PERSPECTIVES ON THE STATISTICIAN'S ROLE IN COOPERATIVE CLINICAL RESEARCH

NORMAN BRESLOW, PHD

Some contributions which the statistician can make to cooperative research are briefly discussed in the hope of furthering communication between statistician and clinician. In the planning stages the major issues are those of sample size, randomization, data collection and early termination. Careful consideration of these points can help ensure that the main questions posed by the trial are answered as economically as possible. During the final stages the major statistical contribution is in analysis and presentation of results in a way which explores the interrelationships between the most important baseline and therapeutic factors and their joint effect on outcome.

Cancer **41:326–332, 1978.**

DURING THE PAST SEVERAL DECADES CLINICAL investigators have become increasingly familiar with the principles of controlled clinical trials as enunciated by the eminent statistical scientist Sir Austin Bradford Hill.[25,26,27] There has been a corresponding evolution in the statistician's role in cooperative research. Formerly called in as an occasional consultant on sample sizes and chi-squares, he is currently more likely to be a full team member sharing major responsibility for study design, data collection and monitoring, and statistical analysis.

Drawing largely on my experience as statistician to the Children's Cancer Study Group (CCSG) and the National Wilms' Tumor Study (NWTS), this paper comments briefly on several areas in which the statistical input to cooperative research has been most important. Specific methodologic techniques are not discussed here as these are available in a recent extensive review by a group of British and American statisticians[34,35] as well as in other cited references. Rather the intent is to present a certain perspective on the statistician's role, with the idea that communication between clinician and statistician may be improved if each has a fuller understanding of the other's viewpoint.

From the Department of Biostatistics, University of Washington, Seattle, Washington.

Supported in part by USPHS Grant CA 11722 from the National Cancer Institute.

Professor of Biostatistics, School of Public Health, University of Washington, Seattle, Washington.

Address for reprints: Dr. Norman Breslow, Department of Biostatistics SC-32, University of Washington, Seattle, Washington 98195.

Accepted for publication May 10, 1977.

HOW MANY SUBJECTS?

One of the first questions facing the statistician in planning a new study concerns the size of the sample since this can indicate the feasibility of carrying out the investigation in reasonable time. Sample size calculations are relatively routine once the goals of the investigation have been explicitly stated.[20] In early (Phase II) studies of new drugs the goal may be simply to screen out those which have demonstrable value, at some minimal arbitrary level of response, from those that do not. Providing that the "spontaneous" response rate from historical series is known to be zero, this is quite easy: a single favorable outcome in a series of proven cases constitutes evidence of treatment effectiveness. On the other hand, if there are no responses in a series of fourteen tries, one can rule out at the 95% level of confidence the possibility that the true response percentage would be 20% or better in an indefinite series of similar cases. The only difficulty in applying this notion is that, when effective therapy does exist, the fourteen subjects selected for studies of a new agent tend to be a heterogeneous collection of patients who have failed prior treatment. An alternative strategy might be to introduce new agents earlier in the clinical course, for example among patients in prognostic categories for which conventional therapy offers little benefit.

The major efforts of cooperative groups have not been in small Phase II studies but rather in the comparison of demonstrably effective agents with other established treatments (Phase III trials). Here statisticians distinguish two types of error in interpreting results: Type I, the error of

0008-543X-78-0100-0326-0085

concluding that the treatment groups differ when in fact they do not; and Type II, the error of concluding "no significant difference" when in fact there is a clinically important difference. The risks, or probabilities, of making Type I and II errors are called α and β: α is known also as the significance level; $1-\beta$ as the power. The process of arriving at values for α and β can be viewed as a contest between statistician and investigator. The statistician, who generally insists on strong objective evidence to support positive results, asks that the investigator only herald his findings when in fact these findings would have been quite unlikely under the null hypothesis that the treatments were the same. The usual criterion is to accept results as reasonable evidence in favor of one treatment over the other providing not more than 5 out of 100 (α = .05) repeat trials would yield similar results under the null hypothesis. Having agreed to this somewhat arbitrary and unreasonable stipulation, possibly under pressure from journal editors, the investigator then charges the statistician with ensuring that the trial have a good chance of meeting the 5% criterion when in fact the treatments differ by a prescribed magnitude. It is in his interest, as well as that of science in general, that such a difference be detected if it exists. Since the risk of *not* detecting the difference decreases as the number of subjects increases, the question of sample size is resolved by specifying a value for β in addition to that already chosen for α.

