AIDS Trials Ethics Questioned

The advent of potent anti-HIV therapies has thrown the world of AIDS drug testing into turmoil, with some researchers lambasting drug developers for offering trial participants "suboptimal" treatments

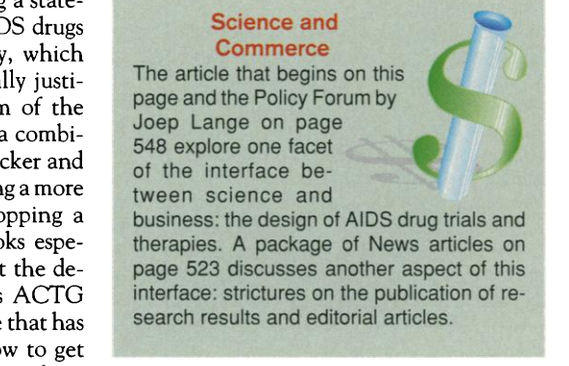
Early this year, researchers running a state-of-the-art, clinical trial of anti-AIDS drugs abruptly called a halt to the study, which they felt could no longer be ethically justified. Patients in the "control" arm of the trial, who were being treated with a combination of two drugs, were getting sicker and dying faster than were those receiving a more powerful, triple-drug cocktail. Stopping a trial when a candidate therapy looks especially promising is not unusual. But the demise of this large study, known as ACTG 320, added fresh urgency to a debate that has been simmering for years about how to get useful results from drug tests without risking patients' lives. It also highlighted how the recent advent of potent combination therapies has dramatically changed the ground rules for AIDS drug testing.

From the start of the AIDS epidemic, researchers have argued over how to test new therapies without putting patients in harm's way. The standard approach is to test candidate drugs against a placebo or a benchmark therapy, using a clinical measure, such as death rates, to evaluate effectiveness. When the first AIDS drugs went into trials more than a decade ago, researchers and AIDS activists agonized over whether it was appropriate, given the lethal nature of the disease, to test potential drugs against placebos. As more candidate therapies came along, they debated the propriety of using clinical measures such as mortality rather than more indirect, "surrogate" measures such as blood levels of critical immune-system cells, to judge effectiveness. The recent development of combination-drug treatments that reduce HIV in the blood to undetectable levels has only sharpened these issues. The central question now facing AIDS researchers and drug testers is whether it is unethical to run a trial that offers patients anything less than these new therapies. Trials in developing countries, where far fewer AIDS drugs are available, are being criticized, too (see p. 519).

AIDS drug developers well recognize that many of yesterday's state-of-the-art clinical trials have become outdated. "We have changed our thoughts dramatically over the last year," says Ann Collier, an AIDS clinician at the University of Washington, Seattle. Collier helps oversee studies staged by the AIDS Clinical Trials Group (ACTG), a network of academic sites funded by the National Institutes of Health, which has been reviewing all of its protocols during the past few weeks. "It is very important that the organization do the most ethical thing and do what is most appropriate *now*," says Collier. Drug companies are furiously revamping trials, too. "We are basically reexamining our whole program," says Lynn Smiley, head of antiviral research at Glaxo Wellcome, maker of the anti-HIV best sellers AZT and 3TC.

Science and Commerce

The article that begins on this page and the Policy Forum by Joep Lange on page 548 explore one facet of the interface between science and business: the design of AIDS drug trials and therapies. A package of News articles on page 523 discusses another aspect of this interface: strictures on the publication of research results and editorial articles.



ANN COLLIER/UNIV. OF WASHINGTON

Swallow this. Complicated new combination treatments, as shown in this patient's pill tray, has complicated the design of clinical trials.

But some AIDS researchers say the pace of change is too slow. One of the most vocal critics has been Joep Lange, a clinician and virologist at the University of Amsterdam in the Netherlands who, last January at the biggest annual AIDS meeting in the United States, lambasted drug developers for continuing to conduct trials with what he describes as "suboptimal" treatments. In a Policy Forum on page 548 of this issue, Lange presents a more measured version of the January talk.

