

Integrity of Research Papers Questioned

Two researchers from NIH allege that the scientific literature may be full of minor errors; others argue they have very little evidence

IN the spring of 1983, Walter W. Stewart and Ned Feder "set aside" their usual laboratory work at the National Institutes of Health to embark on a wholly unusual investigation of the validity of the scientific literature. Through that they hoped to shed light on the question of whether coauthors of scientific papers, journal editors, and referees are "meeting the standards conventionally accepted as necessary" to assure that the literature is accurate.

Stewart and Feder were interested in a number of points including the factual consistency and accuracy of papers, duplicate or "republication" of data in abstracts and articles, and the appropriate assignment of credit and responsibility when it comes to deciding whose name will go on a manuscript and whose will be left off. They took as their study sample 109 papers written by John Darsee and 47 coauthors, some from Emory University School of Medicine, others from the Harvard Medical School laboratories of cardiologist Eugene Braunwald. In 1981, Darsee was discovered to have committed fraud at Harvard and it was later discovered that he also committed fraud in medical school at Emory, and even before that as an undergraduate at Notre Dame.

Using the Darsee papers as a "sample of convenience," Stewart and Feder concluded that those papers are full of errors and inconsistencies, quite apart from the frankly fraudulent data, and that the coauthors could have spotted those errors had greater care been exercised in preparing the manuscripts for publication. Stewart and Feder also assert that journal editors and referees failed to attend sufficiently to detail to catch what the NIH scientists regard as obvious errors.

A manuscript reporting their observations was first submitted to *Nature* in the fall of 1983, but was subsequently withdrawn when the authors concluded that changes mandated by legal considerations of libel watered it down too much. Stewart and Feder went on to offer their paper to several journals, including *Science* and the *New England Journal of Medicine* where it was rejected on editorial grounds. In a letter that has become part of the public record through Freedom of Information Act requests for

access to legal and editorial documents on the Stewart-Feder manuscript, *Science* editor Daniel E. Koshland, Jr., wrote them in July 1985 and said, "I do not concur with you that details of this case are extraordinarily important. . . ." *New England Journal* editor Arnold S. Relman rejected the version he saw. "What was true wasn't new, and what was new wasn't true, or in any event very misleading," he recently told *Science*. But other editors told Stewart and Feder they thought their paper was important and in the public interest, but too expensive to handle because of its potentially libelous nature.

Almost everyone agrees that coauthors could take their responsibility for the accuracy of data more seriously.

In addition to submitting their manuscript to many journals, Stewart and Feder also reported the essence of their findings to a committee of the Congress on 26 February 1986, when they asserted that fear of libel was preventing publication. And, within the past few months, they sent copies unsolicited to every member of the National Academy of Sciences. Because this study had NIH approval, NIH paid copying and postal costs.

Finally, the saga of the Stewart-Feder manuscript now comes full circle with publication of a much edited version in the 15 January issue of *Nature*. In an accompanying editorial, John Maddox calls attention to a companion article in which Harvard's Braunwald offers a rebuttal to allegations that he and his colleagues allowed sloppy work to be reported in the scientific literature. (Researchers at Emory have yet to be heard from.)

In his comment in *Nature*, Braunwald also recalls, "When we discovered and ultimately exposed all of Darsee's misconduct we were staggered at its extent. . . . Perhaps one positive outcome of the unfortunate Darsee affair is that it has helped focus

attention on fraud in science, in part prompting both research institutions and granting agencies to develop more formal rules and procedures for dealing with suspected fraud." (The Association of American Medical Colleges and the Association of American Universities each produced guidelines. Formal procedures were recently established by NIH.)

One of the more unusual features of this event is the suggestion, implicit in the paper and its publication, that one can generalize from the faults in the Darsee papers to the scientific literature generally. If these papers are full of errors, is it not possible that the scientific literature is chockablock with papers whose data may be inaccurate? To this, almost everyone answers—"You cannot say."

Stewart and Feder take this on themselves. In their *Nature* paper, "The integrity of the scientific literature," they write, "The results of our study naturally raise the question: what fraction of papers in the biomedical sciences are not supported by primary data at the time of publication? The question cannot be answered with the papers in our sample, since most are known to be partial or complete forgeries and thus are atypical." Further, Stewart and Feder make this observation about the nature of many of the problems they identified in that sample: "We emphasize again that many of the errors and discrepancies are minor, for example a small discrepancy between a numerical value printed in a table and what is supposed to be the same value printed in the text."

Why, then, publish it at all. In his editorial, *Nature's* Maddox says, "The truth is that the document by Stewart and Feder has by now acquired a notoriety comparable with the Darsee affair itself."* Maddox goes on to say, "Stewart and Feder have complained that they could not publish what they chose because of 'threats of libel.' They have not understood that the unfettered right to publish scientific data does not equate with a

*The Darsee affair was widely reported in the press, including *Science*, which published the following News and Comment articles: 29 January 1982, pp. 478-482; 12 February 1982, pp. 874-876; 1 April 1983, pp. 31-35; 27 May 1983, p. 936.

right to denigrate others' characters."

Nevertheless, the Stewart-Feder paper focuses attention on questions about scientific publication that are widely regarded as important. Maddox writes that "the general conclusion at which Stewart and Feder arrive, that coauthors should be more zealous in their scrutiny of what is about to be published in their names, anodyne though it may be, cannot be a bad prescription."

