THE JOURNAL OF MEDICINE AND PHILOSOPHY

Sponsored by the Center for Ethics, Medicine, and Public Issues, a joint project of Baylor College of Medicine, the Institute of Religion, and Rice University, in collaboration with the School of Medicine, University of Connecticut [illegible] Center.

This quarterly publication has been established under the auspices of the Society for Health and Human Values to explore the themes and concerns of philosophy and the medical sciences. Central issues in medical research and practice have important philosophical dimensions, for in treating disease and promoting health, medicine involves presuppositions about human goals and values. Conversely, the concerns of philosophy often significantly relate to those of medicine, as philosophers seek to apprehend the nature of knowledge and the human condition in the modern world. In addition, recent developments in medical technology and treatment raise ethical problems that overlap with philosophical interests. *The Journal of Medicine and Philosophy* aims to provide an ongoing forum for the discussion of these themes and issues.

Unsolicited manuscripts are considered for publication for future thematic issues. Prospective authors should contact issue editors for further information. Closing dates for receipt of manuscripts are indicated in parentheses.

'Justice Between Generations and Health Care for the Elderly'
Prof. Norman Daniels
Department of Philosophy, Tufts University, Medford, MA 02155, U.S.A.
(January 1, 1987)

'Ethical Issues in Research with Animals'
Prof. Tom L. Beauchamp
Kennedy Institute of Ethics, Georgetown University, Washington, D.C. 20057, U.S.A.
(April 1, 1987)

'Comparative Medical Ethics'
Prof. Robert M. Veatch
Kennedy Institute of Ethics, Georgetown University, Washington, D.C. 20057, U.S.A.
(July 1, 1987)

'The Structure of Clinical Knowledge'
Prof. Kenneth F. Schaffner
Department of History and Philosophy of Science, University of Pittsburgh, 312 Loeffler Building, Pittsburgh, PA 15260, U.S.A.
(January 1, 1988)

'Commissions and Biomedical Ethics: International Perspectives'
Prof. LeRoy Walters
Kennedy Institute of Ethics, Georgetown University, Washington, D.C. 20057, U.S.A.
(April 1, 1988)

'Neurobiology and the Future of Psychiatry'
Prof. Stuart F. Spicker
Department of Community Medicine and Health Care, University of Connecticut Health Center, Farmington, CT 06032, U.S.A.
(July 1, 1988)

'Philosophical Critique of Bioethics'
Prof. K. Danner Clouser
Department of Humanities, Milton S. Hershey Medical Center College of Medicine, Pennsylvania State University Hershey, PA 17033, U.S.A.
(January 1, 1989)

'Toward a Medical Axiology'
Dr. Edmund D. Pellegrino
Kennedy Institute of Ethics, Georgetown University, Washington, D.C. 20057, U.S.A.
(July 1, 1989)

In addition to issues of the *Journal* centered on selected themes, the editorial policy also provides for the publication of nonthematic issues. While a certain number of papers are invited for thematic issues, unsolicited papers may be addressed to the preannounced topic of a specific issue or to any significant topic in the relationships of medicine and philosophy. Especially suitable for nonthematic issues would be commentaries, critiques, and responses to previously published papers. In this way the Editorial Board hopes to establish an active forum for a continuing dialogue on fundamental philosophico-medical questions to which any scholar may contribute. Unsolicited papers should be sent to:

Managing Editor, JMP,
Center for Ethics, Medicine, and Public Issues,
Baylor College of Medicine,
One Baylor Plaza,
Houston, TX 77030, U.S.A.

The Editorial Board acknowledges with gratitude a grant from the Upjohn Company that makes possible the publication of the Journal

The Journal of Medicine and Philosophy is published 4 times per annum: February, May, August, and November.
Subscription prices, per volume: Institutions $ 47.00. Individuals $ 22.50.
Second-class postage paid at New York, N.Y. USPS No. 589-490.
U.S. Mailing Agent: Expediters of the Printed Word Ltd., 515 Madison Avenue (Suite 917), New York, NY 10022.
Published by D. Reidel Publishing Company, Spuiboulevard 50, P.O. Box 17, 3300 AA Dordrecht, Holland, and 190 Philip Drive, Norwell, MA 02061, U.S.A.
Postmaster: please send all address corrections to: Kluwer Academic Publishers Group, c/o Expediters of the Printed Word Ltd., 515 Madison Avenue (Suite 917), New York, NY 10022, U.S.A.

III-15

ETHICAL PROBLEMS IN CLINICAL TRIALS

I. INTRODUCTION

A. *Historical Background and Definitions*

Clinical trials, for all their widespread use in contemporary medicine, are surprisingly a mainly postwar phenomenon. Though there are earlier antecedents to the randomized, double-blind, controlled clinical trial (see Armitage, 1984, p. 68; Feinstein, 1986, pp. 684–688; Miké, 1982, pp. 113–118; and Schaffner, 1983, pp. 197–198), the first large scale randomized controlled trial involving the antibiotic streptomycin began only in 1946 under the statistical guidance of A. B. Hill (Medical Research Council, 1948).

