

Trials on Trial

Research ethics generally fails to capture public attention and scrutiny. But a debate over clinical trials in developing countries moved suddenly into the public domain last fall, when an editorial in the *New England Journal of Medicine* criticized studies designed to test the efficacy of antiretroviral drugs in reducing mother-to-infant transmission of HIV. The editorial objected to the trials because they included placebo-control groups, in which HIV-infected pregnant women were given a dummy pill rather than the drug zidovudine (AZT). The criticism was especially pointed because in nine of the fifteen trials then under way, funding had been provided by U.S. health agencies—the National Institutes of Health (NIH) and the Centers for Disease Control and Prevention (CDC).

Recently, the CDC announced that it was stopping a placebo-controlled trial in Thailand, not because of ethical objections, but because the question under study had now been answered to the agency's satisfaction. Data from the trial showed that short courses of AZT, administered prenatally and during delivery,

All clinical research on human subjects strives to balance the interests of the study subjects in receiving the best available treatment for their illnesses against the interests of society in acquiring the most dependable information at the earliest time. One can grasp this tension by thinking about one's own experience *as a patient*. When ill or in pain, the patient wants the best available treatment. Since a controlled clinical trial requires the random assignment of patients to two or more arms of the study offering different therapies, it is only when there is no reason to believe that one therapy is superior that it is not against the patient's interest to participate. If one treatment is better, why should a patient take a chance of not getting it? At the same time, *as a patient* one depends on the doctor having

As a patient, one hopes that prior patients have participated in studies that show clearly and decisively what treatment is best.

All clinical research on human subjects strives to balance the interests of the study subjects against the interests of society.

reduce mother-to-infant transmission of HIV by one half. As a result, other trials supported by the NIH and CDC will no longer treat HIV-infected pregnant women with placebos. Both critics and defenders of the trials are pleased with this development; the critics because use of placebos has been halted, the defenders because the trials produced decisive, reliable data that policymakers in developing countries can use.

The Basic Dilemma

The research that sparked the controversy presents the classic conflict inherent in all studies of promising new therapies carried out on human subjects. As we shall see, the fact that these trials were conducted in developing countries does add complexity to the issue. But the basic dilemma of appropriate research design is familiar.

reliable information about what treatments are useful. One hopes that prior patients have participated in studies that show clearly and decisively what treatment is best.

The AZT trials in developing countries replay this classic dilemma. The HIV-infected woman wants the therapy most likely to prevent transmission of the disease to her infant. The community (including its public health policymakers) wants reliable information about the effectiveness of treatment so that it can implement programs to help the many other infected pregnant women and their babies.

Several facts about these trials heighten the familiar conflict, however. HIV infection is a deadly disease. Hence, the use of a placebo in the case of HIV especially compromises the interests of the study subject (here, the baby). Since the subject in these trials cannot provide consent, participation requires consent by surrogate (the mother). Yet in cases involving surrogate consent, researchers generally have a special responsibility to safeguard the well-being of the dependent.

The interests of the community are also substantial. HIV infection rates in many developing countries are extremely high. For example, in Kampala, Uganda,

one-sixth of the adult population is infected. In addition, most developing countries have few resources available for public health. It is thus imperative to know whether a treatment is effective before diverting extremely scarce resources from other health programs.

The use of a placebo was especially controversial because an effective means of reducing the likelihood of transmission has been identified. In 1994, the results of a study carried out in the U.S. and France demonstrated that administering AZT to the mother pre- and perinatally, and to the infant after delivery, dramatically reduced transmission. This regimen, known as the ACTG 076 protocol, cut the incidence of HIV transmission from mother to child by two thirds. As a result of the clear benefit of the 076 regimen, it is the recommended treatment for HIV-infected pregnant women in the developed world.

