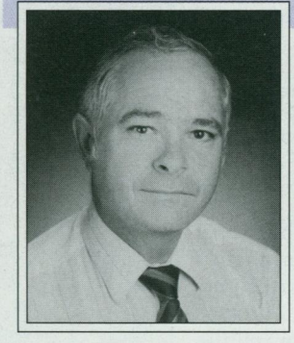
Searching for Markers on the AIDS Trail

The urgent need for new AIDS therapies has caused a head-on collision between compassion and rigorous science; smack in the middle are "surrogate markers"

Imagine a disease, invariably fatal but so slow in developing that it could take a decade to carry out clinical trials that would show conclusively which drugs were effective in preventing death. What should researchers do? Should they insist on scientific accuracy, designing long-term trials (or enormous shorter ones) and withholding unproven drugs from patients on the grounds that untested therapies could do more harm than good? Or should they heed compassion and release drugs as soon as they show any hint of effectiveness, running the risk that—in the absence of carefully controlled trials—it may never be possible to tell which drugs are actually the most effective?

"For the sake of people's lives, I'm willing to use surrogates."

–Robert Biggar



That may sound like an impossibly tortured problem from an ethics seminar for first year medical students, but it's a puzzle that's far from theoretical. In fact, the entire AIDS research and policy community is facing this dilemma as the imperatives of science and mercy collide in case after case. The drug AZT, for example, escaped the most rigorous of scientific tests because early controlled trials were stopped for compassionate reasons after the first sign that the drug was working. When that happened, the control group immediately received treatment, and as a result no one really knows whether AZT extends life. To complicate matters further, two other anti-AIDS drugs, ddI and ddC, have since been released on an accelerated basis, with even less data, using what researchers call "surrogate markers" to evaluate the drugs' impact on the progression of the disease.

Indeed, the AIDS research and policymaking community has seized on "surrogate markers" as the best means of reconciling compassion with rigorous science. Surrogate markers are biological indicators that—researchers hope—will reflect the step-by-step progression from symptomless HIV infection to fullblown AIDS. If the markers are good, they can be exploited as substitutes for clinical endpoints such as AIDS symptoms or death. But there's a catch: Nobody knows what the best surrogate markers are for AIDS progression. And that's created tangles, as the Food and Drug Administration (FDA) attempts to get drugs to people suffering from HIV, and the scientific community struggles to understand how the disease progresses and decide what the best markers are.

In spite of these uncertainties, the FDA has decided it has to move quickly. In the next few weeks, the agency will adopt a regulation making it easier to approve AIDS drugs rapidly based only on data from surrogate markers. The marker it has relied on most heavily to date is the one that's received the most scientific attention: the number of CD4 cells, key immune system cells that HIV infects and destroys. But CD4 counts have serious limitations when it comes to predicting a treatment's clinical efficacy, and researchers are currently examining more than a dozen other markers (see box on next page). FDA Commissioner David Kessler acknowledges that in the arena of surrogate markers, policy has begun to outrace scientific knowledge: "We're certainly pushing the limits of the envelope here," he says. "We're approving [treatments] on less data than ever before." But, given the urgent need for better AIDS treatments, he adds, "we're trying to do this as thoughtfully as possible."

Some researchers, however, think the push to use surrogate markers to approve AIDS therapies is premature and could ultimately be harmful. Deborah Cotton, an infectious disease specialist at the Harvard School of Public Health, is a member of the FDA advisory panel that recommended approval of ddI and ddC on the basis of surrogate markers (and the fact that the drugs operate by a similar mechanism to the already approved AZT). Cotton voted against both approvals and says she felt the panel was asked to "pound [the data] into a scientific conclusion."

"We really have to ask whether relying on surrogate markers will hasten a cure or hinder it," says Cotton. "We're getting into a situation of such complexity that we may have a large number of agents being used and no way of distinguishing among them." She's also concerned that effective drugs might even be lost in the surrogate marker shuffle—wrongly rejected because they don't influence the surrogate marker that has been crowned as the preferred index.

Like all other experts interviewed by *Science*, Cotton firmly acknowledges the need for better drugs and she believes that ultimately surrogate markers will be the way to find them; but she doesn't think the markers on hand are good enough. "We've got to keep looking for surrogate markers," she says. "Right now, CD4 isn't a very good marker." At bottom, she says, the problem is a very human propensity: the desire for pat solutions. "I don't think there's an easy answer, and that's what people want."

