Fraud and the Structure of Science

Is fraud a trivial excrescence on the process of science or do the recent cases have deeper roots?

There is little doubt that a dark side of science has emerged during the past decade. In ever-increasing detail, the scientific and general press have reported the pirating of papers and the falsification of data. Four major cases of cheating in biomedical research came to light in 1980 alone, with some observers in the lay press calling it a "crime wave." Federal investigators say two of these cases may end in criminal charges.

In a profession that places an unusual premium on honesty, the emergence of fraud has created something of a stir. Scientific societies are holding symposia on the subject. The National Institutes of Health has taken administrative steps to cope with the putative rise in cheating. And Congress, as it does with issues significant or otherwise, is preparing to hold hearings* on the falsification of data in biomedical research.

Is the issue important? After all, reported cases of cheating are few, and NIH funnels government funds into nearly 18,000 extramural projects. *Science* recently put the question to a dozen research directors, NIH officials, bench scientists, and sociologists. One recurring observation was that fraud has always been around, but not always advertised. A Nobel laureate, for example, was said to have coauthored a paper in the early 1960's that was retracted due to the cheating of a junior associate. The unseemly details of the Nobel retraction never went out of the lab, and therein, say some observers, lies one difference between the finagling of yesterday and today. Changes in comtemporary science and its interactions with society are making fraud in the labs more visible.

- John Long, a researcher with $750,000 in federal funds at Massachusetts General Hospital, forged data and for 7 years watched over a cell line for the study of Hodgkin's disease that proved useless (*Science*, 6 March 1981).
- Vijay Soman, a researcher at Yale medical school, plagiarized a rival's paper, fabricated data, and received for 1980 alone some $100,000 in NIH support. Eleven papers were retracted. He ultimately returned to his home in India, but left his coauthor and boss, Philip Felig, in an administrative and ethical tangle (*Science*, 3 October 1980).
- Elias A. K. Alsabti, a young researcher from Jordan, pirated almost word-for-word at least seven papers and published them in obscure journals. (*Science*, 27 June 1980).
- Marc Straus, a Boston University researcher who in 3 years was awarded nearly $1 million in cancer research grants, submitted reports containing repeated falsifications. He resigned under fire, insisting that he was the victim of a conspiracy by select members of his 20-person staff. More than 2 years later, after the Boston *Globe* ran a five-part series on the affair, the National Cancer Institute initiated an investigation.

In response to these and a few other incidents, Congress has invited two witnesses to the falsification drama, Long and Felig, to come and give their views on what, if anything, is happening to U.S. biomedical research. Also invited are a bevy of NIH officials, research directors, and bioethicists. Cheating is also being discussed during symposia at the upcoming annual meeting of the Council of Biology Editors. Meanwhile, at the Harvard School of Public Health, the seventh national conference on Public Responsibility in Medicine and Research just held a session on "How to detect and prevent fraudulent or unethical research."

Until recently, charges or even discussions of scientific fraud were seldom aired in public. Most scientists, conscious of their image and eager to avoid political interference, tried to stay out of the limelight. Control was an internal matter. An informal group of scientists could hold court and decide to ban an offender from the realm of research. More fundamentally, science was said to be self-correcting. If an experiment was important enough, other scientists would try to repeat it. This self-correcting mechanism would expose cheating and encourage honesty. It would detect and deter. Dubbed "organized skepticism," this view was originally set forth by Robert K. Merton, the father of the sociology of science. "Scientific inquiry," he wrote, "is in effect subject to rigorous policing, to a degree perhaps unparalleled in any other field of human activity." Initially propounded in 1942, this view has become the conventional wisdom. Donald Fredrickson, director of NIH, today puts it this way. "We deliberately have a very small police force because we know that poor currency will automatically be discovered and cast out."

A sterling example of such self-correction comes from the case of the Nobel retraction. The incident unfolded at Yale in the late 1950's, with the arrival of a young graduate student in biochemistry. Working in the lab of Melvin Simpson, the student quickly made significant gains in the cell-free synthesis of cytochrome *c*, a key protein in cellular energy-releasing reactions. In early 1960, Simpson and the student coauthored a paper on the successful experiments that received wide attention because it was the first time such a single, highly purified protein has been synthesized outside a cell. The success carried the student, now equipped with a Ph.D. from Yale, to the lab of Fritz Lipmann at Rockefeller University, where he coauthored a paper with the Nobel laureate. The promising career, however, soon suffered a setback.