Unfortunately, 076 is not practical for developing countries. Women in these countries usually do not seek prenatal care as early as the regimen envisioned. Many of these countries lack intravenous equipment used to deliver AZT perinatally. The drug-related cost of the 076 regimen, and of the requisite health care delivery systems, is too high for most developing countries. Moreover, 076 required that women not breast-feed their children, since breast-feeding increases the likelihood of HIV transmission. This is no small matter in countries where polluted water may make formula dangerous and the increased immunity provided by breast milk is especially valuable.

These two facts—the dramatic success of 076 and its impracticability for developing countries—led researchers to the following question: would a different regimen of antiretroviral drugs be effective and useful in the developing world? Researchers created several different shorter courses of treatment that did not depend on the use of IVs and did not require that women refrain from breast-feeding. Next, they needed to design studies to evaluate these courses of treatment. With the exception of one Harvard-NIH study in Ethiopia, all the original NIH- and CDC-sponsored research tested a particular short course regimen against a placebo. The controversy turned on whether this research design was ethically permissible.

The Equipoise Standard

According to the generally accepted standard, clinical research on human subjects is only permissible when scientists don't know what therapy is best. The state of medical knowledge must rest in "equipoise." Placebo-controlled trials are permitted, therefore, only if scientists don't know whether the new therapy is better than nothing. Actually the standard is slightly more complex, in that side effects and other harms are considered along with anticipated benefits. But the

basic principle remains. A placebo-controlled trial is permissible when it is unclear whether a new therapy is better, counting burdens as well as benefits, than no treatment.

The equipoise standard offers a resolution of the tension between individual and communal interests inherent in all clinical trials. On its face, the standard seems to rank the interest of the individual in treatment above the interest of the community in knowledge. Only when the patient has nothing personally to lose by getting one therapy rather than another are researchers permitted to use the patient to advance the common good.