How CD4 became THE marker

The problem of finding surrogate markers for AIDS progression has been kicking around the research community since the first successful anti-AIDS drug—AZT—was tested. The first patients began receiving AZT in a small safety study in 1985. It quickly became clear that AZT was not the hoped-for magic bullet against AIDS, but it did seem to be helping: After 6 weeks, most of the 35 people in the trial had gained weight and shown increases in CD4 counts. Those hopeful indicators led to a larger, placebo-controlled trial in 281 patients, which got under way the following February. Seven months later, the plug was pulled on the larger study because only one person receiving AZT had died, compared to 19 deaths in the placebo group. In March 1987, the FDA approved AZT for people with advanced HIV disease.

That wasn't the end of the AZT story, and it was only the beginning of efforts to find reliable surrogate markers, because the trial was interrupted before enough deaths had occurred to be sure that the drug actually lengthened life. Subsequent trials showed AZT does reduce the risk of opportunistic infections in healthier people, but those trials did not demonstrate that the drug prolongs life in healthy patients; furthermore, improvements in CD4 counts were small and temporary.

Nevertheless, CD4 counts became the benchmark of clinical trials. The reason was that a decrease in CD4 numbers was a clear cellular sign of AIDS, and it seemed likely that any drug that increased the count was

Exploring Other Surrogate Markers

Though CD4 counts are the best-known surrogate marker for AIDS, the body offers other clues about the rate at which HIV is dismantling the immune system, clues that may help researchers identify effective treatments. Here is a sampling of lesser-known surrogate markers that researchers are exploring.

■ **β_2 microglobulin and neopterin.** β_2 microglobulin is part of the class I major histocompatibility complex found on the surfaces of many cells; neopterin is produced when white blood cells called macrophages are stimulated. Both are elevated in HIV-infected people and in some studies have been better disease predictors than CD4 counts. Though they offer similar predictive value, β_2 microglobulin is more informative in late stages of the disease. A few AZT studies have shown that both can help evaluate therapies. Yet researchers are uneasy because neither seems directly implicated in the disease process. As Anthony Fauci, head of the National Institutes of Allergy and Infectious Diseases (NIAID), says: "People don't die from elevated neopterin or β_2 microglobulin levels."

■ **Interferon.** Interferon, a protein that carries signals to white blood cells, is elevated during HIV disease, especially in sicker patients. In the *Lancet* earlier this year, Donna Mildvan of New York's Beth Israel Medical Center and her colleagues reported a small study in which AZT markedly reduced levels of two types of interferon, an effect associated with decreased risk of death. "If confirmed in larger studies, it could be a very powerful marker in advanced disease," says Mildvan.

■ **Delayed-Type Hypersensitivity (DTH).** As their disease progresses, HIV-infected people lose the ability to respond to foreign antigens. DTH tests capitalize on this by injecting harmless antigens into a person and then measuring the resulting sore. DTH is a gauge of a type of immunity known as cell-mediated immunity (as distinct from the antibody response).

■ **Early clinical symptoms.** Symptoms like thrush, herpes zoster, weight loss, and rashes do not define AIDS but can serve as useful markers of a treatment's effectiveness.

It's also more than possible that the best markers haven't been discovered. "Surrogate markers are the clinical application of pathogenesis research," says Jonathan Kagan, head of clinical sciences at NIAID's Division of AIDS. "The more we understand pathogenesis, the more we'll find markers."

And, as many researchers emphasize, markers become more powerful as you combine them. "Anything as complicated as HIV can't be totally captured in one silly little marker," says biostatistician Stephen Lagakos of Harvard University. "That's naive. It's going to take a battery."

–J.C.

likely to help HIV-infected people. In October 1991, the FDA approved ddI based on CD4 counts, with the manufacturer promising to continue evaluating clinical endpoints and to take the drug off the market if it turned out not to work. To the relief of many associated with ddI's approval, postmarketing data presented to the FDA's antiviral advisory committee in April, 1992, suggested that ddI does reduce the number of AIDS-related opportunistic infections. At the same April meeting where the ddI data were presented, slight improvements in CD4 counts led the FDA's advisory committee to recommend approving ddC (but only for use in combination with AZT).

"Anything as complicated as HIV can't be totally captured in one silly little marker."

–Stephen Lagakos

These decisions left many observers wondering whether too much weight was being placed on CD4 counts alone. Activist Mark Harrington of New York's Treatment Action Group, who serves as a "community representative" on the FDA's advisory committee, maintains there's been an "overreliance" on CD4 counts by the researchers who make approval decisions. "We're not just using the marker to evaluate the drugs, but we're also using the drugs to evaluate the marker," Harrington says, explaining that the shaky faith in CD4 as a surrogate marker seemed to be confirmed when ddI appeared to bring improvements in the clinical condition of those who received it. But is Harrington, who isn't a researcher, barking up the wrong tree in his reservations about CD4 as a marker? As interviews by *Science* show, many qualified scientists agree that CD4 alone isn't the answer.

