Scientific Rigor and Medical Realities: Placebo Trials in Cancer and AIDS Research

David J. Rothman and Harold Edgar

from Elizabeth Fee and Daniel M. Fox, AIDS: The Making of a Chronic Disease, University of California Press, 1992

In the recent debates on the ethics of placebo-based trials in the evaluation of new drugs to combat AIDS, a sharp line is often drawn between the need to satisfy the principles of “sound science” and the readiness to satisfy “humanitarian impulses.” The proponents of sound science contend that the only procedure that will demonstrate the efficacy of a new agent is a placebo trial in a population that is randomly selected; in such a trial both the subjects and the investigators must be ignorant, or blinded, as to who is receiving the active agent and who the inert substance. Although half the subjects will receive an inactive ingredient, proponents believe that the long-term good of establishing knowledge outweighs all other considerations. Those opposed to placebo trials contend that, although the new substance is of unknown efficacy, it may work, and therefore may give persons with AIDS an opportunity, both psychological and pharmacological, to extend their lives. But an evaluation of the ethics of placebo trials in the AIDS era does not require us to pursue an either/or approach, a rigid opposition of scientific progress and compassion, with no ground between these two extremes. In fact, the choices are not so stark, mostly because the principles of sound science are not so rigid and immutable as many of its advocates insist. The placebo-based random clinical trial does not have the hegemony in drug development that its proponents suggest — and the departures from the standard have not come at the price of ignorance or maleficiency. The evidence for this proposition comes most powerfully from cancer research. The way cancer researchers have pursued drug development casts a very different light on the AIDS controversies and deserves sustained analysis.

Long before the AIDS crisis, the issue of scientific rigor and patient needs was confronted in the cancer research field, surfacing most notably in the disputes that marked the relationship of the National Cancer Institute (NCI) and the federal Food and Drug Administration (FDA). The NCI, a government-funded research organization, actively develops and tests new drugs. In formal terms it has the same relationship to the FDA that any other drug manufacturer has; that is, the drugs it develops must be licensed by the FDA before distribution. But in reality the FDA-NCI relationship is far more complicated, and the NCI generally follows special procedures that depart from FDA requirements. Probably the most important difference between the two organizations is on the matter of demonstrating drug efficacy, in effect, on the kinds of clinical trials appropriate to demonstrating efficacy.

A 1982 congressional hearing clarified these differences. The hearings were occasioned by a series of articles in the Washington Post, describing protocols in which cancer patients were ostensibly used as guinea pigs in research. These patients, the article contended, were receiving drugs that investigators knew were too toxic, or ineffective, and the FDA was failing to supervise or regulate their work. (Even as late as 1981, the predominant fear among outsiders was not that patients were unable to enroll in protocols but that patients would be misused by researchers.) The cancer investigators, for their part, insisted that the patients were fully informed about the risks and benefits, that drugs ineffective against one type of tumor might be effective against another type, and that high drug toxicity was unavoidable in light of the present state of knowledge. But what emerged most vividly in the course of the hearings was the shared commitment among cancer researchers to doing something, anything, for the terminally ill cancer patient. When death was the alternative, they were ready to try new and admittedly dangerous drugs on patients who wanted a shot, even a long shot, at a remission or cure, and if this commitment brought them into conflict with the FDA, or with the gold standard of random clinical trials (RCTs), so be it. The first loyalty was to the patient.

Vincent DeVita, director of the NCI, explicated this position fully. “The most serious toxicity of all,” he declared in his testimony, “is the unnecessary death from cancer.... Any system of drug distribution we develop that denies any cancer patient access to these resources is wrong.” The NCI arrangements aimed to maximize distribut
out sacrificing oversight. New cancer drugs (all cytotoxic—that is, by definition injurious to normal cells) were designated A, B, or C, with a rough but not complete analogue to Phase I, II, III categories. Drugs in category A were tested first on patients with advanced disease by oncologists in ten designated institutions; should any of the drugs appear promising (by evidence of tumor shrinkage or improved quality of life), they moved to category B, to be tested by a larger group of selected clinical investigators on a wider range of patients. The drugs that demonstrated effectiveness were then promoted to category C, to be distributed to a still wider network of designated practicing physicians (those sponsored by NCI grants or contracts). This distribution was akin to the FDA’s “compassionate use” procedure but was much more extensive and systematic.