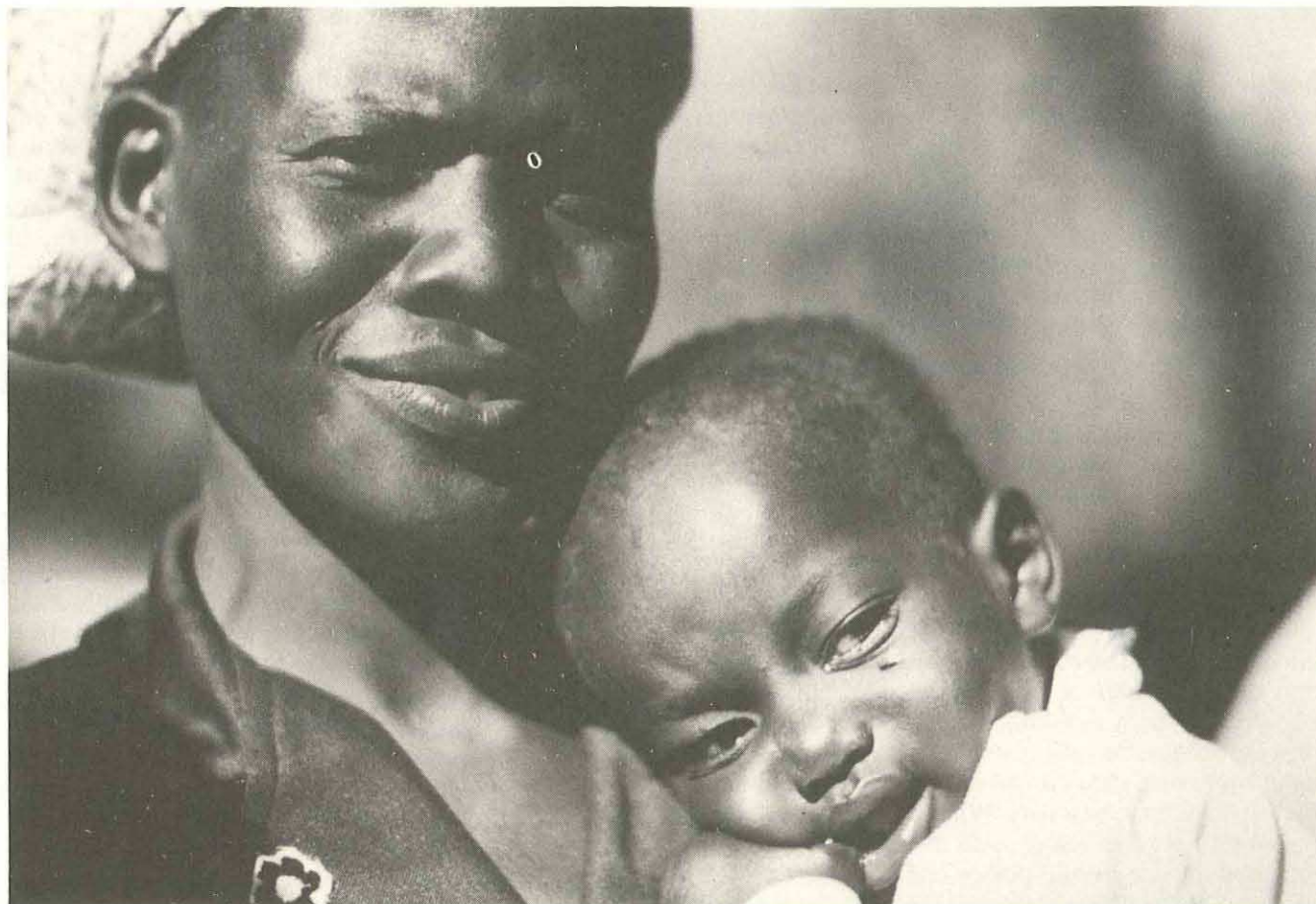
However, this preference for the individual over the community is only apparent. In practice, communal interests trump individual interests. The equipoise standard requires that no patient in a trial get a treatment *known* to be inferior. If we don't *know* that a new therapy is better than the old until we have tested it in a rigorous clinical trial, then patient interest is not sacrificed by participating in such a trial. Until the results are in, we have equipoise. But this argument neglects the patient's perspective. For the patient, well-grounded *belief* is also valuable. If a new therapy is probably better than the old, the patient's interests are sacrificed in a trial where there is a chance she won't receive it.

So, while the articulated standards place the interests of the individual above those of the community, this commitment is eviscerated by the manner in which the equipoise standard is actually employed. The level of certainty required to unsettle equipoise is key. Today,

A placebo-controlled trial is permissible when it is unclear whether a new therapy is better, counting burdens as well as benefits, than no treatment.

the research scientist's standard of knowledge prevails. For the scientist, a hypothesis is proved on the basis of a randomized clinical trial. Moreover, the probability that the results of the trial are due to chance must be less than one in twenty. These are stringent criteria. By importing this certainty standard into the ethicist's understanding of equipoise, we undermine it. Since the patient wants whichever therapy is *believed* best, she sacrifices in joining a trial to establish, for others, that the therapy really is best.

As Samuel Hellman and I have argued elsewhere, clinical research is often ethically problematic because of a conflation of roles. In doing clinical research, the doctor acts simultaneously as a physician and as a scientist. Each of these roles rightly requires a different degree of certainty before action can be taken. In the AZT trials,



Kay Chernush
Archival photograph/World Bank

there is yet a third role in the picture, that of the public health official. The ethical quagmire of these trials is the result of the conflict between these three roles.

Medical researchers rightly aim at knowledge in itself and therefore correctly employ very stringent criteria. Before closing a question and moving on, the

*The equipoise standard must
incorporate the physician's understanding
of when information is valuable
to the patient.*

research scientist must be very sure to have the right answer. Public health officials aim at communal well-being. Because they are under time pressure that the scientist is not, a public health official may employ a slightly less rigorous standard for considering a hypothesis proved. But as public health decisions affect large numbers of people and require the expenditure of often very limited societal resources, communal interest demands a still quite rigorous standard for

medical knowledge. The physician treating an individual patient works toward the health of that patient. As a result, the degree of certainty a physician requires before acting is considerably lower. From the perspective of patient health, the doctor ought to provide that treatment which is most likely to be best.