Vague shapes, cloudy rooms

The main scientific difficulty underlying the entire surrogate marker dilemma is that AIDS researchers are still trying to decipher how HIV undermines the immune system; and they don't fully understand the roles surrogate markers play in that process. And those two big areas of ignorance compound the problem of evaluating clinical trials. "It's like trying to see a vague shape across a cloudy room," says Dan Hoth, head of the Division of AIDS at the National Institute of Allergy and Infectious Diseases (NIAID). "There are two sources of variability that make it hard to see the truth."

In spite of these uncertainties, researchers know what they would like to have in a surrogate marker. A "good" surrogate marker, investigators say, is one that has a biologically plausible connection to the disease process. CD4 counts certainly meet that criterion, since the destruction of CD4 cells is thought to be central to the weakening of the immune system that characterizes AIDS. In untreated HIV-infected people, the number of CD4 cells declines steadily over time. "HIV is a disease about CD4 deficiency," says Hoth. "It's clearly bad to have low CD4s."

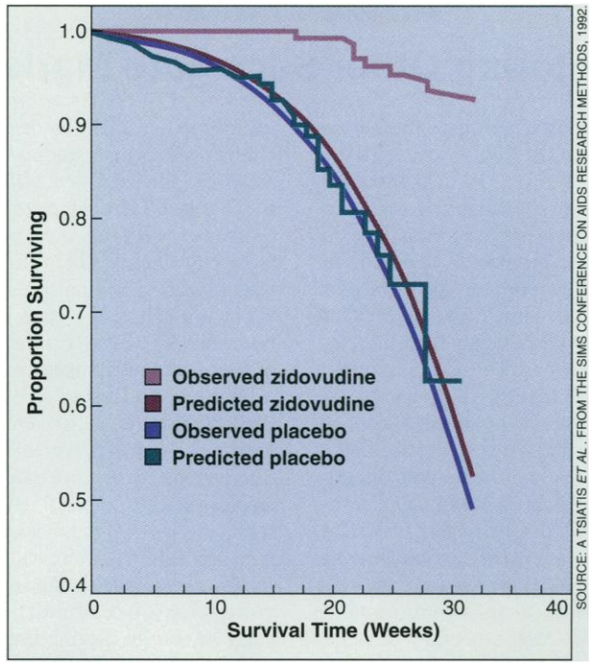
But CD4 counts, as even Hoth agrees, have shortcomings as a way to evaluate therapies. It isn't sufficient for a surrogate marker to be plausibly implicated in the disease. A good marker should also improve quickly with an effective treatment—and the change in the marker should be directly correlated with improvements in the patients' clinical conditions. And that's where CD4 falls short. Biostatisticians at the Harvard School of Public Health have examined data from the AZT trials and calculated that improvements in CD4 counts are responsible for at most one-fourth of the benefits seen in the AZT-treated patients. "Treatment-mediated changes in the [CD4] marker do not explain much of the treatment effect," says Victor Degruttola, a biostatistician at the Harvard School of Public Health.

What is more, CD4 counts are far from being ironclad predictors of progression to AIDS. HIV-infected people with very low CD4 counts sometimes remain healthy; conversely, ailing people sometimes have comparatively high levels of CD4s. In addition, CD4 counts are notoriously

fickle: They can vary widely between labs or because of a person's age, the time of day a measurement is taken, and even whether the person smokes.

Despite these limitations, many researchers feel that the approvals of AZT, ddI, and ddC were appropriate and that use of CD4 as a marker has helped lead to sound decisions. "We're asking a lot of the CD4 marker at this point," says the University of Miami's Margaret Fischl, who has been involved with several groundbreaking AZT clinical trials. "CD4 is probably a very good surrogate marker," Fischl adds, but the anti-HIV drugs tested so far may not be potent enough to consistently send CD4 levels flying up.

Other researchers agree with Fischl that, no matter what defects surrogate markers may have, they're a necessity. National Cancer Institute (NCI) epidemiologist Robert Biggar, who has analyzed the performance of several surrogate markers in studies of the natural history of AIDS progression, says he feels the "tremendous need for drugs" makes the desire for definitive studies based on clinical endpoints a "luxury." "For the sake of people's lives, I'm willing to use surrogates," Biggar says.



Small change. HIV-infected people who received AZT (zidovudine) survived far longer than would have been predicted solely by using the small and transient changes seen in their CD4 counts—showing the difficulties of using CD4 as the sole surrogate marker for evaluating AIDS therapies.

