlington RB, Trochim WMK. Power analysis of cutoff-based randomized clinical trials. Eval Rev. 1994;18:141-52.
19. Williams SV. Regression–discontinuity design in health evaluation. In: Sechrest L, Perrin P, Bunker J, eds. Research Methodology: Strengthening Causal Interpretations of Nonexperimental Data. Washington, DC: Agency for Health Care Policy and Research, U.S. Public Health Service, 1990.
20. Cappelleri JC, Trochim WMK. An illustrative statistical analysis of cutoff-based randomized clinical trials. J Clin Epidemiol. 1994;47:261-70.
21. Berk RA, Rauma D. Capitalizing on nonrandom assignment to treatment: a regression-discontinuity of a crime control program. J Am Stat Assoc. 1983;78:21-8.
22. Carter GM, Winkler JD, Biddle AK. An evaluation of the NIH research career development award. Santa Monica, CA: Rand Corporation, 1987.
23. Cappelleri JC, Trochim WMK, Stanley TD, Reichardt CS. Random measurement error doesn't bias the treatment effect estimate in the regression-discontinuity design: I. The case of no interaction. Eval Rev. 1991;15:395-419.
24. Trochim WMK, Cappelleri JC, Reichardt CS. Random measurement error doesn't bias the treatment effect estimate in the regression-discontinuity design: II. When an interaction effect is present. Eval Rev. 1991;15:571-604.
25. Goldberger AS. Selection bias in evaluating treatment effects: Some formal illustrations. Discussion paper #123. Madison, WI: Institute for Research on Poverty, 1972.
26. Goldberger AS. Selection bias in evaluating treatment effects: The case of interaction. Discussion paper #129. Madison, WI: Institute for Research on Poverty, 1972.
27. Rubin DB. Assignment to groups on the basis of a covariate. J Educ Stat. 1977;2:1-26.
28. Rubin DB. Bayesian inference for causal effects: the role of randomization. Ann Stat. 1978;6:34-58.
29. Meinert CL. Clinical Trials: Design, Conduct, and Analysis. New York: Oxford University Press, 1986.
30. Singer DE. Problems with stopping rules in trials of risky therapies: the case of warfarin to prevent stroke in atrial fibrillation. Clin Res. 1993;41:482-6.
31. Truog RD. Randomized controlled trials: lessons from ECMO. Clin Res. 1992;40:519-27.
32. Parker LS, Arnold RM, Meisel A, Siminoff LA, Roth LH. Ethical factors in the allocation of experimental medical therapies: the chronic left ventricular assist system. Clin Res. 1990;38:537-44.
33. Hellman S, Hellman DS. Of mice but not men. Problems of the randomized clinical trial. N Engl J Med. 1991;324:1585-9.