The articulated standards for ethical research on human subjects command that the role of the physician must predominate. For example, the World Health Organization's Declaration of Helsinki, which is widely recognized as authoritative, provides that "the interests of the subject must always prevail over the interests of science and society." However, this public commitment to put the interests of the individual over those of the community is but empty rhetoric unless researchers interpret the equipoise standard in a way that supports it. For the interests of the study subject to trump communal concerns, the equipoise standard must incorporate the physician's understanding of when information is valuable to the patient.

Ranking individual over communal well-being will have its costs. For example, to test the efficacy of short course AZT therapies, researchers must adopt a study

design without a placebo control. Critics of the trials suggested testing short courses against the 076 regimen. But this design may not be helpful. If the short course therapy proved less effective than 076, as is likely, investigators would still not know whether the short course was more effective than no treatment at all. Alternatively, one could give all study participants short course therapy. Researchers would then evaluate how the results compare with background data on

*Ranking individual over
communal well-being
will have its costs.*

transmission rates from the particular country. The information gathered in this way would be useful and informative, but less certain and reliable than information attained with a placebo control. Differences between the study population and the general population would not be controlled for, and might be relevant. Moreover, data on general rates of transmission within a country may vary. What this means in practice is that studies using such data may have to be repeated before public policy can be based on the results. This lost time may well cost lives as we wait for more certainty before taking action.

Were These Trials Ethical?

The controversial NIH and CDC trials randomly assigned HIV-infected women to either a short course of AZT or to a placebo. To assess whether these trials were ethical, we need to answer two questions. First, how strong were the reasons, pre-trial, to believe that the short course therapies would be more effective than no therapy at all? Second, was that belief strong enough to disturb equipoise?

Prior to the trials, scientists knew that the 076 regimen had been extremely successful in the U.S. and France. They also knew that some participants in the 076 study hadn't received the full regimen but showed some benefit nonetheless. Still, a short course administered in developing countries might be less effective. Oral intake of the drug might not work as well as IVs. Breast-feeding might destroy the gain of treatment. And, finally, other background health and nutrition problems might make AZT simply less effective in the developing world than it proved to be in Western countries. In part, the debate about these trials is over the significance of these differences.

But that is only a part of the disagreement. Given that researchers have *reason to believe* that the short course will decrease the likelihood of transmission—as

they must, or there would be no incentive to conduct the trials in the first place—what follows? For the study's defenders, the fact that the therapy shows promise provides a reason to do the trial, but no more. For the critics, the belief that the therapy will work means that randomization to placebo is unethical. In this disagreement we see the two issues (familiar from introductory philosophy classes everywhere) that I have been discussing.

First, critics and defenders disagree about the appropriate balance between individual interest and communal well-being. Peter Lurie and Sidney Wolfe of Public Citizen, joined by Marcia Angell of the *New England Journal*, condemned the use of babies born to women in the control group for the benefit of babies to be born tomorrow. Study defenders Harold Varmus (NIH director) and David Satcher (then CDC director, recently confirmed as Surgeon General) emphasized the enormous devastation wrought by HIV in these countries and the great need for a decisive answer to an important research question.

The second familiar philosophical issue is epistemological. Critics and defenders of the trials disagree about the value of information not yet validated by a randomized trial. Research scientists, steeped in the values of science, rightly assert that before a trial we

*Critics and defenders disagree
about the value of information
not yet validated by a randomized trial.*

don't know whether a new therapy is valuable. Patient advocates assert that we don't need to *know*. Of course, the scientist doesn't really know that the trial-validated conclusion is accurate, either. The standard of a clinical trial (less than a one-in-twenty probability of a chance result) itself reflects a judgment about the usefulness of this information and the costs of demanding more. Since scientific information is produced by induction, we deal always in degrees of certainty.