But there is a hot debate about what constitutes a suboptimal trial. To Lange and like-minded researchers, the line between optimal and suboptimal is clear: "Anything that is not designed to completely suppress viral replication is suboptimal," asserts virologist Douglas Richman, an expert on HIV drug resistance at the University of California, San Diego (UCSD). Like Lange, Richman further argues that studies of AIDS, tuberculosis, and other infectious diseases have proven that combining agents delays drug resistance. Putting off resistance is crucial, he and others argue, because resistance to one drug often entails resistance to similar compounds. "You have to be very careful in taking drugs so that you maximize your options," says AIDS clinician Daniel Kuritzkes, who stages drug trials at the University of Colorado Health Sciences Center in Denver.

Other researchers, however, say that to know how best to maximize patients' options, clinical trials must study a wide variety of treatments, including those that may have only modest firepower against the virus. These investigators note that anti-HIV drugs affect people differently depending on how damaged their immune systems are, which treatments they already have exhausted, and how able or willing they are to swallow daily a dozen or more pills that can come with debilitating side effects. Some AIDS researchers also argue that tests that rely solely on a treatment's effectiveness in suppressing levels of HIV may not provide a reliable indicator of death rates over the long term. "The biology is really screaming at us to get everyone to undetectable levels [of HIV], but there is more to medicine than biology," says Michael Saag, an AIDS clinician at the University of Alabama, Birmingham.

Science has explored these contentious issues with more than 50 leading clinicians, drug developers, regulators, and activists, closely examining a dozen controversial trials. While there are no easy fixes, many creative ideas about ways to improve clinical trials are being floated, from making it easier for studies to incorporate new findings, such as the ACTG 320 data, to lowering the hurdles a drug must clear before it is sold.

Bad news, good news

Four years ago, the catchwords in AIDS drug development were doom and gloom. The only compounds on the market—AZT, ddI, and ddC—were those that interfered with HIV's reverse transcriptase enzyme, preventing the virus from injecting itself into the host's DNA. And all proved to be only moderately effective, even in combination. By the end of last year, however, the pendulum had swung to the other extreme, with talk of treatments leading to Lazarus-like rebounds in many people and even the possibility of a cure.

Two major advances triggered the excitement. One was the development of sensitive tests that can measure minute amounts of virus in the blood. The second was the discovery of the stunning impact on HIV of reverse transcriptase inhibitors (RTIs) combined with drugs that disable HIV's protease enzyme, which helps the virus copy itself after it infects the host DNA. During the past year, a dizzying assortment of combinations have been shown to drive down the amount of HIV in the blood—the "viral load"—to below the detection limits of the new tests—and keep it there. Although people with no detectable HIV in their blood can still harbor plenty of virus in their lymph nodes and other sites, an undetectable viral load has become the new gold standard for HIV treatments and clinical trials.

What happened with the ACTG 320 study exemplifies this sudden shift in thinking. ACTG 320, a comparison of a combination of the RTIs AZT and 3TC with a cocktail consisting of these two drugs plus the Merck protease inhibitor indinavir, was launched in January 1996. Later that month, researchers presented early data from a small pilot study sponsored by Merck comparing the same drugs. After 24 weeks, the triple combination had whacked HIV to undetectable levels in six of seven patients, versus zero of eight people in an AZT-3TC "control" arm. "There were groups who said even back then that they did not want to take [just] AZT-3TC based on those [data]," recalls Martin Hirsch, a virologist at Massachusetts General Hospital. Not only did the triple drug combo appear to be dramatically more powerful, but the higher viral levels in people receiving only two drugs suggested that they had become resistant to the AZT-3TC combination. Researchers already had established that resistance to 3TC developed in a few weeks when the drug was used alone, and in a few months when used with AZT. Still, Hirsch felt the trial was "perfectly ethically reasonable," as did many other influential clinicians.