The reason for the widespread acceptance of clinical trials is largely based on the belief that such trials are the best scientific means of assessing the efficacy of putative therapies. The history of medicine is replete with clinical therapies that have won the uncritical acceptance of physicians.[1] The need for clinical trials is also indirectly supported by several professional ethical codes. The Declaration of Helsinki, for example, states:

> Biomedical research involving human subjects must conform to generally accepted scientific principles and should be based on adequately performed laboratory and animal experimentation and on a thorough knowledge of the scientific literature. (1964, I.1)

Though these general injunctions are unquestionable, whether "accepted scientific principles" require a "clinical trial", exactly

Kenneth F. Schaffner, M. D., Ph.D., Professor of History and Philosophy of Science. Co-Director, Center for Medical Ethics, 1017 CL, University of Pittsburgh, Pittsburgh, PA 15260, U.S.A.

The Journal of Medicine and Philosophy 11 (1986), 297–315.

when and under what specific circumstances a clinical trial should begin, and what type of clinical trial it should be (e.g., should it be a fully randomized trial?) are intensely debated topics.

The randomized, controlled clinical trial (hereafter RCT) refers to that form of investigation that involves (1) one or more treatment groups and a (placebo or standard therapy) control group, (2) randomized (possibly after stratification) assignment of patients to the two or more groups, sometimes referred to as "arms" of the trial, and (3) often a single- or double-blind design in which the assignments of the agents or procedures being tested are not known to the patients nor possibly to the treating health professionals.

For the purposes of perspective, as well as to be better able to appreciate scientifically respectable alternatives to the RCT, it will be useful to provide a schematic summary of various research designs in addition to the RCT. The following Figure representing various forms of the epidemiological study is based in part on Lilienfeld and Lilienfeld (1980, p. 192) and in part on Fletcher *et al.* (1982, p. 193):

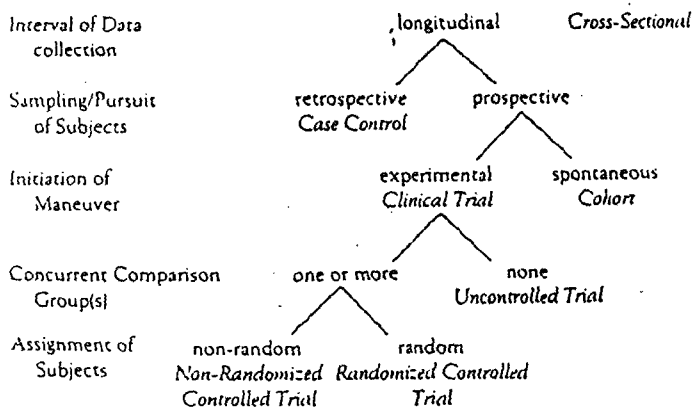

Fig. 1. The epidemiological study: Various research designs (in italics) used to establish causation in medicine.

B. *Ethical Problems*

Almost from their inception, medical research in general and clinical trials in particular have been associated with a number of ethical problems. These have ranged from abuses, such as the Tuskeegee Syphilis Experiment and other unethical research practices (see Jones, 1981 and Beecher, 1966) to more current debates discussed at length in the articles in this issue of the *Journal*. Such abuses have resulted in more stringent federal regulations that require the implementation of specific safeguards for human subjects involved in clinical trials (see DHHS, 1981; and FDA, 1981). One of the recurring ethical themes associated with medical research in general but that arises with special force in clinical trial methodology can be termed "the dilemma of the healer and the scientist", a problem that will be pursued in the following section. I will also consider ethical problems associated with the clinical trial in its traditional form, as well as review several alternatives to the standard RCT which various authors have felt are "more ethical" clinical forms of investigation. These issues will occupy me in later sections of this essay.

II. THE DILEMMA OF HEALER AND SCIENTIST

As healer, the physician is bound by various professional codes to act in the patient's best interest.[2] As scientist, the physician's main concern is to respect the scientific canons of valid experimental design, which may require that the physician sacrifice the goal of individualized best treatment for statistical efficiency. Lellouch and Schwartz (1971) develop a similar distinction between an "ethique collective" and an "ethique individuelle". Clayton (1982) characterizes these two different ethical responsibilities of the physician as follows:

(1) *Individual ethic*

> It is the duty of the doctor to apply existing knowledge for the best possible treatment of each individual patient.

(2) *Collective ethic*

> (a) It is the duty of the doctor to acquire new knowledge so that, by such advance, future patients might benefit, and,
>
> (b) having acquired new knowledge, to accurately communicate it to other doctors (1982, pp. 470–471).

The two roles or "ethics" of the physician are not necessarily in harmony. All of the authors represented in this issue of the *Journal* comment on this dilemma and on various ways in which the tension may or may not be resolved ethically.

Kopelman, in her article, argues that though there are forms of clinical trials that are ethically incorrect, there are also other forms in which RCTs can be an "honorable and cooperative venture" between investigator and subjects. Some of these forms are traditional RCT designs, and others are recent modifications. Each will be commented on in relevant subsections below. Suffice it to note that Kopelman is doubtful whether any *general* approval can be given for a specific research design schema, since "their suitability cannot be determined abstractly as they have different strengths and weaknesses in relation both to the consent requirements and the structural integrity of the RCT" (1986, p. 342).