Even if many of the errors Stewart and Feder identify can be explained, the fact remains that the errors are there. In one instance, which they describe as "so glaring as to offend common sense," Stewart and Feder call attention to a family pedigree of individuals with heart disease that shows one man having fathered four children ages 8, 7, 5, and 4 by the time he was only 17 years old. That the ages noted on the pedigree were *meant* to indicate the age at which heart disease was diagnosed is of little help to readers of the original article who would have no way of knowing that.

An interesting aspect of the Stewart-Feder story is the NIH's official role in all this. As intramural scientists, the two researchers enjoy substantial freedom to tackle whatever projects they think are important, and at first there was no objection to their undertaking an analysis of the literature along with their laboratory research. But as the Darsee study came to occupy more and more time, questions arose within the National Institute of Diabetes and Digestive and Kidney Diseases, where they work, about the fact that their laboratory science had all but come to a stop.

The matter was taken to the office of the NIH director where, in the spring of 1984, Joseph E. Rall, deputy director for intramural research, and others made an affirmative decision that Stewart and Feder should be permitted to complete their literature study. William Raub, newly named deputy director of NIH, was in on the decision as head of the NIH office on research fraud. "We agreed that it was a proper scientific inquiry," Raub told *Science*, adding that no one envisioned that completion of the project would occupy Stewart and Feder full time for two-and-a-half years more. Nor did they appreciate the amount of time the NIH's own lawyer would spend helping them produce a legally defensible manuscript.

Only a couple of the coauthors retained legal counsel to block publication of the original versions of the Stewart-Feder paper, but it only takes one to raise roadblocks that can cause interminable delay unless allegations are provable absolutely. Bancroft Littlefield of Boston took the lead as Braunwald's counsel.

In the end, where earlier versions and

congressional testimony talked about "serious misconduct" and "statements for which the evidence indicates that they knew or should have known of their falsity," the published paper uses less damaging language, referring instead to "incomplete" rather than "misleading" statements, for instance. Stewart and Feder are not entirely happy with the changes and *Nature* notes that some editorial changes (including the latter) were made without their consent. But any threat of suit seems to have been averted, at least from the Harvard coauthors, and Littlefield says it is better now that the paper be published than to allow further charges that Braunwald and others are using legal threats to stifle the free flow of information.

What now? Stewart and Feder will face demands from their institute that they return to laboratory science, and they have said they want to do so. Their own work has been with a class of agents known as "Lucifer dyes" that enable visualization of the branching patterns of nerve cells. Feder, 59, graduated from Harvard College and Harvard Medical School. Stewart, 42, graduated

from Harvard College but lacks a doctorate. The two have worked together at NIH for many years. In 1981 Stewart published two highly regarded and cited research papers on the dyes.

But they have also expressed an interest in continuing to work on questions about the integrity of scientific literature to answer questions that the study of the Darsee papers could not answer. The appropriate balance between laboratory research and other professional inquiry by intramural scientists is an issue that NIH officials will have to grapple with.

At the conclusion of their *Nature* paper, Stewart and Feder suggest auditing a random sample of the literature to check on its integrity. But they also note that there are potentially serious costs to "examining the practices of scientists" this way. "Systematic examination of scientific practices might even weaken the fabric of trust that is essential to the functioning of science," they write, adding that even though science is vulnerable to abuse, it is "perhaps even more vulnerable to harm by regulation." ■

BARBARA J. CULLITON

EPA Finds Western Lakes Free of Acid Pollution, But Vulnerable

Unlike lakes in the industrial East, those in the West remain unspoiled by acid precipitation, according to a survey by the Environmental Protection Agency (EPA). But EPA researchers warn that many western lakes are susceptible to long-lasting damage because they have little acid buffering capacity. At present, there is less acid pollution in the West, but the West's vulnerability may be greater.

The survey, costing \$7 million, is part of EPA's continuing study of the effects of acid rain on surface waters. It was released on 15 January. An admitted weakness, EPA researchers say, is that the data represent a one-shot random sample taken in the fall. They do not describe the worst chemical shocks that befall lakes during the spring when acid-bearing snow melts.

EPA published a similar analysis of eastern lakes in 1985 (*Science*, 2 September 1985, p.1070). The eastern study (costing \$12 million so far) revealed that 10% of lakes in the Adirondack Mountains of New York were verging on acidic, with a pH level of 5 or less. Around 10% of the lakes in the Upper Peninsula of Michigan and in Florida are also in this problem category. In contrast, in the West, only one—Fern Lake, a sulfur hot spring in Yellowstone National Park—had a pH below 5.

Richard Linthurst, manager of this research program, thinks the best measure of the environmental threat to lakes is "acid neutralizing capacity" or ANC. It reflects buffering capacity—the ability to absorb acid without a change in pH. A lake with an ANC rating below 50 is considered to be at risk of acidification. In California, 37% of the lakes fell in this category, and in other western states, the number at risk ranged from 5% to 20%. In the East, the proportion of low-ANC lakes ranged from 9% in upper Wisconsin to 36% in the Adirondacks.

EPA researchers think western lakes may be especially vulnerable because they lack the surrounding vegetative watershed that acts to protect other lakes. In the East, rain or snow loaded with acidic compounds may be buffered by passing through the watershed. This is less likely to happen in the West. As evidence of the difference, Linthurst pointed out that conductivity values (a measure of the amount of chemical material dissolved in water) were "very low" in one-quarter of the lakes in the West. This indicates that the surrounding watersheds contribute little to lake chemistry. Linthurst concluded that western lakes are likely to be more directly sensitive to the effects of acid deposition. ■ ELIOT MARSHALL