The early Merck study, Hirsch and others reasoned, involved too few people to answer definitively even the limited biological questions it was asking. The ACTG 320 study, in contrast, was large enough that it stood a good chance of determining whether the triple combo actually delayed disease and prolonged life. Moreover, at the time the study began, most researchers considered AZT-3TC the standard of care, which made it an appropriate benchmark against which to test candidate therapies.

But, as new data came in, Hirsch became seriously concerned about ACTG 320. Late last year, he completed a trial comparing AZT-3TC with and without indinavir in people with fewer than 50 CD4s, immune-system cells that are prime targets of HIV. (Healthy people have 600 to 1200 cells per cubic millimeter of blood.) At 24 weeks, 65% of the people on triple drug therapy had undetectable levels of virus as compared to 0% on AZT-3TC alone. Hirsch contacted the researchers running ACTG 320, who, in turn, asked the independent Data and Safety Monitoring Board (DSMB) overseeing their trial to peek at the data and see whether there was yet a statistically significant difference between the treatment and control arms. There was not.

Hirsch presented the data from his study 1 month later at the same meeting at which Lange gave his blistering talk. Soon thereafter, anxious ACTG investigators asked the members of the DSMB to take another look at ACTG 320. Disease had progressed in 66 people receiving AZT-3TC, 18 of whom died. Only 33 cases of disease progression and eight deaths occurred on the AZT-3TC-indinavir arm. On 18 February, the DSMB recommended pulling the plug.

In the minds of many researchers and AIDS activists, ACTG 320 marked the end of large trials that rely on clinical outcomes such as disease progression and death to evaluate a new treatment. "From this point," says the University of Colorado's Kuritzkes, "I don't think [an ACTG 320] type of study will likely be done again. When there are clear indications that a drug can decrease the viral load and increase CD4 counts," it is not ethical to wait for "body count" data, he says. The trial's results also immediately threw into question every protocol that had an AZT-3TC control arm.

ANTI-HIV DRUGS

	Date Approved by the FDA	Time to Resistance (monotherapy)
Reverse Transcriptase Inhibitors		
AZT	3/87	medium
ddI	10/91	slow
ddC	6/92	slow
d4T	6/94	slow
3TC	10/95	fast
1592U89	E	unknown
bisPom PMEA	E	unknown
Non-Nucleoside Reverse Transcriptase Inhibitors		
Nevirapine	6/96	fast
Delavirdine	4/97	fast
DMP-266	E	fast
Loviride	E	fast
Protease Inhibitors		
Saquinavir	12/95	medium
Ritonavir	3/96	medium
Indinavir	3/96	medium
Nelfinavir	3/97	medium
141W94	E	unknown

E = experimental

Trials and tribulations

Broadly speaking, three trial designs recently have been denounced as suboptimal: those that do not allow participants to take protease inhibitors, and trials that give participants either AZT-3TC or any single drug (the quickest way to develop resistance). But some researchers who defend such protocols raise difficult questions of their own. If a drug has never been tested in humans, can one learn how the body processes it without testing it on its own? If a questionable trial is expected to yield important results soon, should it be allowed to continue? Is the trial short enough that drug-resistant HIV strains will not have time to gain a foothold? Is the study long enough to satisfy the Food and Drug Administration (FDA)?

The National Cancer Institute's (NCI's) Robert Yarchoan, a veteran AIDS-drug developer, gave a talk at a Gordon Conference last month about the difficulty of testing new drugs in the post–triple-combo era. "If you really take a hard line and say everyone should be on triple therapy and drive the virus down to zero and everything else is amoral, then you get into a situation of how do you test new drugs," says Yarchoan. "I don't have the answer."

Answers may be in short supply, but several studies have raised such serious questions that the trials have been scotched or modified during the past few weeks. A protocol proposed by Bristol-Myers Squibb, for example, that called for comparing ddI to ddI-d4T (an RTI) and to ddI-d4T-indinavir never left the runway. At a meeting before the trial was scrapped in late March, Howard

Grossman, a New York City AIDS clinician, and others slammed the design. "I don't think at this point that I could ethically recommend monotherapy to anyone," says Grossman. "I would have a couple of months ago. The world has changed."