Simpson, back at Yale in late 1960 after spending several months on sabbatical in England, reassembled his lab and started trying to extend the successful experiments with cytochrome *c*. His efforts met with failure. "I had gone all around Europe giving seminars on our success," he recalls. "And now I couldn't repeat it. Imagine the agony." A call from Lipmann at Rockefeller revealed that people in his lab were also having difficulty repeating the student's work. The student was called back to Yale and told to duplicate the cytochrome *c* experiment. He worked under

*The House Science and Technology subcommittee on investigations and oversight, chaired by Albert Gore (D–Tenn.), will hold hearings on 31 March and 1 April, as this issue of *Science* is going to press. Testifying will be Donald Fredrickson (NIH), Philip Handler (National Academy of Sciences), LeRoy Walters (Kennedy Institute of Ethics at Georgetown University), Philip Felig (Yale), John Long (formerly of Massachusetts General Hospital), Ronald Lamont-Havers (Mass General), Patricia Woolf (Princeton), Stuart Nightingale (Food and Drug Administration), Alexander Capron (President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research), and William Raub (NIH).

0036-8075/81/0410-0137$01.25/0

around-the-clock supervision, and failed. He was subsequently told to leave research in general. Two retractions, one from Simpson and one from Lipmann's lab, were published in late 1961. Sometime later it was discovered that the student's undergraduate college in Massachusetts had no record of his ever receiving a degree.

Since that time, revelations of cheating—but not necessarily cheating itself—seem to have slowly but steadily increased. The cause? According to Robert H. Ebert, former dean of the Harvard medical school, part of the reason may be increasing pressure. Writing in the *New York Times* about the fabrication of data by John Long at Mass General,

MX Lobotomized by Air Force, Critic Says

The Air Force is merely doing a little budget trimming, according to spokesmen. But Richard Garwin, IBM scientist, consultant to the Pentagon, and a frequent critic of the Air Force, says there is more than this in a recent cutback ordered on the MX missile program.

Garwin says there may be a special policy slant in the decision announced 12 March to cancel work on part of the electronic brain of the MX. The Air Force wants to base this newest and biggest nuclear weapon on land in America's southwestern deserts. Garwin argues that it makes more sense to put the MX in the ocean, and he thinks that the program cut will make it difficult to keep the sea-based option available.

Defense Secretary Caspar Weinberger has said that he wants to keep open all the options for siting the missile bases. President Reagan also has suggested that he is ready to scrap the basing scheme chosen by the Air Force. A group of 15 civilian weapons specialists has been convened to reconsider all the proposals and report back to the secretary of defense by July. They may find that the Air Force has decided already what the outcome will be.

The MX contract in question called for the Charles Stark Draper Laboratory of Cambridge, Massachusetts, which is affiliated with MIT, to design two things: part of the missile's inertial guidance system and a radio receiver that could be used to read signals from a network of satellite beacons known as the NAVSTAR Global Positioning System. The Navy uses NAVSTAR as a navigational aid. By interpreting the Doppler shift of the radio signal, one can obtain a precise fix on one's location and direction anywhere on the globe.

Garwin is enthusiastic about giving missiles the capacity to receive NAVSTAR signals. Doing this would greatly increase confidence in missile accuracy, he says, and it would bring about a big change in strategic planning. In land-based missiles, the improvement would be important, but only incremental. In missiles for submarines or other compact mobile systems, it would be revolutionary, Garwin says.

The potential improvement in guidance could make submarine-launched missiles as accurate as land-based missiles. This would transform the submarine in strategic terms from a blunt retaliatory weapon into a precise instrument of war. Because missiles at sea would be both accurate and invulnerable to attack, they would become more important than land missiles, according to Garwin. This shift of emphasis would demolish the theory that the Soviets might be tempted to launch a surprise attack on bases in the United States in order to knock out America's most threatening missiles. Garwin says that the Air Force ought to put a NAVSTAR receiver on the MX to make the best use of the missile and keep options open.

The Draper laboratory heard formally in December that it had won the $41-million contract to design part of the guidance system and a NAVSTAR receiver for the MX. The radio system is known officially as the Missile Accuracy Evaluator (MAE). According to the Air Force, it was meant to be a tracking mechanism that would let observers follow the course of the MX in tests, supplementing data given by radar tracking stations. According to Garwin and other experts, MAE has another value: it could serve as the first generation of a radio link with NAVSTAR, and be used eventually to help guide a missile to its target.