Gifford (1986) notes the existence of the physician's "therapeutic obligation" (TO), a term he borrows from one of Marquis's earlier (1983) articles, and discusses various defenses of RCTs, which I shall return to further below. Ultimately for Gifford, however, there is an uneliminable conflict between RCTs and the TO. However, the conflict can be mediated, and to a degree ethically resolved, by an appeal to some form of contract theory. Gifford does not as yet propose the details of such a resolution, though he does cite Rawlsian and Nozickian possibilities.

Marquis and Kadane in their essays are each sensitive to the healer-scientist dilemma as it develops in the context of the RCT. Marquis is primarily concerned with a special form of the RCT known as the pre-randomized design, having previously (Marquis, 1983) criticized standard RCTs as "incompatible with the ethics of the patient-physician relationship". Kadane argues that a Bayesian departure from the classical RCT can permit a "more ethical" methodology for clinical trials. I shall discuss each of their contributions in more detail later.

III. MAKING CLINICAL TRIALS ETHICALLY ACCEPTABLE

Several authors have examined the ethical foundations of clinical trials including Fried (1984), Levine and Lebacqz (1979), and Beauchamp and Childress (1983). As I have written elsewhere

(Coulehan *et al.* 1983), general ethical principles such as autonomy and respect for persons, nonmaleficence, beneficence, and justice, can be utilized in evaluating the design of various clinical trials. Clinical trial designs can accordingly be judged as more or less ethical in terms of the net benefits they provide for patients and society, and the extent to which they respect persons by providing a full opportunity for informed consent.

The ethical acceptability of various forms of clinical trials can be examined under a number of different rubrics as a review of the articles in this issue of the *Journal* will indicate. One way for us to proceed in this Introduction is to consider the topic under several general questions: (1) when is it ethical to start a trial?, (2) can there really be informed consent in a clinical trial?, (3) when should a clinical trial be stopped for ethical reasons?, and (4) what modifications might be made in the traditional clinical trial to make it "more ethical"? Under these general queries, I shall also consider several more specific questions.

A. *When is it Ethical to Start a Trial?*

(1) *Can Historical Controls be used in Place of RCTs?* I will say only a little about this difficult issue since our authors do not extensively examine it. There are two generally opposing views. Chalmers (1972), on the one hand, believes that there are good arguments for always beginning the use of a new therapy with an RCT. On the other hand, there are theorists such as Freireich and Gehan (1979) who contend that RCTs are highly overrated and that more extensive use of historical controls is desirable. Most commentators' sentiments appear to take a middle position, and it should be recalled that for *drug* trials, a three phase investigation is required by the FDA, beginning with tests for toxicity and proceeding to RCTs only in phase III. Gifford comments on trial initiation in the context of a discussion of the need for clinical trials, and also in connection with the issue of when randomization should begin. A more specific question is associated with trial initiation involves the issue of known or suspected benefits and harms of the arms of the trial. I turn next to this problem.

(2) *Is the Initial "Null Hypothesis" Ethically Null?* There is a fairly standard type of reply that is given to the above discussed "Dilemma of the Healer and Scientist." This is that, certainly when

a trial begins, it is not know that there is any *real difference* in efficacy (and safety) between the standard therapy (or if none exists, a placebo and the experimental therapy. This defense can be found in Chalmers (1972). Gifford (1986) refers to this as "the knowledge argument". Others including Levine (1981) and Kopelman (1986) characterize this as the problem regarding whether a "null hypothesis" can honestly and reasonably be stated.

Even if such an initial null hypothesis can be defended on the basis of objective grounds, Fried (1974) has argued powerfully that important *subjective* considerations may come into play. (See Fried, 1974, pp. 153–154.) He writes:

> Even in medically equivalent cases, patients may have quite different value systems: their life plans may have quite different structures. And though the overall prognosis, the overall expected value of the therapies may be practically the same, the composition of the risks and benefits of each therapy might be different. Thus for instance surgery for heart disease in some cases might involve a very high initial risk of surgical mortality followed by a very good risk for five years of survival after surgery, while the standard medical treatment for the same condition may have the same overall mortality expectation, but a risk of death distributed more evenly over a period of years. Different people might quite rationally have different preferences about this. (1974, p. 153)

This is a problem that is of special interest to Marquis and to Kadane, though whether any statistically efficient general form of clinical trial can adequately respond to Fried's point is debatable, and may, as Kopelman seems to suggest, ultimately depend on the particulars of a trial.

B. *Can Consent to Participate in a Trial be Truly Informed?*

The question of informed consent is one that pervades a number of the ethical problems with clinical trials. I will comment on three.

(1) *Must Randomization be Disclosed?* One informational issue that was a subject of fairly vigorous discussion in the mid- and late 1970s is now, I believe, a matter of settled consensus. This is the question of whether the subject should be explicitly informed if an assignment was to be done *randomly*. Arguments such as found in Fried (1974) and Levine and Lebacqz (1979) appear to have convinced most writers (and IRBs) that the method of assignment is material to subjects' decisions to participate in a trial (see Levine, 1981, p. 131). A personal review of a number of consent forms indicates that language such as "the treatment I will receive has been made by a process called randomization", and explanatory expressions such as "this means ... by chance", now typically appear in such consent forms. The interesting issue concerning randomization now is no longer whether the procedure should be disclosed, but rather whether randomization is really necessary to assure sound scientific inference. This is a topic currently discussed in the medical journals (see Chalmers *et al.*, 1983 and Wahrendorf *et al.*, 1985), in the philosophy of science literature (Urbach, 1985), and in the papers by Kopelman, Gifford and Kadane in this volume.