When observations are costly, or patients scarce, it is sometimes necessary to make compromises. A not uncommon choice in clinical trials is β = .20 to accompany the α = .05, so the statistician has apparently chosen to compromise the investigator's interests rather than his own. However, it should be borne in mind that such an analysis is made only to give a rough idea of the requisite n, that it is actually very much in the investigator's interest to keep α small (that way he has a better chance of convincing his colleagues), and that he is still afforded a good (80%) chance of detecting the prescribed difference and an even better one of detecting a more extreme difference.

Suppose the treatment groups are compared simply in terms of percentage response. Then, using the above specifications and assuming that the sign of the difference is unimportant, 164 patients in *each* of two groups are required to detect a difference of 55 versus 70%, 424 in each to detect a difference of 90 versus 95%. These figures, taken from a convenient set of tables prepared by Natrella,[33] give some indication of why a multi-institution cooperative effort is needed to complete trials of relatively rare diseases in reasonable time.

WHEN CAN WE STOP?

In addition to making projections at the start of a trial regarding its ultimate size, one must consider also the question of early termination. Sometimes this is because one or more regimens turn out to be too toxic, or because a new treatment is introduced which renders the trial obsolete. More problematic is when differences between regimens start to appear sooner than anticipated. At this point the participants may be faced with a real conflict of interest: of choosing to treat their patients with what appears at the moment to be the best agent *versus* continuing to assign some patients to an apparently inferior regimen in the hopes that the information gained will be of benefit to future patients with the same disease. If one is 70% sure that drug A is better than drug B, that may suffice in choosing drug A for a single patient; however, one would want to be very much more certain if the choice were to be applied to a population of 1,000 patients.

One technique introduced in large scale heart disease trials to ease the ethical dilemma of participants, and to relieve pressure for early termination, is simply to keep the interim results of the trial secret.[24] An external monitoring committee is given sole access to the data as they accrue and is charged with deciding when sufficient information has been collected to justify closing the trial. Of course this committee itself must come to grips with the issue of just how much information that is.

Statisticians have in fact developed formal stopping rules for clinical trials which are based on consideration of Type I and II errors similar to those just outlined.[2] Considerable debate has taken place within the statistical community as to the appropriateness of such designs.[1,12,15] Part of the problem stems from the fact that the arbitrary choices of significance level and power, formerly used to provide broad guidelines as to feasibility, are now translated into inflexible decision rules. A sequential design of this type,[5] for example, was proposed for the NWTS.[16] The reason was a desire not to withhold prophylactic radiation if it should become evident that such therapy was indeed of significant value, even for patients with completely localized and resected disease. During the course of the study, the main

criterion for evaluation was changed from relapse in any site to local recurrence in the treated flank, which rendered the original design inoperable. Increasing awareness of the long term sequelae of radiation therapy eventually led to a desire to terminate the trial for precisely the opposite reason as originally envisaged, *viz.* that some patients were being needlessly irradiated.

In view of the difficulties of applying formal stopping rules, and philosophical disputes about their relevance, attempts have been made to develop other methods of rationalizing the decision to terminate or modify a trial. One of these[13] involves the calculation of relative betting odds in favor of one or another treatment. Together with other relevant information, for instance regarding toxicities, these are used informally by the monitoring committee in its deliberations. However, it is fair to say that the issue of when precisely to stop a trial has not yet been resolved to everyone's satisfaction.