Several aspects of this system made clear the extent to which cancer drug development was treatment oriented. First, this considerable distribution of drugs took place before the FDA actually licensed them. Second, drugs were moved into category A without extensive animal tests (since they were known to be toxic). Again and again cancer researchers made the point that the true toxicity was cancer. As Dr. James Holland of New York City’s Mt. Sinai Hospital put it: “Can it be more ethical to deny the possible good effects to most, by avoiding all toxicity in order to do no harm to one? The unmitigated disease must be calculated as a toxic cost of cancer. Underdosing, in an attempt to avoid toxicity, is far more deadly.”

Third, no one at NCI disputed that “leakage” occurred; that is, physicians who received drugs in A or B category and did not exhaust their supply sent the remainder on to still other physicians for use with their patients. DeVita was not very apologetic about the leakage, insisting that “sometimes patients benefit.”

Finally, and perhaps most important, and we will return to explore this point in more detail later, the trials with cytotoxic drugs against advanced cancers were almost always single armed—that is, not controlled and not placebo based.

The cancer investigators in their testimony made no secret of their disdain for the FDA regulatory apparatus. “Innovation and regulation are constantly in conflict,” argued Dr. Emil Freireich, of the University of Texas System Cancer Center, and formerly at the NCI. “In our country we have gone extremely to the side of regulation, much to the detriment of innovative creative science. . . . It is truly ironic that the mechanisms designed for protection create serious harm to thousands of individuals with cancer without any potential for benefit . . . . Speaking as a physician-scientist . . . there is continuous frustration resulting from excessive regulation. . . . It is clear that any new knowledge requires additional risk.”

Indeed, these researchers were impatient not only with the FDA but with the idea of government paternalism, and if some of this attitude may have been the product of professionals wanting to maintain ample discretion, it also reflected a deep concern for the desperately ill patient. When Dr. John Ulmann, director of the Cancer Research Center at the University of Chicago, was asked whether in category A or Phase I studies the researchers might be sacrificing patient welfare for scientific knowledge, he insisted that “throughout this process, above all else we are doctors.” And when California congressman Henry Waxman invoked the need for the government “to protect the public from drugs that are going to kill them, poison them, maim them,” Dr. Holland reminded him that with cancer drugs the injunction to “do no harm” was meaningless, for “all the patients who would have benefited will be undertreated.”

In the immediate aftermath of the 1981 hearing, the NCI and the FDA established a joint task force. Its report, aptly titled Anticancer Drugs: The NCI’s Development and the FDA’s Regulation, spelled out further differences between the more patient-centered risk-taking procedures at the NCI and the more paternalistic and “sound science” oversight at the FDA. For one, the cancer researchers were so committed to patient care that they were unwilling to continue to test a new drug against all types of tumors when the drug had shown little efficacy in its initial tumor screens. Conceding that some drugs had proven effective only against one or two types of tumors (and if the screening had not been complete, this efficacy would have been missed), the researchers were nevertheless “reluctant to enroll patients with a given tumor in a study of a drug already shown ineffective in several other tumors; they prefer, instead, to try a drug with which there is little prior experience.”

For another, the FDA required that, before any drug could be licensed to be used in combination therapies (together with other drugs), its own individual efficacy had to be established. By the gold standard, drug X should not be added to drugs Y and Z unless drug X had independently demonstrated its efficacy. The cancer researchers took a contrary position; they were ready to go with what worked, regardless of testing requirements. Thus, in Phase III studies the NCI focused on the patient and the disease, not on the drug. “Most research oncologists are convinced that they will obtain the best therapeutic results with drug
combinations, so that the design at this stage that will clearly illustrate the value of a drug may appear unethical." The FDA staff accused the NCI of being unwilling to do proper testing; the NCI responded that over the past decade, as a result of its testing methods, a number of drugs had found "secure places in the practice of clinical oncology and ... overall survival of cancer patients has improved." 9