(2) *Should Interim Data be Disclosed?* There is another related difficulty with RCTs that is touched on only briefly in the papers in this issue of the *Journal* and deserves a short comment. This is the problem of what to do about interim data. Kopelman addresses this issue citing federal guidelines that suggest "it may be appropriate to inform subjects of 'significant new findings developed during the course of the research which may relate to the subject's willingness to continue participation ...'". Many investigators worry that disclosure of early trends to subjects would both be unreliable and misunderstood. Following classical RCT methodology, attainment of a significance level of ≤ 0.05 is often a requirement for publication of the trial's results.

Kopelman's response is to cite instances in which subjects are willing to waive the right to such interim data. In this position, she seems to agree with Veatch (1979) and Levine (1981, pp. 134–135) who opt for an informed "contractual" agreement to "incomplete disclosure" between investigator and subject. Interestingly, Kadane's proposal for his form of trial automatically updates patient assignment to reflect such interim data, though he also would impose restrictions on the full availability of information generated in the course of a trial.

(3) *Can a Contract Model Really Work in a Clinical Trial Context?* At several points in the above discussion we have encountered the suggestion that a contract model might be invoked, for example, for resolving problems such as a patient not necessarily receiving the best therapy or not being informed about interim data generated in a trial. Such a contract model typically assumes that the similarly situated contractees have each given their informed consent to such an agreement. Given the signifi-

cant differences between the frequently undereducated and often quite ill patient on the one hand and the trained staff on the other, it is questionable whether potential contractees bargain from similarly situated positions. In addition, it is still a major *empirical* question about whether subjects agreeing to participate in a clinical trial have given a true informed consent. In point of fact, the empirical data we possess at this time seems to argue the contrary except for relatively few individuals (Lidz *et al.*, 1983). This issue of informed consent is one that concerns all of the authors in this volume, and the reader should note the broad variety of contexts in which the question arises.

C. *When Should a Clinical Trial be Stopped?*

Most commentators feel that though disclosure may be restricted as regards preliminary trends favoring one mode of treatment over the other, they are much more uncomfortable with non-disclosure of significant unanticipated harmful side effects. There is in addition, the desire not to prolong a trial unnecessarily long — thus depriving others of the better therapy or subjecting patients in the trial to a less advantageous treatment. There is also, however, the countervailing consideration not to end a trial too soon, before the new therapy is given an appropriate chance to prove itself. These problems have been addressed almost from the inception of clinical trials under the rubrics of "sequential trials", trial "monitoring", and "stopping rules". Armitage, an important contributor to the sequential approach wrote in 1960 that " ... the organizer of a medical trial, anxious to avoid any use of poor treatments will often find it useful to examine the results, serially, as they become available. The ethical need to avoid extensive experimentation when treatment differences are large is characteristic of medical trials" (1960, pp. 16—17). The discussion of these issues is often quite technical, and a good overview of some of the alternative statistical proposals in this area can be found in Gail's (1982) essay. Writing on the subject of stopping rules, Armitage cautions that *formal* statistical rules for trial termination should be only one component of a judgment to stop a trial (1984, p. 69).

In his essay in this *Journal*, Gifford comments on one such recent stopping rule proposed by the statistician, Paul Meier. Meier's rule is novel and suggestive, and Gifford finds it "promising", though he does discuss one difficulty with it. Meier himself notes in his (1979) that there are at least three problems with implementing the rule. The consensus here appears to be that stopping rules are important but not yet sufficiently well understood to be employed in an algorithmic fashion and, accordingly, that they need more research.

D. *What Modifications Might be Made in the Traditional Clinical Trial to Make it 'More Ethical'?*

(1) *Adaptive Designs.* There have over the past two decades been a number of suggestions made to modify the traditional design of RCTs so as to ameliorate some of the ethical problems with such trials. One class of novel designs addressed the concern that as a trial generated data, one of the therapies would gradually be seen as the better one, with its competitor(s) thus constituting *inferior* treatment(s) for the patients in that arm(s). The proponents of "adaptive designs" sought to develop modified forms of RCTs that would minimize the numbers of patients receiving the inferior treatment and at the same time satisfy sound principles of statistical inference. Gifford refers briefly in his essay to the "two-armed bandit" design, which was but one of a number of adaptive designs which included Zelen's (1969) "play the winner" strategy. Kadane in his essay also comments on some of the difficulties with adaptive designs. A thorough review of these designs can be found in the article by Hoel *et al.* (1975). Weinstein (1974) and Byar *et al.* (1976) are also worth consulting. Suffice it to say that the designs did not live up to their promise and to the best of my knowledge have never been used in actual clinical trials. Significantly, Kadane's design discussed in his article in this *Journal* contains a variant of such "adaptation", as we shall see below.