WHY RANDOMIZE?

Without doubt the keystone of the modern clinical trial is its insistence on concurrent randomized controls.[27] While the hazards of dispensing with such controls, of relying on collections of case reports or on comparisons with historical series, have been articulated on many occasions, they nevertheless seem to need frequent and repeated emphasis. It has been estimated that 60% of patients in cancer chemotherapy trials reported in three major journals during 1970–1972 were entered in "uncontrolled" studies.[3] A typical example is that of L-asparaginase, which early on was given very enthusiastic reports yet has failed so far to fullfill this promise.[11] The optimistic results achieved with small nonrandomized trials of new agents are often followed by more sober assessment in properly conducted investigations.

The object of using concurrent randomized controls, of course, is to ensure that the groups of patients being compared are as similar as possible in all respects except the treatments administered. Some researchers[40] have questioned whether unrestricted randomization is the best way to achieve this goal. Formation of relatively homogeneous patient strata, on the basis of prognostic factors such as age and stage of disease at diagnosis, and restricted randomization so as to achieve equal numbers of treatment and control patients within each strata, may achieve greater comparability.[42] Unfortunately, this method becomes unworkable if the number of factors increases to the point that there are scarcely more patients than strata. While new methods of "data-dependent" treatment allocation[37,40] enable larger numbers of such factors to be accounted for, these can be difficult to administer and tend to invalidate the assumptions of random sampling which underly conventional statistical procedures. A reasonable practical compromise is simply to balance the numbers of patients assigned to treatment and control arms within each of the cooperating institutions. The effects of prognostic factors may be accounted for when the data are analyzed by "retrospective stratification," with little loss of information from what would have been achieved by perfect balance.[34]

Some researchers[18,21] have questioned whether concurrent controls, however randomized, are a necessary ingredient of all trials. They argue that our current knowledge of prognostic factors is sufficiently precise to enable "comparable" control patients to be selected from historical series. A drawback to this approach is that the information needed to select such historical controls simply may not be available. As a case in point, the attempt to compare recent CCSG experience with metastatic neuroblastoma to an earlier series[41] faltered because "insufficient data (were) available to use current staging criteria even if the original hospital records were consulted."[29] With rapid advances in diagnostic techniques leading to greater understanding of the disease, it becomes less likely that truly comparable controls can be selected from past experience. While the use of historical controls may be advocated in a few clearly defined circumstances,[3] many statisticians agree that randomization remains the most reliable method for evaluating therapies.[9,34]

KILLING TWO BIRDS WITH ONE STONE

Little use seems to be made of factorial designs in medical work, although they are widely used in agriculture, industry, and other branches of scientific inquiry. This is due in part to a desire to keep protocols as simple as possible. It also attests to the difficulty of getting a large number of participants to agree to randomize patients to more than two or three treatment regimens. Nevertheless, certain therapeutic questions would seem to cry out for the use of such designs.

The concept of a factorial design is illustrated in Figure 1. In this hypothetical study it is de-

sired to determine the value of radiation therapy, and also to compare two chemotherapeutic agents A and B. Rather than conduct separate studies, one of radiation and another of chemotherapy, one may address the two questions simultaneously to the same group of patients. The sample size need be scarcely greater than that used in either two arm study. Comparison of regimens 1 versus 2 and 3 versus 4 tests the radiation question, while 1 versus 3 and 2 versus 4 tests chemotherapy. Moreover this is the only type of design which provides for the detection of synergism (or antagonism) in the therapeutic or toxic effects of the two modalities.