The real question that these and other trials raise is: How certain must the physician-scientist be that a new therapy is valuable before a randomized trial becomes unethical? There is a continuum stretching from a doctor's hunch to reasonable belief to solid conviction. Instead of asking when we know that the new therapy is better than the old, we should ask how much certainty it makes sense to demand, given the intended purpose. In practice, even the defenders of the trials recognize this. On the basis of the Thai study, the CDC is stopping the placebo arm of a trial in the Ivory Coast, and NIH is stopping the placebo arms of all of its studies. But there is an important difference

between the women enrolled in the CDC-Thai study and in these others: the Thai women did not breast-feed. So, should we say that we still don't know whether short course will be effective in other countries where breast-feeding is the norm? In stopping the other trials on the basis of the Thai study, the CDC and NIH demonstrate that they too recognize that we make a moral judgment when we decide whether randomization to placebo may continue. As the level of our

*We make a moral judgment when
we decide whether randomization
to placebo may continue.*

confidence in the effectiveness of short course increases, the interests of the babies in that therapy become more insistent, and thus it becomes harder and harder to sacrifice their welfare for the welfare of other babies born tomorrow.

Thus, the two philosophical questions are intertwined. If our aim is community welfare, we need a fairly high level of certainty before taking action. The actor must have a solid conviction that short course therapy will reduce transmission rates significantly in order to justify spending scarce resources and treating thousands of people. But if our purpose is to treat the individual patient, a well-grounded belief is enough. Each HIV-infected woman wants the best chance for her baby to be HIV-free. Research results prior to the controversial trials surely gave her and her doctor good reason to believe that short course therapy would be of value.

In order to fulfill our commitment, as stated in the Helsinki Declaration, to placing individual interests over communal well-being, the physician's certainty standard must prevail. If a doctor would recommend short course therapy over no treatment, then randomization to placebo is unethical.

The Clash of Worlds

These trials also raise a special problem, however. Their aim is not the usual aim of clinical research: a medically superior therapy. Instead, their goal is to find an effective therapy practical for general use in countries where health resources are extremely limited. The equipoise standard is built to handle a different kind of case. In general, researchers look for new therapies that will produce better results: more health, fewer side effects, less pain. In that context, one can only test such a therapy in a randomized trial if the physician does not have good reason to believe that one therapy will offer more to the individual patient.

But in the trials under discussion, scientists are looking for a regimen that will work for a specific population. They recognize that such a therapy might not work as well as 076 works in a different population. Given this research goal, equipoise (in its usual sense) is impossible. The researcher simply cannot say she does not have good reasons to believe one treatment (here 076) is superior. The best known therapy for the individual, however, is not practical for general use.

This same paradox explains the debate over the ethical requirement that all study subjects "be assured of the best proven diagnostic and therapeutic method," as the Helsinki Declaration requires. The trials have been criticized on the grounds that use of placebo denied participants the "standard of care" established for HIV-infected pregnant women. But this argument may prove too much. If the principle articulated in the Helsinki Declaration requires that all study participants get the standard of care in the developed world, then trials of short course therapy against a control

*The commitment to individual well-being
does not require that researchers
give each woman the best therapy
available anywhere in the world.*

arm getting 076 are also unethical. While short course therapies may be highly effective, there are still good reasons to believe they may be less effective than 076. Therefore, participants getting short course would get less than the standard of care.

An ethical requirement that each study participant get as good as is offered in the developed countries would make it impossible to test new therapies in any manner. Each participant would be required to receive 076. This counterintuitive result should make us reconsider the argument that produced it. Where the research goal is to find a practical therapy for general use that may be less effective than treatment available elsewhere, the usual interpretations of both the equipoise principle and the standard of care principle are inapt. Surely we do not wish rigid adherence to these principles to entail that no testing of the much-needed short course therapies is permissible at all. This result would serve neither individual nor communal interests.