(2) *Prerandomization Designs.* A variant of the traditional design that has been employed in a number of important RCTs involves an assignment procedure termed "prerandomization". The general idea appears to be due to Zelen (1977, 1979) though there are importantly different forms of prerandomization as clearly depicted in Kopelman's Figures 3, 4, and 5 in this *Journal* and which are sometimes conflated under the general term of "prerandomization".

Prerandomization has been implemented both because of the reluctance of physicians to enroll their patients in certain types of

clinical trials (such as a test of different surgical therapies for breast cancer) and because of patient resistance to participation in standard RCTs. Two statisticians working with the National Surgical Adjuvant Project for Breast and Bowel Cancers (NSABP) and involving a breast cancer surgery trial NSABP B-06 of total versus segmental mastectomy (see Fisher *et al.*, 1985) wrote in 1979 that:

> A major problem with the (initial standard RCT) protocol appeared to be the lack of acceptability of the randomization. Physicians were reluctant to approach patients at the time of operation about chance assignment to surgical therapies that involved either removal or cosmetic preservation of the breast. Patients also had difficulty dealing with randomization. In many cases, patients were not even certain whether or not they had a breast cancer and yet they were being asked to consider quite dissimilar surgical procedures if cancer was found at the time of surgery. Further, even when the patient knew the diagnosis, it was disquieting not to know which surgery would be performed, i.e., whether she would wake up with or without a breast (Redmond and Bauer, 1979, pp. 1–2).

More recently, several articles in the *New England Journal of Medicine* provided additional information as well as several different perspectives on these problems. Taylor *et al.* (1984) report that they received questionnaire responses from 94 of the 97 principal investigators of the total versus segmental trial. Physicians who did not enter all of their patients into the trial gave as their reasons: concern that the doctor-patient relationship would be affected by a randomized trial (73%), difficulty with informed consent (38%), dislike of open discussions involving uncertainty (22%), and perceived conflict between the roles of scientist and clinician (18%). Angell states that after the prerandomization procedure was adopted the "accrual rate increased sixfold" (1984, p. 1386).

The prerandomization design used in the NSABP B-06 trial involves randomization of the patient to an arm of the trial, with consent *then* sought for that randomly assigned procedure. The entire protocol is explained to the patient along with all treatment options. "The patient is informed, however, that if she agrees to participate in the trial, she will receive the treatment that has been randomly selected for her. It is only at this point that an informed consent is obtained" (Redmond and Bauer, 1979, p. 2). If the patient elects not to receive this randomly assigned treatment, she may refuse it and be given the treatment of her choice. She will be asked if the projects' investigators may follow the results of her chosen therapy. Obtaining this follow-up information is crucial in order to avoid bias.

In this issue of the *Journal*, Kopelman generally approves of this design whereas Marquis argues at length that prerandomization is ethically flawed. Marquis argument is too long to do adequate justice to it in this introduction but several comments are worth making. First, it is important to distinguish carefully the various forms of prerandomization designs so that the flaws in one form (for example Zelen's "single arm consent" design) are not unfairly applied to the NSABP B-06 form. Marquis's argument does not depend on this conflation, since though he discusses the single arm form, he also directs his criticism against Zelen's multiple arm design. However, comments in the literature about prerandomization can lead to this conflation. A close inspection of Kopelman's Figures 3, 4, and 5 will aid the reader in this discrimination.

Second, the question of *why* prerandomization is so successful in increasing accrual rates is one for which we do not have a clear answer. This is a most important answer to have in formulating a response to Marquis's argument. The worry is that there *may* be some "informing bias", to use Kopelman's term, either consciously or unconsciously at work in the consent process. Ellenberg in her (1984) article expressed this concern clearly when she wrote of the ways that different strengths and weaknesses of the standard and experimental therapies *could* be presented to prerandomized patients (1984, pp. 1406–1407, quoted by Marquis, 1986, on p. 377 below). Kopelman feels that the "informing bias" may be offset by the fact that the NSABP trial provided "the same excellent three-page consent form for all persons in all groups", though she also appears to agree with Ellenberg (1984) that prerandomization is of rather limited usefulness. Marquis, on the other hand, argues that prerandomization can only succeed either if it is done unethically or that it was unnecessary and thus could have been replaced by a standard RCT. (Marquis does not, I should add, consider standard RCTs as ethical — see his (1983).) Thus he concludes in his essay that "prerandomization is always wrong".

The importance of securing agreement from the prerandomized subjects to accept their prerandomized therapy is forcefully underscored by an examination of the extremely rapid rate of additional required accrual if the refusal rate exceeds approximately 30%. A table from Ellenberg (1984) is given below:

TABLE I
Sample-size 'inflation factor' according to overall refusal rate in prerandomized studies (after Ellenberg, 1984)

Refusal rate	Inflation factor
0.02	1.09
0.05	1.23
0.10	1.56
0.15	2.04
0.20	2.78
0.25	4.00
0.30	6.25
0.35	11.11
0.40	25
0.45	100
0.50	—[a]

[a] If half the patients on each arm refuse the randomly assigned treatment and receive the other one, the determination of differences in treatment effect is impossible regardless of the sample size.