Early this year the Air Force told Draper laboratory to reduce expenses on the project. Then in March the Air Force ordered work to stop on the MAE system, even though MAE had proved viable and was ready for flight packaging. This decision may have saved the Air Force about $22 million. The planned cost of the total MX project, including bases, is more than $34 billion.

If the MX were based at sea, says one of the nation's top specialists in missile guidance (not Garwin), it would be useful to have a link between NAVSTAR and the missile in order to help it get its bearings. "But this wouldn't be the only way to do it," he said. One could also build the necessary guidance control systems into the ships that carry the missiles, but that would take up space aboard the ships and be "very expensive."

Colonel Neil Buttimer of the Air Force's Ballistic Missile Office in San Bernardino, California, says the decision to cancel work on MAE has nothing to do with basing options. An "updated assessment" of the quality of the missile's primary guidance system "indicates that the likely error sources for MX would be sufficiently integrated so that the MAE program would be of only marginal value." Other sources of flight data could be used in place of MAE, the Air Force has decided.

In basic terms, Buttimer says, "Our budget was cut by Congress. We were looking for ways to save money, and we lined up our programs and asked what are the most important and what can we get rid of with the least injury?" As it happened, MAE came out at the very bottom. Buttimer adds, "If it were important, which it isn't, it wouldn't be difficult to start it up again. . . . We'll know by June or July which way we're going to go" on missile basing.

Wouldn't it make sense to trim the budget somewhere else, rather than to call off work on what could be a very useful innovation in missile guidance? Buttimer says the Air Force does not want to rely on NAVSTAR for missile guidance in any case. It is too vulnerable to radio jamming or preemptive attack by the Soviets.

Garwin, who regularly finds himself disputing official wisdom, claims there is no technological threat to the NAVSTAR link that cannot be solved relatively cheaply.

—ELIOT MARSHALL

Claudius Ptolemy reportedly invented data to support his own astronomical theories.

Ebert said "it would be a mistake to consider this an example of human frailty and nothing more. Medical schools and academic research centers have inadvertently fostered a spirit of intense, often fierce competition, which begins during the premedical experience and is encouraged thereafter. . . . There is intense pressure to publish, not only to obtain research grant renewals but in order to qualify for promotion."

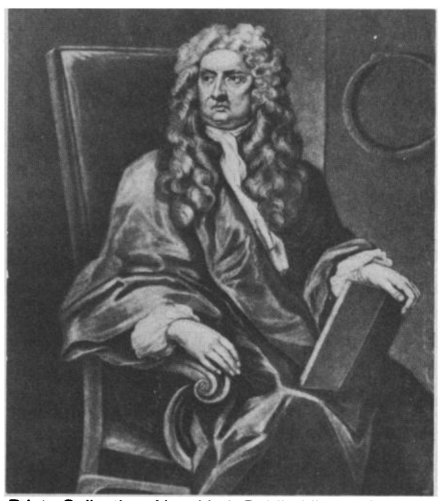
The implication in this account of a rise in cheating itself is dismissed in many quarters. Pressure, say a chorus of commenters, has always been around. Moreover, many hold that the rate of finagling has remained roughly the same throughout the years, and cite the purported cooking of data by Mendel, Newton, and Ptolemy to back up their beliefs.

A radical view of the ubiquity of fraud comes from philosopher of science Paul Feyerabend (*Science*, 2 November 1979), who holds that small-scale cheating is essential to the advancement of science. He argues that no theory, no matter how good, ever agrees with all the facts in its domain. A scientist must therefore rhetorically nudge certain facts out of the picture, defuse them with an ad hoc hypothesis, or just plain ignore them. A similar but less polemical view is expressed by philosopher Thomas S. Kuhn (*Science*, 8 July 1977). Kuhn divides the history of science into periods of normal and revolutionary activity, arguing that during normal periods, anomalies observed by the scientist must be suppressed or ignored.

If finagling of one sort or another is endemic, what then causes increased exposure? Here it is necessary to make a distinction: exposure of fraud to other scientists and exposure to the public.

In the first instance, one mechanism that may bring cases of cheating out in the open is the denunciation of scientists by one another due to cutbacks in research funding, according to Ronald Lamont-Havers, a former NIH official who witnessed the Long affair from his position as director of research at Mass General. If this is indeed the case, troubled times may lie ahead. Since 1979, NIH has had a drop in purchasing power, and this year the percent of approved grants lucky enough to get funded has dropped to 30—an all-time low.