Bristol clinical trial executive Laurie Smaldone says resistance to ddI develops slowly and the trial would have addressed vital questions about how the drug worked. "Single-agent therapy for some physicians still has value," says Smaldone. But the company recognizes that it is "appropriate to revise [the protocol]," she adds.

Trials in error? Joep Lange.

Virologist Maureen Myers of Boehringer Ingelheim Pharmaceuticals says her company a few weeks ago made the costly decision to modify a large, ongoing study of its drug nevirapine, which has been on the market since June of last year. The compound is a "non-nucleoside" RTI (NNRTI), a family to which HIV quickly develops resistance. The study, begun before the advent of protease inhibitors, initially compared nevirapine plus other RTIs to various combinations of RTIs. Now, participants in both arms can take protease inhibitors also. "We ended up biting the bullet," says Myers, who predicts the change will make it harder to see a difference between control and treated arms, requiring the company to enroll an additional 500 participants. (To see smaller differences, trials must run longer, enlist more participants, or both.)

Lange has been most critical of an ongoing, 99-person, 64-week trial called Quattro, which is sponsored by the U.K.'s Medical Research Council (MRC). Designed in 1993 and launched in July 1995, Quattro compares a cocktail of AZT, 3TC, ddC, and the NNRTI loviride to AZT-3TC and to the four drugs given sequentially as monotherapy. Lange says AZT-3TC is now a suboptimal arm of any trial, and he argues that the sequential monotherapy arm is even riskier because it may quickly make patients resistant to some of these compounds. "I just think [the involved researchers] are wrong, and they are taking undue risks in the face of biology," said Lange in a recent interview. "It's ruining [patients'] future options."

Janet Darbyshire, head of the MRC's HIV Clinical Trials Centre, stresses that while the trial might have a dated design, it is asking valid questions about the possibility of reducing toxicity by giving drugs sequentially. If the sequential monotherapy reduces the viral load in a stepwise fashion, reasons Darbyshire, over the long term, "the approach may still be as effective as the concurrent approach." She further notes that Quattro's Data and Safety Monitoring Committee reviewed the data in March and saw no reason to stop the trial.

Robin Weiss, a virologist at Chester Beatty Laboratories in London who advises that committee, adds that the trial may help determine whether the best way to achieve sustained, undetectable viral loads is to treat people aggressively up front. He agrees that giving three drugs at once should initially lead to a more dramatic drop in viral load. "But it is pure speculation whether a year later or 2 years later these patients are going to have a lower load than those patients given drugs sequentially," says Weiss. "It's just a rationale, and it's turning into a dogma. Quattro might give [some] answers."

Lange is hardly alone in his criticism of Quattro. And there are several other trials that he and others interviewed by *Science* say contain suboptimal arms, including:

■ *HBY097 (an NNRTI) vs. HBY097-AZT.* "As a general statement, we know enough about NNRTIs to say that monotherapy is a problem," asserts ACTG head Robert Schooley of the University of Colorado Health Sciences Center. Oncologist Dagmar Oette, who oversees the trial for HBY097 developer Hoechst Marion Roussel, argues that the monotherapy arm lasts only 12 weeks and that test-tube studies show that resistance is less of a problem with this drug than it is with its cousins.

■ *ddI vs. hydroxyurea (a cancer drug)-ddI.* A committee of the ACTG, which is running this study, recently considered stopping it because of its monotherapy arm, says the University of Washington's Collier: "But it came to the conclusion that it is very important to finish this relatively short-term study, because there is so much speculation about what this combo can do." Although HIV is slower to develop resistance to ddI than to other RTIs, some researchers say the single-therapy arm still poses an unnecessary risk for trial participants. Schooley says the trial "is probably stopping itself" because people are not enrolling in it.