The task force also had to address the issue that the Washington Post had raised about the appropriateness of using cancer patients in Phase I, or NCI category A, tests for toxicity. Here, too, it concluded that patients should be permitted to make their own determinations of risk and that the FDA should not decide what risks were or were not allowable. "While the Task Force recognized that people do not have an absolute right to harm themselves consciously, neither should they be absolutely precluded from seeking treatment which holds out hope of benefit." Even an overall response rate of 9.5 percent (the average response rate to Phase I drugs) was reason enough to let the patient make the choice.10

The task force then confronted two especially controversial aspects of NCI procedures. First was the NCI's unwillingness, and the general unwillingness of cancer researchers, to adhere to the placebo trials. Although it acknowledged the widespread perception that "NCI protocols are not scientifically adequate; they are biostatistically flawed," the task force unapologetically defended the NCI procedures in language that is well worth scrutinizing: "There are difficulties in creating ethical controlled trials in a uniformly fatal disease, and there are restrictions on the number of patients to be studied because of the known drug toxicity. Ideal experimental design must be compromised to achieve the best possible patient care. As a result, many Phase II studies have used historical controls, and Phase III studies [have used] combination therapies. Thus they may not be compared with experiments that can be performed in other kinds of illnesses." 11 In other words, the need to treat desperately ill patients ruled out the use of placebos or the testing for efficacy of the individual drugs that went into combination therapies.

This same rationale supported the distribution of the drugs in group C testing. Although the efficacy of these drugs had not been proved by FDA standards, and although the drugs were being distributed to hundreds of physicians, the task force defended the procedure. First, it noted that the drugs were distributed only to a selected group of qualified physicians; second; "The Task Force believes that Group C status is an appropriate method for bringing important medications to patients who need them." 12 Once again the needs of patients took first priority.

The task force's endorsement of the practice of distributing cytotoxic drugs whose efficacy had not been established in placebo-based clinical trials was only the latest entry in a decade-long debate on the standards that should be satisfied before drugs were made available. This same controversy erupted, with even more heat, around the release of AIDS drugs. Were the new agents to undergo placebo trials? Should the active agent be given to all subjects, and its efficacy measured against past knowledge of the course of the disease?

The FDA, in fact, does not insist on placebo-controlled studies or rule out the use of historical controls. Its 1985 regulations defining "adequate and well-controlled studies" (section 314.126) open with the statement: "The purpose of conducting clinical investigations of a drug is to distinguish the effect of a drug from other influences, such as spontaneous change in the course of the disease, placebo effect, or biased observation." It then lists five types of controls that are "recognized." The first is placebo concurrent control; the others include a "no treatment concurrent control," or control through comparison with another active agent. Fifth—and by no coincidence last on the list—is the historical control: "The results of treatment with the test drug are compared with experience historically derived from the adequately documented natural history of the disease . . . in comparable patients or populations. Because historical control populations usually cannot be as well assessed with respect to pertinent variables . . . historical control designs are usually reserved for special circumstances. Examples include studies of diseases with high and predictable mortality (for example, certain malignancies)" (italics added). Hence, the FDA does accept historical controls as a type of control in clinical trials, in contradistinction to the reliance on "isolated case reports" or "random experience."

Many investigators object to this position, insisting that historical controls are never an adequate base for measuring the efficacy of a drug. One of the most persistent critics has been Thomas Chalmers, who was dismayed to report, on the basis of a survey of abstracts presented to the 1971 meeting of the American Association for Cancer Research, that only 21 percent of the protocols had used clinical trials. He found it "surprising that this crucial concept has not caught on to a greater extent" and marshaled arguments for its use. Noting that "many clini-
cal investigators believe that they cannot deprive their patients of the opportunity to receive a new drug," he countered that "the experience with every new cancer drug, when it is introduced into man, is such that either the risk of drug toxicity and mortality is greatest during its early use, or impotent doses are used at first to avoid unknown toxic effects. In either case little benefit to the first patients treated with the new agent can be anticipated." Moreover, Chalmers continued, if a drug shows some signs of early efficacy, investigators will then not undertake randomized trials but will accept the pilot test results as definitive. From Chalmers's perspective, the only way to avoid the predicament is to randomize from the first patient; otherwise, a state of ignorance is certain to prevail.13