The increasingly close scrutiny of research that has direct implications for public policy or public health is also a factor in inter-scientist exposure, according to Columbia University sociologist Harriet Zuckerman. This clearly seems to be the case in the Straus affair



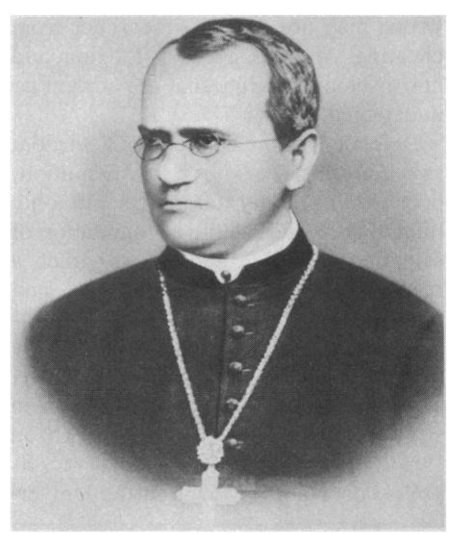
Prints Collection, New York Public Library, by permission of *New Republic*

Isaac Newton in his Principia *relied in places on an unseemly fudge factor.*

at Boston University. Data from about 200 patients studied by Straus and his team were kept in the computer files of the Eastern Cooperative Oncology Group, a 40-hospital consortium funded by the National Cancer Institute to conduct large-scale testing of new cancer treatments. In 1978, five members of Straus's team disclosed to officials at Boston University problems with the data. Falsifications ranged from changing a patient's birthdate to reporting treatments and laboratory studies that were never done and inventing a tumor in a patient who had none. Boston University says a detailed study of medical records found no evidence of patient mistreatment or inappropriate care. Disagreeing with this view is a Food and Drug Administration official familiar with an ongoing investigation of the Straus affair: "To say the least, some of this had serious clinical implications, both in the sense that the patient in the study was endangered, and that data generated would present conclusions that were poorly founded."

Concerning exposure to the public, one factor repeatedly singled out is the growth of a vigorous scientific press. Indeed, the National Association of Science Writers, founded in 1934 by 15 reporters, now has more than 1000 members. And clearly, the NCI investigation of the Straus affair would never have materialized had it not been for the series in the *Globe*. Some observers, however, suggest that the press for the most part tends to purvey, rather than initiate, exposures.

A general rise in social consciousness among scientists may account for some of the increasing public exposure, according to E. Frederick Wheelock, a microbiologist at Jefferson Medical College in Philadelphia whose work was pirated by Jordanian researcher Alsabti. "In the past," he says, "the system was much more closed. People were afraid to call attention to cheating." In his own case, Wheelock at first hesitated to charge Alsabti with piracy. Wheelock had kicked Alsabti out of his lab after two young researchers came to him with proof that Alsabti was making up data. Later, when Wheelock saw his work being published in the scientific literature by Alsabti, he discussed the problem with his program manager at the National Cancer Institute, who suggested that he alert the wider community. After first writing to Alsabti and demanding retractions (that did not materialize), Wheelock wrote letters to *Nature*, *Science*, *Lancet*, and the *Journal of the American Medical Association* and de-



The Bettmann Archive, Inc.

Gregor Mendel reported data that some researchers say was too good to be true.

scribed ways for researchers to "avoid such episodes in the future."

The list of possible reasons for increasing exposure rambles on, most everyone having their own pet speculations. Lurking in the record of events, however, is an intriguing contradiction. A review of the cases where cheating has come to light during the past decade shows that the failure to duplicate experiments plays a relatively minor role in uncovering fraud. This self-correcting mechanism "worked" in earlier episodes: in the cases of Mendel, Newton, and Ptolemy (though it took two millennia), or in the case of the Nobel retraction. During the past decade, however, other means have predominated, the mechanism often being the detective work of young lab assistants or young scientific rivals who have extra-experimental evidence of cheating, who have some independent reason for suspicion. This was the case in all four of the 1980 fraud episodes. It was also the case with the Sloan-Kettering affair and the painted mouse of William Summerlin (*Science*, 14 June 1974), although Summerlin's work was also under fire because it could not at the time be repeated.

This gap between real and ideal ways of detecting and preventing fraud (what sociologists of science euphemistically refer to as the "social control of science") has helped fuel a heated critique* of the conventional wisdom during the past decade.