Can you put the genie back?

It is that kind of awareness of lives hanging in the balance that led the FDA to its accelerated approval program based on surrogate markers, which Kessler expects will clear its final bureaucratic hurdles in a few weeks. Accelerated approval is not limited to AIDS, but it applies only when the disease is "serious or life-threatening" and "a serious medical need is not met by currently available therapies." After the FDA grants an accelerated license, the drug's sponsor must conduct postmarketing studies to prove actual clinical efficacy. If the drug fails those tests, the FDA can yank it off the market. "The key to this is the follow through," says Kessler, or what he calls "putting the genie back in the bottle."

To get the genie back in the bottle, the accelerated approval program pumps up the legal muscle available to the agency to remove a drug from the market if it doesn't show clinical benefits. That's one important departure from previous practice. Another key change is directly relevant to surrogate markers: The new regulations downgrade the standard for surrogate marker data from being "very" to being "reasonably" likely to predict clinical benefit.

But many AIDS investigators interviewed by *Science* argue that, in the future, CD4 counts alone are unlikely to meet even this weakened standard. In particular, they argue, another marker—the total amount of HIV present in a person's system at a given moment—is likely to complement, if not supplant, CD4 counts as a way of predicting the progress of infection. (This indicator is often referred to as "viral load.") David Ho of New York's Aaron Diamond AIDS Research Center argues that "one should not call viral load a surrogate marker. It is *the* marker you want if you believe AIDS is a viral disease."

Robert Coombs of the University of Washington agrees with Ho: "As people look harder and harder at CD4, it will become less important as a marker," says Coombs. "From my point of view as a virologist, the goal here is to turn the virus off totally."

It's not surprising that Coombs and Ho are in agreement on this subject, since it was their combined work that turned the attention of the research community to viral load. In 1989, they published back-to-back articles in the *New England Journal of Medicine* that made many other researchers believe direct measurement of HIV in the blood was a sensitive and meaningful marker. Before that, the most popular measure of viral load was an assay for the HIV core protein, p24. But p24 levels could be found only in a small percentage of patients and they didn't show levels of infectious virus.

Believing viral load is a sensitive marker is one thing, however, and deciding just how to measure it is another. A variety of high-tech methods are now becoming available to measure the level of virus. One of the newest and most intriguing is a quantitative assay of HIV nucleic acids by polymerase chain reaction (PCR). Though the newly developed quantitative PCR promises to be a powerful tool, Deborah Birx notes that in studies she's conducting at the Walter Reed Army Institute of Research, the changes in viral load between early and late stage patients are clear but of a relatively small magnitude. "[The] change is difficult to assess," says Birx.

Even if such sensitive tests could measure clear differences in viral load as HIV infection progresses, they might not be definitive, because all of them use peripheral blood as their raw material. And, as NIAID director Anthony Fauci and his colleagues have shown, in the early stages of HIV disease, there is little virus in the blood but large amounts sequestered in the lymph nodes. A new NIAID-sponsored trial and another at Walter Reed are now biopsying lymph nodes from infected people receiving different treatments to compare changes in lymph and blood viral loads.

All these complexities suggest that it won't be a snap to use viral load as a surrogate marker. Yet it does seem that, for the moment, viral load, in combination with CD4 counts, is becoming the marker of choice for clinical trials—especially for preparations that work by mechanisms different from that of AZT. Last month, at NIAID's annual AIDS vaccine conference, Dennis Klinman of the FDA, said that "if you have to go with surrogate markers," for therapeutic vaccine trials, he would favor a combination of viral load and CD4. Though Klinman stressed that he was not speaking officially, it's clear that his informal remarks reflect the latest thinking in AIDS clinical trials.

That kind of thinking continues to draw criticism from those who, like Deborah Cotton, view it as short-sighted, putting an immediate form of compassion first and, in the end, depriving the very patients it's designed to help. "It's sad that we may have nothing to offer people in 1992," she says. "It's sadder that in 2000 we may have nothing, too. In 2000 we'll look back and say, 'If only we'd done this in a more rational way.' "

But that kind of hard-headed thinking is difficult to maintain in the face of thousands of people who are getter sicker and dying. Cotton is likely to remain a skeptical voice of the minority. The majority is more likely to heed those like Robert Biggar of NCI, who thinks the risk of using surrogate markers is well worth it. "If someone comes to me and says, 'Look, you were wrong [to rely on surrogate data],' I'd say, 'Well, we did the best we could.' " The sad fact, says Biggar is that "there isn't much of an alternative." And, in the absence of alternatives, the reliance on surrogate markers, flawed though they may be, will surely continue.

–Jon Cohen