Franz Ingelfinger, then editor of the New England Journal of Medicine, ran an accompanying editorial to the Chalmers article, supporting his insistence on the randomized clinical trial. Noting the "ethical and emotional" objections to the trials, Ingelfinger declared: "It is an investigator of strong moral and intellectual fiber who would resist the urge to 'do something' for a fatally sick patient... who would use 'cold science' when the pressures are all on the side of warm hope." But researchers must rise to the challenge: "Ethical, as well as scientific, considerations require that medicine depend on the most reliable and best controlled data available—the kind of data that is sought by the randomized clinical trial."14

The types of arguments that Chalmers and Ingelfinger raised in defense of the RCT are familiar and have been often repeated. What is more needed is a full explication of the counterposition, one that goes beyond "warm hope" for the subject or the weak fiber of the researcher-clinician. At its core is the proposition, conceded by the FDA but ignored by Chalmers and the others, that placebo-based trials are ethically inappropriate in the case of a "uniformly fatal disease." In the standard medical text on cancer, a chapter on the design of clinical trials declares: "To determine whether a new treatment cures any patients with a disease that is uniformly and rapidly fatal, history is a satisfactory control... Are randomized trials necessary for identifying major advances in treatment? No. There are many examples of therapeutic breakthroughs that were recognized without randomized trials. For the most part, however, these occurred in diseases where the prognosis was 100% predictable before the advent of the new therapy, and hence there was no possibility of bias with regard to patient selection."15

This position was also advanced by David Byar, of the NCI's Clinical and Diagnostic Trials Section, in an essay on the "Necessity and Justification of Randomized Clinical Trials." Byar listed six difficulties with historical controls (from missing data to failure to convince others of the results), presented another six arguments in favor of randomized trials (bias is avoided, time trends are no problem, fewer patients need to be treated to get a convincing answer), and comfortably declared that the ethical dilemmas in RCTs were resolved because "there is always some cost in learning something." But Byar was also prepared to support nonrandomized studies "when a new treatment appears that is markedly effective for a disease which before that time was virtually incurable." Those situations, he cautioned, might be rare, but when they arose, "it would be difficult or impossible to justify a randomized study... from an ethical point of view."16

Some cancer researchers were even prepared to go further in undercutting an exclusive reliance on RCTs. Drs. Edmund Gehan and Emil Freireich, of the M. D. Anderson Hospital and Tumor Institute, writing in 1974, insisted: "A clinical investigator has an ethical responsibility for his patients when they are involved in a clinical trial... to administer... the treatment that gives him the highest probability of a successful outcome... If preliminary clinical studies suggest that a new treatment is significantly more effective than a standard... the physician would not be fulfilling his ethical responsibility if he planned a randomized comparative trial." Hence, the authors concluded: "In clinical trials it is unwise to assign patients to treatments by any single method. In the field of cancer chemotherapy, effective new therapies have been detected, confirmed, and applied widely in practice as a result of prospective and quantitative clinical trials that have not used random allocation of patients... The widespread acceptance of the randomized comparative trial seems based... more on the intuitive attractiveness of the technique than on any objective scientific evaluation of the methodology."17