On the deterrence side of the debate, critics have argued that the self-correcting mechanism does not distinguish between error and fraud. In the published literature, an experiment is only found right or wrong. Given the ever-present academic pressure to succeed in a spectacular way, this chance of being found wrong may not deter a researcher from cheating. After all, guesses, fudging, and unconscious finagling that are correct go undetected.

Defenders of the conventional wisdom say that this weakness, by definition, does not make any difference. The only thing that matters is the accumulation of scientific "truth," and not whether a falsifying researcher is caught and punished.

It is here, on the detection side of the debate, that critics rail most vehemently. The acceptance or rejection of claims in science often depends not so much on "truth," according to such observers as philosopher Ian I. Mitroff at the University of Pittsburgh, but on who makes the claim and how well the claim fits prevailing beliefs. In short, the goodness of a reputation or the attractiveness of a theory often gives immunity from scrutiny.

This circumvention of the idealized mode of detection was probably a factor in why the problems with John Long's contaminated cell lines at Mass General escaped detection for so many years. He worked in a prestigious lab at one of the world's leading teaching hospitals. "With the credentials of background and training that Long presented, the study section would expect that he would be aware of this [contamination] problem," says Stephen Schiaffino of the NIH division of research grants.

Of late, failure to duplicate experiments has played a minor role in uncovering fraud.

Immunity from scrutiny was also clearly a factor in the case of Cyril Burt (*Science*, 26 November 1976), the English psychologist whose studies of identical twins supported his theory that intelligence is determined partly by heredity, and whose work went unchallenged during his lifetime. As a government adviser in Britain in the 1930's and 1940's, Burt was influential in setting up a school system in which children were assigned to one of three educational levels on the basis of a test given at the age of 11. According to Leon Kamin, a psychologist at Princeton, Burt's data remained unchallenged for so long because they confirmed what everyone wanted to believe. "Every professor knew that his child was brighter than the ditchdigger's child," he says, "so what was there to challenge?"

Burt's work was picked up by researchers in the United States, and figured prominently in the debates over whether heredity might underlie racial differences on IQ scores. Eventually, after a reign of nearly 40 years, his data were found to be riddled with internal implausibilities and basic methodological oversights. Some researchers concluded that Burt may have doctored or even invented his collection of IQ data.

Critics for the most part do not argue that the conventional wisdom is wrong, but rather, taken alone, it is inadequate to explain how science really works on a day-to-day basis. Perhaps the most troubling observation is that even when the self-correcting mechanism works, it addresses only experiments and observations that are "important" to pure science, to the accumulation of scientific truth. No one, after all, takes much time to repeat clinical trials of new drugs, therapies, or treatments. Replication of a multi-institutional clinical trial, such as the one at Boston University that Straus worked with, is financially and structurally impossible. In terms of the self-correcting mechanism, these are not applicable areas of research, although they may be important in terms of patient welfare.

Just as there was no scientific or institutional mechanism to detect or deal with fraud in the Straus affair, neither was there a federal mechanism. When three top officials at Boston University medical center flew to Washington to tell the NCI director about their rapidly unfolding problems, NCI told them there was nothing the government could do.

The slow response of the federal bureaucracy, the questioning of the self-correcting mechanism, and the emergence of a few graphic examples of fraud have combined to stir considerable activity concerning data abuse. At Boston University, the multi-hospital group that got stuck with the project's bad data has set up a system of random audits to ensure that the program will never again be vulnerable to such falsification. Congress is in the process of holding hearings. The President's commission on bioethics plans to hold a number of sessions at "sites of controversy involving the conduct of research."

Confronted with the increasing reports of fraud-related incidents, NIH recently took steps to prevent abuse in the future. In November 1980, debarment regulations went into effect that allow the government to cut off an entire institution from NIH grants if just one researcher is caught misusing grant money or falsifying reports (*Science*, 14 November 1980, p. 746). This sweeping mechanism was needed, says NIH associate director William Raub, in order to put the onus for prevention and detection of fraud on the institution. Previously, institutions might have been tempted to look away. Over the years, the administrative costs charged by universities for nurturing the research enterprise have risen so that they now average more than 27 percent of a grant.