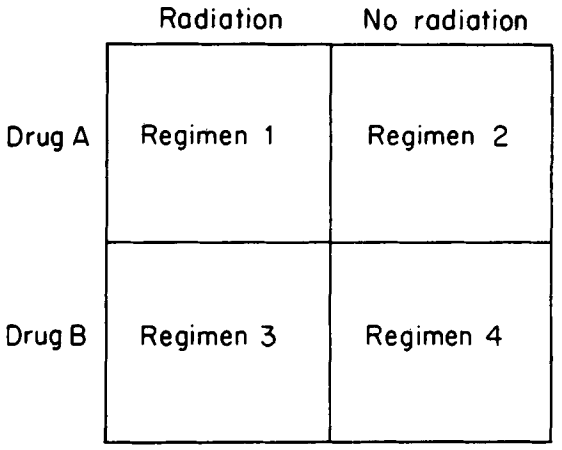

FIG. 1. A 2 × 2 factorial design for the simultaneous study of radiation and chemotherapy.

MAINTAINING DATA QUALITY

Working in close consultation with medical colleagues, the statistician can make valuable contributions to protocol writing and forms design. These should encourage adherence to the agreed plan of the study, facilitate recording of the data, and ensure that they are ultimately analyzable. Clarity is the key to definitions and descriptions of procedures. Forms can be constructed so as to reinforce the protocol *vis-á-vis* eligibility criteria, treatment schedules, required observations and their timing. Increasing use is made of "pre-coded" forms, meaning simply that the rules for transcribing data into numbers have been specified in advance and incorporated in the form. The quality of response on important items is improved if one avoids the "shotgun" approach to data collection, and concentrates on the main issues. With the growing capacity of data storage devices, it is feasible and desirable to record, store and analyze original measurements rather than derived "scores" for such items as blood chemistries and size of lesions. Grouping of these measurements into categories can be done later, after preliminary data analyses, rather than fixing them arbitrarily in advance.

STATISTICS AND COMPUTERS

A common misunderstanding of the statistician's role concerns his relation to the computer. While modern statistical methods place increasing reliance on the computational power of electronic computers, data processing *per se* is a speciality field quite separate from that of statistical methodology. In order to meet his professional responsibilities, of course, the statistician must be sufficiently conversant with computer hardware and software that he can coordinate the activities of data processing personnel. In some cases statistician and data processor are one and the same person. However, if a professionally trained statistician's major attraction to medical colleagues is his ability to feed data into the computer, ultimately both parties are liable to be disappointed in their relations.

This is not to decry the very real and important advances in computer technology, which have resulted in substantial savings in investigator time, in improved data accuracy and quality, and in a greater variety in the statistical procedures used to analyze clinical trials data. A modern software system[28] specifically designed for medical records has been used successfully to process information for several complex clinical studies. In this system data are automatically edited to ensure that values of all items are within specified ranges. Registration cards, surgery and pathology checklists, multiple follow-up reports and other records are easily incorporated in patient's computer file. "Code sheets" for each record type are maintained in a separate computer file and used to interpret the numerical codes in lists and tables. The system provides for summarization of the data on each individual into a single record suitable for statistical analysis, thus relieving the investigator of the time-consuming task of compiling and recording summary information himself.

ANALYZING THE DATA

Undoubtedly the most visible contribution the statistician makes to cooperative clinical studies comes when it is time to analyze the data. While his first job is to present answers to the main

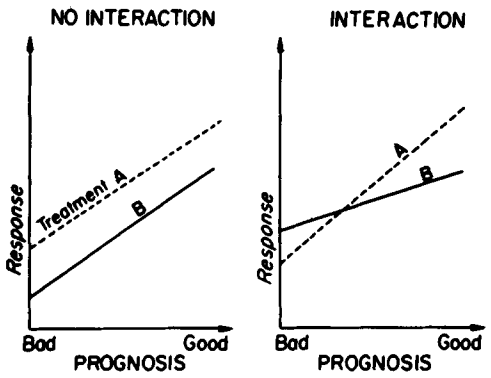

FIG. 2. Schematic drawing of response as related to treatment and baseline variables, contrasting interaction and no-interaction situation.

questions posed by the study in as simple and straightforward a manner as possible, these results may then be checked, refined, and extended using more complicated procedures. Data from several studies or sources may be pooled for comprehensive retrospective studies of epidemiological and prognostic factors. Careful stratification on such factors during analysis makes for more precise treatment comparisons by decreasing the amount of "unexplained" variation in outcome.