Thus, for a number of reasons, the placebo-based RCT is not the gold standard in cancer research, not in principle and not in fact, when the disease is uniformly fatal or virtually incurable. Cancer researchers have openly made the case against the monopoly of RCTs in trial design and in practice have prepared to avoid them, even at the risk of not satisfying FDA procedures. As Dr. Robert Wittes of the NCI concluded: "The placebo or no-treatment control has always had a very limited role in the evaluation of cytotoxic therapy in advanced cancer... Clini-
cal oncologists in the United States have been generally unwilling to randomize a patient with advanced progressive cancer to placebo or observation alone.” Instead, “the clearest demonstration of a beneficial effect on survival might only come from a comparison with a carefully selected and characterized historical control group.” And by 1989 at least, the FDA was on the whole ready to accept the position. In a “Talk Paper” on “Approval of New Cancer Drugs,” issued March 3, 1989, as part of a series of papers “to guide FDA personnel in responding with consistency and accuracy to questions from the public on subjects of current interest,” the FDA declared that in decisions to approve new drugs “neither safety nor effectiveness is absolute, but must be weighed in particular cases” and that “although randomized clinical trials . . . are the preferable means of evaluation, other study designs may be acceptable, especially for refractory diseases (those malignancies which do not respond to standard therapy), where a clear response may be apparent even without a randomized control.”

With the cancer research model to mind, let us now examine the controversies around research design and AIDS drugs, focusing first on the AIDS Drug Development hearings conducted in July 1986 by Congressman Ted Weiss, and then on the ethical dimensions of the decision to make the first large-scale AZT trial placebo based.

The Weiss hearings confronted directly the issue of placebo-based trials, and the testimony split along the lines that we have been tracing. Proponents of the classic-style RCT came predominantly from the realms of infectious diseases, the FDA, and the drug companies (whose products, after all, must pass FDA review). Their model was not the cancer model; the research in AIDS was not to follow on the designs for research in cancer. “We have learned,” declared Harvard professor of medicine and infectious disease specialist Martin Hirsch, “in clinical trials of antiviral agents against other fatal diseases . . . that placebo controls are mandatory until an effective agent is found. The same procedures must be followed in HIV infections, or we will pay the price in unnecessary delays and unwarranted deaths. . . . Until you have some evidence of efficacy of a drug you are still justified in doing placebo-controlled trials even in a fatal condition, such as AIDS, because you may do harm with any of these drugs.”

Anthony Fauci, director of the National Institute of Allergy and Infectious Diseases, and Dr. Harry Meyer from the FDA both tried to differentiate AIDS from intractable cancer. “Although one can project that within a five-year period most of the patients will succumb to disease,” argued Fauci, “in fact, to those of us who see AIDS patients every day, it becomes very clear that the natural history is quite variable. One of the great problems that we could create for ourselves would result from using a control that is not an adequate control and feeling that an agent was helpful when it really was not.” This position was also defended by Dr. David Barry, vice-president of research, Burroughs Wellcome (the manufacturer of AZT): Stating unequivocally that AIDS was unlike cancer, he argued that “because of the waxing and waning of some of the clinical manifestations of AIDS, we could not do an uncontrolled study.”

Dr. Mathilde Krim, herself a cancer researcher before she established the American Foundation for AIDS Research, most explicitly made the case for having AIDS research follow on the cancer model. (Surprisingly, she remains one of the very few commentators in this debate to do so—the “plague-like” quality of AIDS apparently made the cancer model seem as irrelevant in the laboratory as in the design of the delivery of care.) Her arguments drew on the traditions in cancer research. Noting that “ethically and scientifically satisfying alternatives to placebo-controlled trials have been devised for the study of experimental drugs in cancer patients,” she asked why they were not being used in AIDS. Observing, as well, that experimental (group C) drugs were made available to cancer patients before FDA licensing, she wondered why AIDS patients were not coming under the same policy. After all, “AIDS is presently more surely lethal, within a shorter time, than most cancers. There is no known accepted treatment.”

The differences that emerged at the Weiss hearing were anything but academic. At that very moment, the first large-scale trial on the new drug AZT was being conducted, and the trial was placebo based. AZT had been first tested on 19 patients with AIDS and ARC; and the highly promising findings from this six-week trial, in which all patients received the drug, were published in March 1986 in Lancet. To review some of the highlights: the patients generally tolerated the drug well, 15 of the 19 had increases in helper T cells, 13 patients had a weight gain of 2 kilograms or more, and 6 patients noted cessation of fever or night sweats and an improvement in their sense of well-being. The published report of the study concluded that the trial did not demonstrate whether immunological improvements would be sustained, whether AZT could be tolerated over a long time, whether viral drug resistance would de-
velop, or whether AZT would affect disease progression or survival. "These are issues which can be resolved only by appropriately controlled long-term studies."23