The NSABP investigators have considered attempting to minimize any informing bias by providing a common videotape to all patients, but this has not as yet been implemented (Redmond, personal communication, July 1986). There also has as yet not been any attempt to gather empirical material from either patients or from patient consent sessions to determine the reason(s) for increased accrual or to test for any informing bias. Marquis does not believe that prerandomization's simple elimination of uncertainty is an adequate explanation for accrual increase, though Redmond and Bauer as noted above suggest this may be at least partially the case. In the light of the dearth of current additional information, these are areas of significant future investigation for individuals interested in assuring the ethical validity of prerandomization.

(3) *Bayesian Alternatives.* The term "Bayesian" refers to a general approach to the foundations of statistics which is quite different from the more "classical" approach typically assumed by biomedical statisticians. Because the Bayesian perspective is so different, often in rather subtle but important ways, a few general comments are in order. Additional details concerning one Bayesian approach can be found in the paper by Kadane in this issue.

In spite of some diversity, Bayesians generally subscribe to several key concepts which are disavowed by classical statisticians. First, Bayesians accept "Bayes' Theorem", whence their name. This theorem, uncontroversial as an uninterpreted theorem provable from the probability calculus, generates significant disagreement between Bayesians and non-Bayesians when it is taken as guide for revising belief in the light of new experimental evidence. The theorem is discussed in an earlier issue of this *Journal* on clinical decision-making. A Bayesian perspective on clinical hypothesis testing has been developed by a number of authors, but Cornfield's (1966, 1969) is perhaps the most cited of earlier approaches. Notions that are standard from the classical statistical perspective become reinterpreted in a Bayesian world, and some writers such as Clayton view this approach as perhaps suggesting a formal model in which we would in effect abandon controlled trials and return to a reliance on clinical judgment (1982, p. 474), a view which partially parallels Gifford's suggestions made at the conclusion of his paper in this volume. Clayton's view is rather extreme, however, and is not subscribed to by most Bayesians.

Kadane and Sedransk (1980) have outlined a Bayesian methodology for what they characterized as "a more ethical clinical trial". In this volume, Kadane elaborates on that design, discusses an actual trial now underway at Johns Hopkins using his new design, and provides a defense of this type of trial against potential objections. In this essay, I will only be able to comment briefly on the design of Kadane's trial to highlight some of the differences with the classical perspective; I will also briefly comment on what I see as some of its ethical strengths and weaknesses.

The Kadane type of trial has several objectives that make it a candidate for a more ethical type of investigation. First, it defines a set of admissible therapies for any given patient which insures that such a patient will receive a treatment which is rated the best for him or her "individually"[3] (1) by at least one of the physician experts, or (2) by a "combination" of the expert physicians.[4] Second, the opinions of the physicians are continually updated as the trial progresses in the light of the data generated by the trial.

This updating is reflected in the class of admissible therapies for any given patient. This accruing data and its supporting or diminishing effects on competing therapeutic modalities, however, cannot easily be made available to either patient or physician. The difficulty (for Bayesian and classical statisticians alike) is that the use of such interim data would likely bias patient selection. To control for this, should the interim information be made available, would require a model of patient and/or physician actions given the data, and such a model is just not available. For similar bias-avoidance reasons, as Kadane notes in his paper, this new design also does not permit a patient to choose his or her own therapy. The Kadane design does, however, open the way for, probably in later phases of development of the methodology, the explicit use of the patient's utilities in defining an acceptable set of treatments for that particular patient.

In contrast to a living flesh-and-blood committee, the Kadane "panel" is a computer-simulated entity that possesses an informational structure initially obtained from several experts in a clinical discipline. Though the analogy is not one Kadane uses, the design is, in a way, a kind of "expert system." The "panel" thus really consists of a set of probability models for a large number of reasonably-finely differentiated patient groups. Kadane describes the "new Bayesian technology" by which this information is elicited in his essay. This modeling utilizes a Bayesian probability approach, and as evidence is developed by the trial, Bayes's theorem is then used to automatically update the belief commitments of the experts, or more accurately the models of the experts. To ascertain that the Bayesian updating accurately reflects the experts actual change of opinion, the experts are periodically reconsulted (probably every six months or so, though this depends on the pace of patient entry into the trial and the results obtained).

A key notion in the Kadane type of trial is that of "admissibility". As defined above, a treatment is admissible or assignable to a patient if *one* or possibly a "convex combination" of experts assess that treatment as best for that type of patient. This design thus protects the patient from the ***worst treatment***, but it does not necessarily assign the "best by majority vote" treatment. This notion of "the best by majority vote" refers to the possibility that four of the five expert models, say, might view treatment B as the best, but the dissenting "expert's" opinion that A is the best *could* result in A being assigned.

This possibility has resulted in spirited debate among the statisticians, lawyers, and ethicists involved in discussions of Kadane's type of trial, and thus far no clear consensus has emerged. We cannot even argue that this failure to assign the "best" treatment is acceptable at the beginning of a trial on the grounds that what is viewed as "best" is based "only on opinions". From a Bayesian perspective even prior opinions are recognized to have some probative force, even though they may be weak and diffuse. One suggestion developed in Kadane's article is to distinguish these expert opinions from "votes" in any traditional sense, since "the expert panel is chosen to represent the *spectrum* of the medical community, but not to reflect the appropriate weight (if such could be specified) that should be given to each view . . . " (Kadane, 1986, p. 399) (my emphasis). We could thus consider treating the four opinions that are in agreement as equivalent to one opinion, since they do not differ conceptually. This, however, discards the kind of information which is represented by "second opinions" in medical cases, and is difficult to defend morally or legally. Morally it appears to diminish autonomy, since many individuals would feel that the four to one distribution was materially relevant to their decision to accept that therapy. I believe it may also be legally suspect on informed consent grounds.