No person or institution has yet been debarred, and NIH officials say they have no plans to make debarment retroactive. If it were, all of Yale University, for example, could well be cut off from

*For a summary, see J. Gaston, "Disputes and deviant views about the ethos of science" in *The Reward System in British and American Science*, (John Wiley and Sons, 1978), pp. 158–184.

the federal research-dollar pipeline. In addition to the threat of debarment, NIH officials say they have now built into their vast computer network an alert system so that NIH administrators are warned if an investigator applying for a new grant is himself under investigation for cheating. Flagged so far by this system are Straus, Soman, and a third, unidentified researcher who is currently under investigation.

Is it important? Perhaps the emerging issue of fraud represents a small, seamy side of science that warrants nothing more than a cursory glance before being tossed onto the pile of passed-over issues. One might argue that the major cases are few, and the minor ones are just that, minor. Science is above it all. Nobel Prizes are awarded and greatness is measured not on the basis of "honesty," but insight. Newton and Mendel may have finagled, but their theories are today committed to memory by every high school student.

In a sense, all this is correct. It is also true that fraud in the literature wastes the time and money of researchers who pursue leads only to find them wrong. Simpson spent 1 year untangling the cytochrome *c* mess, and, because of this unanticipated chore, lost a priority battle in a different area of biochemistry. Similar amounts of time are probably wasted in other fabrication episodes. Further, in a profession where "organized skepticism" is meant to be the rule, the emergence of a type of fraud not detected by this self-correcting mechanism may prove especially corrosive to community ideas. This mechanism did not and could not deter data fabricators at Boston University, with the result that patient safety was probably jeopardized. And the fact that immunity from scrutiny often seems to supersede any kind of "organized skepticism" can only lead to the discouragement of the young, who tend to be far from immune. In the case of the imbroglio at Yale, it was a 29-year-old NIH researcher who brought charges against Soman, an assistant professor, and Felig, a professor with an endowed chair. "I just found it hard to believe that Felig had engaged in any hanky panky," said an appointed NIH auditor who, after a wait of 6 months, decided not to investigate the data of Soman and Felig. During this noninvestigation, the young researcher quit NIH and research in general.

No matter why they come forth, the recent cases illuminate much. They disclose a gap between the ideal and the real, between reliance on automatic self-policing and the fact that mechanisms such as immunity from scrutiny often prevail. They hint at support of philosophical views that say finagling of one sort or another may be endemic to the research enterprise. Perhaps further study of the dark side will disclose more about the structure of science. At the very least, the recent cases illustrate that "organized skepticism" and the self-policing nature of science need themselves be taken with a little more skepticism.—WILLIAM J. BROAD

Interferon: No Magic Bullet Against Cancer

It may be medically and commercially more important in fighting viral infection

Daytona Beach, Florida.—Interferon as an anticancer treatment has not lived up to expectations; however, its ability to fight off viral infection may prove to be very important both medically and commercially, according to Frank J. Rauscher, Jr., senior vice president of research of the American Cancer Society, at a meeting for science reporters here.

Interferon does not appear to be any better than available chemotherapeutic agents in treating non-Hodgkin's lymphoma, multiple myeloma, breast cancer, or melanoma, said Rauscher. Of the 82 patients tested thus far in the cancer society's program, only about 25 to 40 percent responded favorably. These responses ranged from stabilization to complete remission in a few patients. The reasons why the remaining 60 percent of the patients failed to benefit warrant further research, he said.

Some of the treated patients relapsed 6 months after remission, although a few remained in remission 8 months after treatment was started. "If there's anything discouraging about interferon, it's that remission doesn't seem to last," said Rauscher, a former director of the National Cancer Institute.

The investigators had hoped that interferon might be free of the side effects associated with the most cytotoxic drugs currently used in cancer treatment. But the fact is, Rauscher said, interferon has side effects similar to those of other medications, although to a much lesser degree. Interferon-treated patients have suffered hair loss, nausea, bone marrow depression, and sudden fever of about 102°F. The most severe side effect has been lethargy, experienced most often by elderly women in the later stages of breast cancer.

Remissions don't seem to last

Frank Rauscher of the American Cancer Society

The toxicity seems to be the same even with interferon that is now 1000 times more pure as a result of recombinant DNA techniques. Less pure interferon was previously extracted from leukocytes. This suggests that the molecule itself, rather than the impurity of the preparation, is the toxic substance. The side effects might be avoidable because as many as eight to ten genes have been pinpointed which all code for interferon. The interferon expressed by the various genes might have different effects in the body. It is not known whether the several companies producing recombinant DNA interferon have been splicing the same gene or not.

Only in the past year has interferon been in great enough supply to begin clinical trials. A year ago, the main supplier of interferon was a laboratory in