On the basis of these findings, a multicenter, placebo-based trial was undertaken in 282 patients; 145 subjects received AZT, and 137 received placebo. Of the patients with AIDS, all had experienced a first episode of Pneumocystis carinii pneumonia (PCP) within 120 days; patients with ARC had notable weight loss or other symptoms, such as herpes zoster or lymphadenopathy. The multicenter study was terminated after twenty-four weeks because the first results demonstrated the efficacy of the drug: over this period, 19 subjects in the placebo group but only 1 in the AZT group died. More generally, in 1986, patients with AIDS and PCP had a median survival of twelve and a half months, and after twenty-two months three-quarters of these patients were dead.

Was the design of this trial ethically proper? Should 137 patients have received placebo? This question was actually the subtext of the testimony at the Weiss hearing. Dr. Krim, in effect, said no, asking why "any AIDS patient should be forced to accept cornstarch pills... This practice has long been abandoned in the experimental treatment of patients with advanced cancer."24 From her perspective, a median survival of twelve and half months made AIDS an intractable and uniformly fatal disease, and historical controls would have been sufficient to establish efficacy. On the other side, the remarks by Drs. Fauci and Meyer about the standards for research and the variability of the disease patterns in AIDS were clearly intended to defend the protocol's design.

Whatever the nature of the dispute, it is apparent that were AZT an anticancer drug, the trial would not have been placebo based. Had 19 patients with advanced cancer and no known therapeutic agent done as well as the first 19 patients on AZT, the next trials would have given the drug to all subjects. The goals of treatment would have taken first precedence. Put another way, that the AZT trials were placebo based testifies to the fact that the treatment of AIDS was based not on the cancer model but on a more generalized medicine model, really an infectious disease model.

This formulation has several implications that merit notice. First, in light of the initial definitions of what constituted the AIDS crisis, it is not surprising that the research design followed an infectious disease model. AIDS was a plague, an infection, the result of a viral agent, not a chronic illness of cellular origin. And those working in infectious dis-

8ases, unlike those in cancer research, generally had considerably less day-to-day contact, and less intense contact, with terminally ill patients than their counterparts in oncology. Most of the research in infectious diseases, although certainly not all, did not involve desperately ill patients willing to take high risks for the slimmest possibility of a gain. Inevitably, in the realm of infectious diseases, the commitment to placebo-based random trials did not have to come up against agonizing questions.

By the same token, the FDA staff, driven for a variety of reasons to maximize safety and minimize risk, were also committed to rigorous RCTs; and the group that stood out against this orientation, the cancer researchers, had over the years been able to insulate their operations, through the NCI, from systematic FDA oversight.25 Hence, it was the infectious disease--FDA model, not the cancer model, that structured the design of the AZT tests.

A recognition of this process has a direct relevance to deliberations on the ethics of clinical trials, for it makes apparent that science comes in a variety of models, and the process by which one or another subsumes a particular area of medicine is determined not by immutable canons of research but by historical and social contingencies, or, if you will, by metaphors. Since the first designation of AIDS was of a plague, not a chronic disease, the models of infectious disease, not cancer, took hold. Put another way, a committee charged to analyze the ethics of trials is confronting a choice not between science and compassion but between which model of science is most appropriate to AIDS.

NOTES

The authors would like to acknowledge research support from the American Foundation for AIDS Research.

2. Ibid., p. 256.
3. Ibid., p. 218.
4. Ibid., pp. 261, 264, 265, 268.
5. Ibid., p. 272.
6. Ibid., pp. 279, 283.
8. Ibid., p. 33.
9. Ibid., p. 36.
10. Ibid., p. 69.
11. Ibid., pp. 90–91 (italics added).


20. Ibid., p. 104.

21. Ibid., pp. 115–16.

22. Ibid., pp. 23, 38–39.


24. Ibid., p. 39.