One solution may be to obtain the informed consent of the patient by presenting the possibility that this situation could occur, and having the patient waive the right to this type of information. (This solution would, however, involve acceptance of the precepts of the contract model, noted earlier to rest on a somewhat shaky foundation.) Obtaining a valid informed consent in such circumstances might require that just the above scenario, or a very similar scenario, be presented to the patient as part of the consent form. Not to disclose such information could violate a 'respect for persons' principle and other ethical principles such an nonmaleficence mentioned earlier. Such disclosure, however, raises the not inconsiderable risk of patient refusal to participate in such a trial. This may require some variant of prerandomization superadded to the Kadane design in order to increase patient numbers. Exactly how such a procedure might work would need

to be investigated in more detail by the proponents of the Kadane methodology.

The Kadane design is quite novel and does insure against a uniform worst therapy assignment. It is still in the process of development, however, and issues such as the appropriateness of the "vote" metaphor and its possibly misleading implications remain under active discussion. The fact that the trial design has been approved by Johns Hopkins's IRB and its successful recruitment of patients to date, as described in Kadane's article, augurs well for its further development and acceptance.

IV. CONCLUSION

Having come this far, the reader has, I believe, a good sense that the ethical issues involved in clinical trials constitutes both a complex and a constantly evolving subject area. In this essay, I have attempted to highlight some aspects of the papers that follow, and to provide a historical and methodological context in which to interpret them. Some issues, such as "deferred consent", could not be addressed here for lack of space, though that does not indicate that these problems are less important. I have also had to treat some arguments in less than full detail in order to provide some critical suggestions which I hope will generate further debate on these topics. Clinical trials will be with us for the foreseeable future, and additional developments on both methodological and ethical fronts are certain to occur.

NOTES

[1] See Lasagna (1962, Ch. 1) and Wulff (1981, Ch. 9) for examples.

[2] The usual reference to support this claim is to the Oath of Hippocrates, though as Fried (1974, n. 6, pp. 50–51) indicates, this is ambiguous. Other codes such as the Declaration of Geneva of the World Medical Association and the International Code of Medical Ethics are quite clear, however, that the physician must act in the patient's best interest. See generally Katz (1972, pp. 311–321).

[3] "Individually" is placed in quote marks because expert opinion is elicited not for a given known patient but rather for relevantly distinguishable patient *subtypes*.

[4] The "combination" of experts is a "convex" combination. See Kadane's response to hypothetical question L in his article in this *Journal*.

REFERENCES

Armitage, P.: 1960, *Sequential Medical Trials*, Blackwells, Oxford.

Armitage, P.: 1984, 'Controversies and achievements in clinical trials', *Controlled Clinical Trials* 5, 67–72.

Beauchamp, T. L. and Childress, J. F.: 1983, *Principles of Biomedical Ethics* (2nd ed.), Oxford University Press, New York.

Beecher, H. K.: 1966, 'Ethics in clinical research', *New England Journal of Medicine* 274, 1354–1360.

Byar, D. P. *et al.*: 1976, 'Randomized clinical trials', *New England Journal of Medicine* 295, 74–80.

Chalmers, T. *et al.*: 1972, 'Controlled studies in clinical cancer research, *New England Journal of Medicine* 287, 75–78.

Chalmers, T. *et al.*: 1983, 'Bias in treatment assignments in controlled clinical trials', *New England Journal of Medicine* 309, 1358–1361.

Clayton, D.: 1982, 'Ethically optimized designs', *British Journal of Clinical Pharmacology* 13, 469–480.

Cornfield, J.: 1966, 'A Bayesian test of some classical hypotheses with applications to sequential clinical trials', *Journal of the American Statistical Association* 61, 577–594.

Cornfield, J.: 1969, 'The Bayesian outlook and its application', *Biometrics* 25, 617–657.

Coulehan, J. L., Schaffner, K. F., and Block, M.: 1983, 'Ethics of clinical trials in family medicine', *The Journal of Family Practice* 21 (3), 217–222.

Declaration of Helsinki (1964), reprinted in Beauchamp and Childress (1983), pp. 339–343.

Department of Health and Human Services (DHHS): 1981, 'Department of Health and Human Services Rules and Regulations', *Federal Register* 46, 8366–8392. Partially reprinted in Levine (1981), 259–273.

Ellenberg, S.: 1984, 'Randomization designs in comparative clinical trials', *New England Journal of Medicine* 310, 1404–1408.

FDA (1981) 'Protection of human subjects: informed consent', *Federal Register* 46, 8942–8980. Partially reprinted in Levine (1981), 275–284.

Feinstein, A. R.: 1986, *Clinical Epidemiology: the Architecture of Clinical Research*, Saunders, Philadelphia.

