Letters

Epidemiologic Investigation

In his article "Scientific standards in epidemiologic studies of the menace of daily life" (2 Dec., p. 1257), Alvan R. Feinstein reviews several axioms of epidemiologic investigation. However, among his examples of objectionable practices he has listed an investigation of some 121,000 American nurses who for many years have supplied data on various aspects of their lives in order to test a number of hypotheses that are of the utmost interest to the public health. Citing one of the findings in this study (*1*)—that there was an association in these women between ingestion of alcohol and cancer of the breast, Feinstein accuses the investigators (epidemiologists of major stature) of "dredging" up the association, presumably by examining the data for a large number of associations, and then finding one by chance. Nothing could be farther from the truth. The hypothesis associating ingestion of alcohol with carcinoma of the breast has been advanced many times. Feinstein cites abstracts of one case-control study and two small cohort studies in which no association was seen, but neglects to mention that in eight of 12 case control, and four out of four prospective, large cohort studies, distinct associations were found between ingestion of alcohol and carcinoma of the breast. Furthermore, had Feinstein inquired of the investigators, he would have found that a question about alcohol was included in the initial questionnaires prepared in 1980, so that the relation between alcohol and malignancies could be investigated, as clearly stated in the initial grant proposal. Also, as stated in their paper, when the investigators of the Nurses Health Study initially found an association between alcohol and carcinoma of the breast on the basis of the first 2 years of accumulated data, they were concerned lest this association be due to chance alone. Accordingly, they set aside the first 2 years of data and collected the next 2 years of information, analyzing the latter independently; again the association was found. None of the investigators who found this association, which now appears not to be due to chance alone, has claimed a causal relation, but this possibility cannot be excluded (*2*).

We are forced to suggest that at best, Feinstein may not have reviewed the literature thoroughly. Any questions that he might have raised concerning the conduct of the study would have been answered courteously and promptly. He chose instead to demean the study and the investigators before he had determined, with the use of his stated criteria of precision, whether his claim was justified.

Large epidemiologic studies are difficult to conduct, and it is easy to become concerned about potential flaws that are quantitatively unimportant. This may have led Feinstein and his colleagues in the past to doubt the validity of a number of conclusions that are now beyond doubt, such as the relation between exogenous estrogens and endometrial carcinoma (*3*), the relation between diethylstilbestrol and vaginal carcinoma (*4*), and the relation between aspirin and Reyes syndrome (*5*).

I believe that the Nurses Health Study, and its investigators, satisfy the most rigorous criteria of scientific accuracy.

EDWARD H. KASS
Channing Laboratory, Brigham and Women's Hospital, and *Harvard Medical School, Boston, MA 02115*

REFERENCES

1. W. C. Willett *et al.*, *N. Engl. J. Med.* **316**, 1175 (1987).
2. M. P. Longnecker, J. A. Berlin, M. J. Orza, T. C. Chalmers, *J. Am. Med. Assoc.* **260**, 652 (1988).
3. R. I. Horwitz and A. R. Feinstein, *N. Engl. J. Med.* **299**, 1087 (1978).
4. R. I. Horwitz, M. J. McFarlane, T. A. Brennan, A. R. Feinstein, *Arch. Int. Med.* **145**, 909 (1985).
5. R. I. Horwitz, A. R. Feinstein, M. R. Harvey, *ibid.* **144**, 1257 (1984).

With respect to reserpine and breast cancer, Feinstein states that "in a retrospective attempt to explain the error, one of the original investigators (S.S.) said that the first reserpine-breast cancer association was probably a "statistically significant" artifact due to multiple calculations done during data dredging" (*1*). (i) The phrase "data dredging" is pejorative; I did not use it. (ii) I referred to multiple comparisons, not multiple calculations—a very different matter, both conceptually and procedurally. (iii) I did not say, or imply, that the finding was an "artifact": I presented preliminary data, since published in full (*2*), and presumably known to Feinstein (*3*), showing that the original positive association was likely to be due to chance, "statistical significance" notwithstanding. (iv) Feinstein does not mention several principles that I and others (*1*) suggested should be applied to multiple comparisons, including, in particular, that no association should be taken seriously unless it can be replicated repeatedly, independently, by means of other methods, and in greater detail.

The association between alcohol and breast cancer was not *found* in the studies (*4*) to which Feinstein refers, but was *confirmed* in them. Earlier, Williams and Horm (*5*) observed an association (and a dose-response relation) with lifetime alcohol con-

sumption; Adelstein and White (*6*) observed an excess of breast cancer mortality among female alcoholics; our group (7) again observed the association independently, and then replicated it in additional data. After making detailed analyses, we published "the hypothesis that alcohol consumption, or related dietary factors, increases the risk of breast cancer" (7). All of the early observations were based on multiple comparisons. Subsequently several studies have been published (*8*), some positive and some null. Whether or not alcohol increases the risk has by no means been established, and none of the investigators has suggested otherwise.

Feinstein is free to express any opinion he wishes, but I question his freedom to back his opinion by distortion or selective citation. The reader should note that, while he now says that the association between smoking and lung cancer is a "splendid achievement" of epidemiologic research, he at one time wrote that it should be regarded with suspicion and suggested that it might be accounted for by cough leading to preferential diagnosis among smokers, or by psychic stress (*9*).