Multivariate procedures, which consider the joint effects of several "independent" variables on the "dependent" variable of interest, constitute a large part of modern statistical data analysis. Thus a retrospective study[7] of children with neuroblastoma showed that age and stage of disease at diagnosis each had independent effects on survival and led to the recommendation that account be taken of both these factors simultaneously in analysis of future studies. Efforts are now being made to try to confirm the effect of circulating lymphocytes[4] as a third prognostic factor.

Many multivariate studies have been made of factors influencing the duration of survival in acute leukemia.[19,23,36,43] Without exception these identify age at diagnosis, cell type, and the initial leukocyte count as important prognostic variables. The results for sex, race, hemorrhagic status, platelet count and organ enlargement are less consistent, but nevertheless significant in some studies.

Classical multivariate analyses are performed in the context of a mathematical model which specifies a linear relationship between the dependent variable and the several independent variables. For many studies of childhood cancer the dependent variable is the elapsed time from diagnosis or treatment until death, or else some transform such as the log or square root of survival. This implies that virtually all children have been followed until death before the study can be analyzed. The desire to extend such models for use with incomplete follow-up data, in which a substantial number of patients are still alive and well, has stimulated a good deal of recent work in statistical theory and methodology.[6,10,14] Some of these methods[30] also offer hope of being able to incorporate into the analysis changes in a patient's status on one or more prognostic variables which occur during the course of follow-up. This could be important, for example, in determining the relationship between episodes of CNS disease and bone marrow relapse or ultimate survival in acute leukemia.[17] Likewise the true significance of bone marrow lymphcytosis occurring during remission might be explored, as previous attempts to answer this question[38] have suffered from deficient methodology.[8]

Another important role for multivariate analysis is in the detection of interaction effects between treatment and baseline variables. The general concept of an interaction is illustrated schematically in Figure 2. If there are no interactions the effect of treatment A versus B is the same regardless of prognosis. Presence of interactions means that treatment effectiveness varies according to prognostic category. There has been a growing realization with acute leukemia, for example, that while patients having a low initial leukocyte count benefit from intensive chemotherapy, patients having a high count do not respond as well.[31,32] Careful study of the relationships between treatment and baseline variables in such circumstances could ultimately lead to guidelines for selecting the "optimum" treatment for each patient, based on his own individual characteristics.[9] This would be an invaluable aid to the intuition of clinical judgment.

CONCLUSIONS

In summary, the role of the statistician in cooperative clinical research is a multifaceted one, dealing with questions of sample size, experimental design, stopping rules, data processing and analysis. Statistical input into programs of cooperative research has resulted in more rigorous protocols, improvements in data quality, and more comprehensive analyses of results.

Protocol studies of particular treatments, as well as large scale retrospective summaries of "natural history" and prognostic factors, have been strengthened in the process.

While the demand for professionally trained statisticians will undoubtedly continue for some time, many medical research workers have learned to carry out a number of these activities on their own. The improvements in computer languages are giving investigators the opportunity to manipulate their own data files without detailed technical training. As more physicians become acquainted with cooperative research programs, the basic statistical concepts inherent in clinical trials methodology gain greater understanding and acceptance. The distinction between statistician and scientist is blurred further as statisticians themselves gravitate toward more substantive and less methodological pursuits, as in genetics and epidemiology. For in the final analysis, as Cornfield[15] has suggested, good statistics are simply good science.