■ *1592 (an RTI)-AZT-3TC vs. AZT-3TC.* In the post-ACTG 320 environment, Mass General's Hirsch, among others, feels "uncomfortable" with this proposed trial protocol, which enrolls people with CD4 counts of 100 or more. Smiley of Glaxo Wellcome—which makes all three of the drugs—says the protocol still may change. She notes that the design includes a "bailout" clause that lets people alter therapies at 16 weeks if their viral loads are increasing. She also says "the jury is still out" about whether resistance to AZT-3TC will develop in that short a study with people who have never received any of the drugs.

■ *Delavirdine (an NNRTI)-AZT-3TC vs. AZT-3TC.* Schooley, who has strong reservations about any AZT-3TC arm, says patients should be given a choice of adding a protease inhibitor. A spokesperson for Pharmacia & Upjohn, delavirdine's manufacturer, declined to discuss the trial design.

■ *DMP-266 (an NNRTI)-AZT-3TC vs. AZT-3TC.* DuPont Merck, DMP-266's developer, announced on 25 March that it was closing this study of patients who, like those in ACTG 320, had previously taken AZT or other RTIs. But the company is continuing a separate trial with people who have never received HIV therapy (and so have not developed resistance to AZT or other RTIs) and have CD4 counts as low as 50. Kenneth Gorelick, a DuPont Merck vice president of clinical research, explains that ACTG 320 did not evaluate the impact of AZT-3TC in this population. "We believe AZT-3TC is acceptable in that case for the first 16 weeks," says Gorelick.

Future fixes

Lange and other investigators contend that FDA rules are at least partly to blame for the number of suboptimal arms in clinical trials. "As long as drug companies need to show their drug is better than existing therapies, they will think they have to use the worst recommended therapy as a comparison," says NCI's Yarchoan. "It really requires some fundamental rethinking on the FDA's part."

One change championed by drug companies and virologists like Schooley and UCSD's Richman is for the FDA to give greater weight to a drug's ability to drive down viral load. Currently, the FDA conditionally approves anti-HIV drugs based on changes in surrogate markers such as viral load and CD4 counts. But the company is required to follow up with large-scale trials measuring impact on clinical indicators such as disease progression and mortality. The nevirapine trial that Boehringer recently modified is just such a postmarket study.

Indeed, an FDA advisory committee will meet in July to discuss the possibility of granting anti-HIV drugs full approval if they offer a sustained reduction in viral load. But some researchers, such as biostatistician Thomas Fleming of the University of Washington and a former member of the advisory committee that advises the FDA about AIDS drugs, warn that surrogates often fail to predict the clinical outcome of a drug treatment, an argument he detailed in the 1 October *Annals of Internal Medicine*.

But others say Fleming's position sidesteps the reality of AIDS drug development today. "It's allowing people to get sick," says Schooley. "A lot of times people sanitize things. It's more acceptable to say that they're doing things for humanity. It's the licensing gap, and you could drive a truck through it," says Schooley.

Richman thinks opponents of surrogates often fail to appreciate the magnitude of the changes seen in viral load and CD4 levels with today's treatments. "All this talk of HIV and CD4 being surrogate markers, that has always bothered me," says Richman. "They are not surrogate markers. They are the measurement of the disease."

Some researchers have strong faith in the virologic marker, but worry about relying on it too heavily. "At the ACTG, we have a number of very important studies with virologic end points," says the University of Alabama's Saag. "All of them will end in 24 to 48 weeks. What are we going to have when that's all over? Highly active regimens that lead to undetectable virus in 90% of patients. What do we do with that in practice? How do we know strategically when to employ one over the other?"

Saag suggests that it would be more useful to do a clinical end-point trial that mirrors the way people take drugs in real life. "In my ideal scenario, it doesn't matter what regimens they are getting," says Saag. His idea, which he calls "Strategic Timing of Antiretroviral Therapy," or START, is to offer patients a menu of treatment options, and to switch to a different regimen whenever the viral load increases above a certain cutoff—a strategy he acknowledges could lead to one very long trial.