Rightly understood, these standards require researchers to put the individual patient's interest ahead of the community's need for medical information. In the special case where adherence to the letter of the standards would frustrate the interests of both the individual and the group, we ought to adhere to the principle from which these standards emanate, rather than to the standards themselves. We can suspend the standards, but not the core principle. Here that principle

would clearly prohibit randomization to placebo. The commitment to individual well-being entails that the investigator treat the research subject simultaneously as a patient. In practice, such a standard does not require that researchers give each woman the best therapy available anywhere in the world. But it does require that researchers studying transmission rates also treat the patient (woman and infant) in a manner designed to reduce transmission.

—Deborah Hellman

Deborah Hellman is assistant professor at the University of Maryland School of Law. Sources: Marcia Angell, "The Ethics of Clinical Research in the Third World," *New England Journal of Medicine*, vol. 337, no. 12 (September 18, 1997); Peter Lurie and Sidney N. Wolfe, "Unethical Trials of Interventions to Reduce Perinatal Transmission of the Human Immunodeficiency Virus in Developing Countries," *New England Journal of Medicine*, vol. 337, no. 12 (September 18, 1997); Harold Varmus and David Satcher, "Ethical Complexities of Conducting Research in Developing Countries," *New England Journal of Medicine*, vol. 337, no. 14 (October 2, 1997); Danstan Bagenda and Philippa Musoke-Mudido, "We're Trying to Help Our Sickest People, Not Exploit Them," *Washington Post* (September 28, 1997); Samuel Hellman and Deborah S. Hellman, "Of Mice but Not Men: Problems of the Randomized Clinical Trial," *New England Journal of Medicine*, vol. 324, no. 22 (May 30, 1991).

Biology, Consciousness, and the Definition of Death

When does a human life end? This question used to be answered quite easily. According to the traditional standard, which has only recently been questioned, a human being is dead when her heart and lungs have irreversibly ceased to function. In some cases, permanent loss of consciousness may precede cardiopulmonary failure. But the interval between these two events has typically been a matter of hours or days, and the traditional standard regards only the latter event as definitive.

Today, however, the development of mechanical respirators, electronic pacemakers, and other medical technologies has created the possibility of a greater temporal separation between various system failures—a patient may lose consciousness a decade or more before his heart and lungs fail, for example. Meanwhile, interest in the availability of transplantable organs has provided an incentive not to delay unnecessarily in determining that a person has died. (Current law, it need hardly be said, embraces the so-called "dead-donor rule": organs necessary for life may not be procured before donors are dead, since the removal of such organs would otherwise cause death—that is, kill the donors—violating laws against homicide.)

Two landmark reports helped to generate a movement away from exclusive reliance on the traditional

standard: the 1968 report of the Harvard Medical School Ad Hoc Committee and a 1981 presidential commission report, *Defining Death*. This second document included what became the Uniform Determination of Death Act (UDDA). Today all fifty states and the District of Columbia follow the UDDA in recognizing *whole-brain death*—irreversible cessation of all functions of the entire brain—as a legal standard of death. The UDDA doesn't jettison the cardiopulmonary standard, however. Instead, it holds that death occurs whenever *either* standard (whichever applies first) is met. One important consequence of this change is that an individual can be legally dead even if her cardiopulmonary system continues to function. If a patient's entire brain is nonfunctioning, so that breathing and heartbeat are maintained only by artificial life-supports, that patient meets the whole-brain standard of death.

Some philosophers and scientists have argued that the whole-brain standard does not go far enough. Several leading authors on the subject have advocated a *higher-brain* standard, according to which death is the irreversible cessation of the capacity for consciousness. This standard is often met prior to whole-brain death, which includes death of the brainstem—that part of the brain which allows spontaneous respiration and heartbeat but is insufficient for consciousness. Thus, a