REFERENCES

1. Anscombe, F.: Sequential medical trials, a book review. *J. Am. Stat. Assoc.* 58:365-383, 1963.
2. Armitage, P.: Sequential Medical Trials. Oxford, Blackwell Scientific Publications, 1960.
3. Armitage, P., Block, J. B., and Saracci, R.: Report of the scientific meeting on some developments in the methodology of controlled clinical trials in cancer. *Int. J. Cancer* 14:576-579, 1974.
4. Bill, A., and Morgan, A.: Evidence for immune reactions to neuroblastoma and future possibilities for investigation. *J. Pediatr. Surg.* 5:111-116, 1970.
5. Breslow, N.: Sequential modification of the UMP test for binomial probabilities. *J. Am. Stat. Assoc.* 65:639-648, 1970.
6. Breslow, N.: Statistical analysis of survival data under the proportional hazards model. *Rev. Int. Stat. Inst.* 43:45-58, 1975.
7. Breslow, N. and McCann, B.: Statistical estimation of prognosis for children with neuroblastoma. *Cancer Res.* 31:2098-2103, 1971.
8. Breslow, N., and Zandstra, R.: A note on the relationship between bone marrow lymphocytosis and remission duration in acute leukemia. *Blood* 36:246-249, 1970.
9. Byar, D. P., Simon, R. M., Friedewald, W. T., Schlesselman, J. J., DeMets, D. L., Ellenberg, J. H., Gail, M. H., and Ware, J. H.: Randomized Clinical Trials: prespectives on some recent ideas. *N. Engl. J. Med.* 295:74-80, 1976.
10. Byar, D. P., Huse, R., Bailar, J. C., and the Veterans Administration Cooperative Urological Research Group: An exponential model relating censored survival data and concomitant information for prostatic cancer patients. *Natl. Cancer Inst. Monogr.* 52:321-326, 1974.
11. Chalmers, T. C., Block, J. B., and Lee, S.: Controlled studies in clinical cancer research. *N. Engl. J. Med.* 287:75-58, 1972.
12. Cornfield, J.: Sequential trials, sequential analysis, and the likelihood principle. *Am. Statistn.* 20:18-23, 1966.
13. Coronary Drug Project Research Group: The Coronary Drug Project: Initial findings leading to modifications of its research protocol. *JAMA* 214:1303-1313, 1970.
14. Cox, D. R.: Regression models and life tables. *J. R. Stat. Soc.* 34:187-202, 1972.
15. Cutler, S. J., Greenhouse, S. W., Cornfield, J., and Schneiderman, M. A.: The role of hypothesis testing in clinical trials. *J. Chron. Dis.* 19:857-882, 1966.
16. D'Angio, G. J., Evans, A. E., Breslow, N., Beckwith, B., Bishop, H., Feigl, P., Goodwin, W., Leape, L. L., Sinks, L. F., Sutow, W., Tefft, M., and Wolff, J.: The Treatment of Wilms' Tumor: Results of the National Wilms' Tumor Study. *Cancer* 38:633-646, 1976.
17. Evans, A. E., Gilbert, E. S., and Zandstra, R.: The increasing incidence of central nervous system leukemia in children. *Cancer* 26:404-409, 1970.
18. Freireich, E. J. and Gehan, E. A.: Historical controls re-incarnated: an examination of techniques for clinical therapeutic research. *Cancer Chem. Rep.* 58:623-626, 1974.
19. Gehan, E. A.: An application of multivariate regression analysis to the problem of predicting survival in patients with acute leukemia. *Bull. Int. Statist. Inst.* 39:173, 1961.
20. Gehan, E. A.: Early human studies with anti-cancer agents: the question of sample size. *J. Chron. Dis.* 15:265-268, 1962.
21. Gehan, E. A., and Freireich, E. J.: Non-randomized controls in cancer clinical trials. *N. Engl. J. Med.* 290:198-203, 1974.
22. George, S. L., and Desu, M. M.: Planning the size and duration of a clinical trial studying the time to some clinical event. *J. Chron. Dis.* 27:15-24, 1974.
23. George, S. L., Fernback, D. J., Vietti, T. J., Sullivan, M. P., Lane, D. M., Haggard, M. E., Berry, D. H., Consdale, D. and Komp, D.: Factors invluencing survival in pediatric acute leukemia. The SWCCSG experience, 1958-1970. *Cancer* 32:1542-1553, 1973.
24. Heady, J. A.: A cooperative trial on the primary prevention of ischaemic heart disease using clofibrate: design, methods and progress. *Bull. WHO* 48:243-256, 1973.
25. Hill, A. B.: The clinical trial. *N. Engl. J. Med.* 247:113-119, 1952.
26. Hill, A. B.: Measurement in medicine; clinical trial. *Br. Med. Bull.* 7:278-282, 1951.
27. Hill, A. B.: Principles of Medical Statistics. New York, Oxford University Press, 1971.
28. Kronmal, R. A., Bender, L., and Mortenson, J.: A conversational statistical system for medical records. *Appl. Stat.* 19:82-92, 1970.
29. Leiken, S., Evans, A., Heyn, R., and Newton, W.: The impact of chemotherapy on advanced neuroblastoma. Survival of patients diagnosed in 1956, 1962 and 1966-68 in Children's Cancer Study Group A. *J. Pediatr.* 84:131-134, 1974.
30. Mantel, N., and Byar, D. P.: Evaluation of response time data involving transient states: An illustration using heart transplant data. *J. Am. Stat. Assoc.* 69:81-86, 1974.
31. Medical Research Council Working Party on Leukemia in Childhood: Treatment of acute lymphoblastic leukemia. *Br. Med. J.* 4:189-194, 1971.
32. Miller, D. R., Sonley, M., Karon, M., Breslow, N., and Hammond, D.: Additive therapy in the maintenance of remission in acute lymphoblastic leukemia of childhood: the effect of the initial leukocyte count. *Cancer* 34:508-517, 1974.