As yet, Saag has won few converts, though. Lange, for instance, counters that it matters greatly how you drive HIV down: As soon as your viral load goes up, you have developed resistance to at least some of the drugs on that regimen, he says. Consequently, he says hitting HIV as hard as possible right away will prolong the use of all of the drugs in a given regimen. Many ACTG leaders share Lange's view, and Saag's START idea was rejected at an ACTG meeting last month. They instead hope to analyze the clinical outcome of all of their trials collectively in a giant matrix.

Harvard biostatistician Victor de Gruttola, who collaborated with Saag on the START idea and is working with colleagues to develop a computer model that can analyze anti-HIV drug options in different patient populations, is disappointed by the ACTG's decision, but allows that the matrix idea might yield useful results. "The whole thing is just in chaos right now, but hopefully it's creative chaos," says de Gruttola.

–Jon Cohen

SCIENCE AND COMMERCE

Publishing Sensitive Data: Who Calls the Shots?

A rash of events in the past few days has thrown a spotlight on the tensions that can arise between science's tradition of open publication and industry's penchant for secrecy. The following articles detail two cases of alleged suppression of unfavorable research findings, a survey indicating that a substantial fraction of researchers in the life sciences have delayed publication or withheld results and materials from colleagues, and a dispute among the top medical journals over rules to guard against conflict of interest in medical publications. But a story on page 527 suggests that not all university-industry interactions are so fraught with problems.

Secrecy Dispute Pits Brown Researcher Against Company

Faculty members at Brown University are in an uproar over what appears to be a classic industry-academic research conflict. At the center of the furor is David Kern, an occupational health physician who claims that his research on an outbreak of lung disease at a local textile plant is being suppressed by a Brown-affiliated hospital and the plant's owner—and that his clinic was closed in retaliation.

Kern, who is employed by the Memorial Hospital of Rhode Island in Pawtucket and is an associate professor at Brown's School of Medicine, conducted the research as a consultant to the textile company. (The parties involved would not reveal the company's identity, but *Science* has learned that it is Microfibres Inc., of Pawtucket, Rhode Island.) Company officials insist that the research is too premature to publish, and Memorial Hospital officials—who deny that they closed Kern's clinic in retaliation—say Kern cannot make his findings public because he is bound by a confidentiality agreement that he signed with the company. Now, the medical school too has been drawn into the dispute: Last week, some of Kern's colleagues both inside and outside the university urged administrators to stand behind Kern over what they see as an issue of academic freedom, and the school has launched an inquiry.

Center of the storm. Memorial Hospital's David Kern.

This tangled saga began last spring when Kern—then Memorial's chief of general internal medicine and head of its Occupational and Environmental Health Service since 1986—examined a young man from the plant. Both the man and another plant employee Kern had examined a year earlier had symptoms of interstitial lung disease (ILD), an inflammation of the alveoli that can lead to permanent scarring and reduced breathing capacity.

When Kern learned that a similar outbreak had occurred at the company's Ontario plant in 1990, he alerted the National Institute for Occupational Safety and Health (NIOSH) and offered to probe the outbreak as a paid consultant to the company. Microfibres agreed and asked Kern to sign a confidentiality agreement to protect the company's "trade secrets." Kern says he found six more employees at the 150-employee plant with what he considers work-related ILD—a far higher incidence than expected, he says, as the disease's incidence in the general population is one in 40,000. But he was unable to link these cases to any specific chemical or airborne material at the plant.

Last October, when Kern showed company officials a draft abstract on the outbreak, prepared for the May 1997 meeting of the American Thoracic Society, he says the company threatened to sue him for breaking the secrecy agreement if he submitted the abstract. Kern then broke off the consultancy agreement—and went ahead and submitted the abstract anyway, because, he says, it does not identify the company and includes no proprietary information. Moreover, a brief report on the outbreak itself, co-authored by Kern and staffers at NIOSH—which also does not identify the company by name—is currently under review in the *Morbidity and Mortality Weekly Report*, published by the Centers for Disease Control and Prevention, NIOSH officials say.

Kern insists that company officials knew he planned to publish his research when his consultancy began and only balked when he turned up additional cases of ILD. But offi-