Fisher B. *et al.*: 1985, 'Five-year results of a randomized clinical trial comparing total mastectomy and segmental mastectomy with or without radiation in the treatment of breast cancer', *New England Journal of Medicine* 312, 665–673.

Fletcher, R. H., Fletcher, S. W., and Wagner, E. H.: 1982, *Clinical Epidemiology – The Essentials*, Williams and Wilkins, Baltimore.

Freireich, E. and Gehan, E.: 1979 'The limitations of the randomized clinical trial', in V. DeVita and H. Busch (eds.), *Methods in Cancer Research: Cancer Drug Development* XVII, Part B, Academic Press, New York, pp. 277–310.

Fried, C.: 1974, *Medical Experimentation: Personal Integrity and Social Policy*, American Elsevier, New York.
Gail, M. H.: 1982, 'Monitoring and stopping clinical trials', in V. Mike and K. E. Stanley (eds.), *Statistics in Medical Research*, Willey, New York, pp. 455–484.
Gifford, F.: 1986, 'The conflict between randomized clinical trials and the therapeutic obligation', this issue, pp. 347–366.
Hoel, D. G., Sobel, M., and Weiss, G. H.: 1975,'A survey of adaptive sampling for clinical trials', in R. M. Elashoff (ed.), *Perspectives in Biometrics*, 1, Academic Press New York, pp. 29–61.
Jones, J. H.: 1981, *Bad Blood: The Tuskeegee Syphilis Experiment*, The Free Press, New York.
Kadane, J. B.: 1986, 'Progress toward a more ethical method for clinical trials', this issue, pp. 385–404.
Kadane, J. B. and Sedransk, N.: 1980, 'Toward a more ethical clinical trial', in Bernardo *et al.* (eds.), *Bayesians Statistics*, University of Valencia, pp. 329–338.
Katz, J.: 1972, *Experimentation with Human Beings*, Russell Sage, New York.
Kopelman, L.: 1986, 'Consent and randomized clinical trials: are there moral or design problems?', this issue, pp. 317–345.
Lasagna, L.: 1962, *The Doctors' Dilemmas*, Harper and Brothers, New York.
Lellouch, J. and Schwartz D.: 1971, 'L'essai therapeutique: ethique individuelle ou ethique collective', *Revue Inst. Int. Statist* 39, 127–136.
Levine, R.: 1981, *Ethics and Regulation of Clinical Research*, Urban and Schwarzenberg, Baltimore.
Levine, R. and Lebacqz, K.: 1979, 'Some ethical considerations in clinical trials', *Clinical Pharmacology and Therapeutics* 25, 728–741.
Lidz, C. W. *et al.*: 1983 *Informed Consent: A Study of Psychiatric Decision-Making*, Guilford Press, New York.
Lilienfeld, A. M., and Lilienfeld, D. E.: 1980, *Foundations of Epidemiology*, 2nd edition, Oxford University Press, New York.
Marquis, D.: 1983, 'Leaving therapy to chance: an impasse in the ethics of clinical trials', *Hastings Center Report* 13 (4), 40–47.
Marquis, D.: 1986, 'An argument that all prerandomized clinical trials are unethical', this issue, pp. 367–383.
Medical Research Council,: 1948, 'Streptomycin treatment of pulmonary tuberculosis', *British Medical Journal* 2, 769–782.
Meier, P.: 1979,' Terminating a trial – the ethical problem', *Clinical Pharmacology and Therapeutics* 25, 633–640.
Redmond, C. and Bauer, M.: 1979, 'Statisticians' Report on Prerandomization', *NSABP Progress Report*, National Surgical Adjuvant Project for Breast and Bowel Cancers, Pittsburgh.
Schaffner, K.: 1983, 'Clinical trials: the validation of theory and therapy', in L. Laudan and R. Coben (eds.), *Physics, Philosophy, and Psychoanalysis*, D Reidel, Dordrecht, Holland, pp. 191–208.

Taylor, K., Margolese, R., and Soskolne, C. L.: 1984, 'Physicians' reasons for not entering eligible patients in a randomized clinical trial of surgery for breast cancer', *New England Journal of Medicine* 10, 1363–1367.
Urbach, P.: 1985, 'Randomization and the designs of experimentation', *Philosophy of Science* 52, 256–273.
Veatch, R. M.: 1979, 'Longitudinal studies, sequential design, and grant renewals: what to do with preliminary data', *IRB* 1 (4), 1–3.
Wahrendorf, J., Blettner, M., and Elder, L.: 1985, 'An investigation of the adequacy of a randomization', *Controlled Clinical Trials* 6, 249–258.
Weinstein, M. C.: 1974, 'Allocation of subjects in medical experiments', *New England Journal of Medicine* 291, 1278–1285.
Wulff, H. R.: 1981, *Rational Diagnosis and Treatment*, 2nd edition, Blackwells, Oxford.
Zelen, M.: 1969, 'Play the winner rule and the controlled clinical trial', *Journal of the American Statistical Association* 64, 131–146.
Zelen, M.: 1977, 'Statistical options in clinical trials', *Seminars in Oncology* 4, 441–446.
Zelen, M.: 1979, 'A new design for randomized clinical trials', *New England Journal of Medicine* 300, 1242–1245.