33. Natrella, M. G.: Experimental statistics. National Bureau of Standards Handbook 91, U.S. Government Printing Office, 1963.

34. Peto, R., Pike, M. C., Armitage, P., Breslow, N. E., Cox, D. R., Howard, S. V., Mantel, N., McPherson, K., Peto, J., and Smith, P. G.: Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Br. J. Cancer* 34:585–612, 1976.

35. Peto, R., Pike, M. C., Armitage, P., Breslow, N. E., Cox, D. R., Howard, S. V., Mantel, N., McPherson, K., Peto, J., and Smith, P. G.: Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. *Br. J. Cancer* 35:1–39, 1977.

36. Pierce, M. I., Borges, W. H., Heyn, R., Wolff, J., and Gilbert, E. S.: Epidemiological factors and survival experience in 1770 children with acute leukemia. *Cancer* 23:1296–1304, 1969.

37. Pocock, S. J., and Simon, R.: Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics* 31:103–115, 1975.

38. Skeel, R. T., Henderson, E. S., and Bennett, J. M.: The significance of bone marrow lymphocytosis of acute leukemia patients in remission. *Blood* 32:767–773, 1968.

39. Sutow, W., Gehan, E., Heyn, R., Kung, F., Miller, R., Murphy, M., and Traggis, D.: Comparison of survival curves, 1956 versus 1962, in children with Wilms' tumor and neuroblastoma. *Pediatrics* 45:800–811, 1970.

40. Taves, D. R.: Minimization: A new method of assigning patients to treatment and control groups. *Clin. Pharmacol. Ther.* 15:443–453, 1974.

41. Wolff, J. A., Brubaker, C. A., Murphy, M. L., Pierce, M. I., and Severo, N.: Prednisone therapy of acute childhood leukemia: Prognosis and duration of response in 330 treated patients. *J. Pediatr.* 70:626–631, 1967.

42. Zelen, M.: The randomization and stratification of patients to clinical trials. *J. Chron. Dis.* 27:365-375, 1974.

43. Zippin, C., Cutler, S. J., Reeves, W., Jr., and Lum, D.: Variation in survival among patients with acute lymphocytic leukemia. *Blood* 37:59–72, 1971.