SAMUEL SHAPIRO
Slone Epidemiology Unit,
Boston University School of Medicine,
1371 Beacon Street,
Brookline, MA 02146

REFERENCES

1. S. Shapiro and D. Slone, *J. Chron. Dis.* **32**, 105 (1979).
2. S. Shapiro *et al.*, *Eur. J. Clin. Pharmacol.* **26**, 143 (1984).
3. S. Shapiro, *J. Chron. Dis.* **38**, 365 (1985).
4. W. C. Willet *et al.*, *N. Engl. J. Med.* **316**, 1174 (1987); A. Schatzkin *et al.*, *ibid.*, p. 1169.
5. R. R. Williams and J. W. Horm, *J. Natl. Cancer Inst.* **58**, 525 (1977).
6. A. Adelstein and G. White, *Popul. Trends.* **6**, 7 (1976).
7. L. Rosenberg *et al.*, *Lancet* **1**, 267 (1982).
8. M. P. Longnecker *et al.*, *J. Am. Med. Assoc.* **260**, 652 (1988).
9. A. Feinstein, *Clin. Pharmacol. Ther.* **14**, 291 (1973).

Response: I am grateful to my respected academic colleagues for making public a set of views that will be enlightening to contemporary nonepidemiologic scientists and perhaps to future historians.

The statements in these and in several unpublished letters are reminiscent of the response offered by members of the medical "establishment" in the mid-19th century (*1*), when Ignaz Semmelweis suggested that the unclean hands of doctors were sometimes giving women fatal infections (puerperal sepsis) after childbirth. The defenders of the status quo attacked Semmelweis for not emphasizing all the successful harm-free deliveries and denounced his scholarship as untrustworthy and perhaps mentally deranged; but they made no acknowledgement of the dirty-hands problem and of his plea for cleanliness.

The current letters are analogous to those responses. Some of the commentators took me to task for not giving suitable credit to the many things epidemiologists have successfully accomplished despite the faulty scientific methods. The published comments refer to my intellectual infirmities in allegedly distorting and inadequately reviewing the literature, and even doubting certain dogmas now regarded as established wisdom. And none of the comments acknowledges or calls for repairing any of the cited flaws in scientific methods.

As for my scholarly malefactions, let me immediately assure Shapiro that I am familiar with the "multiple comparison" problem in "statistical significance"; I used the word "calculations" because I thought it would be easier for nonstatisticians to understand. The term "artifact" seemed appropriate for a spurious finding that arose, as Shapiro says, by chance. I did not quote him as using the words "data dredging," and I am sorry he thinks it is a pejorative term for data dredging. Shapiro's principles for allowing an "association [to] be taken seriously" do not seem to have been applied to explain the sources of error in the two concomitantly published studies that "replicated" his original fallacious conclusion.

I assure Kass that I did not intend to demean either his investigative colleagues or their work. In an era when almost any feature of modern life has been accused of causing almost any selected disease, investigators having enough data can readily examine hundreds of hypotheses. Since tests of more than 13 different hypotheses have now been reported from that single project, I doubt that each of the individual hypotheses was specifically identified in the original research protocol. If so, I wonder how the investigators planned to deal statistically with the multiple comparison problem.

As for my heresies, they arise because my colleagues and I have given careful thought to the problems produced by absent or low scientific standards in epidemiologic studies of cause-effect relationships. We have developed new methods, using improved standards, that have been applied in our own research. Thus, we have now shown (*2*) that lung cancer is indeed underdiagnosed in noncoughers and nonsmokers. My remark about psychic stress, which Shapiro appears to have misunderstood, was intended to refer not to lung cancer, but to coronary disease, which has been inadequately investigated for the role of certain forms of psychic stress in possibly causing both smoking and coronary disease. Kass writes that the causal relationships between "exogenous estrogens and endometrial cancer" and between "diethylstilbestrol and vaginal carcinoma . . . are now beyond doubt." My colleagues and I disagree; and we have recently (*3*) reviewed the evidence, stated the reasons for our disagreements, and indicated why the problems will not be solved without new studies, using better methods. As for the aspirin–Reyes syndrome relationship, we have now carried out a study (*4*) using improved scientific standards. In this instance, we confirmed the original statistical association.

The most remarkable feature of the letters to *Science* is the absence of concern for the fundamental scientific defects I cited in epidemiologic methods. After completing their attacks, the critics do not seem upset by investigators making changes in control groups after the results have been analyzed, by large numbers of studies with unresolved and unreconciled contradictions, by the infrequent precautions against ascertainment bias, by statistical maneuvers that are substituted for a true dose-response curve, or by the credulous acceptance of erroneous death-certificate diagnoses. No one seems troubled by the persistent scientific neglect of detection bias, which may also be responsible for yet another recently publicized "menace" of daily life: the relationship of breast cancer and oral contraceptive agents.

I hope my colleagues will forgive me, but I could not have asked for better illustrations of the type of scientific complacency I lamented.

ALVAN R. FEINSTEIN
Clinical Epidemiology Unit,
Yale University School of Medicine,
New Haven, CT 06510–8025

REFERENCES

1. S. B. Nuland, *J. Hist. Med. Allied Sci.* **34**, 255 (1979).
2. M. J. McFarlane *et al.*, *Arch. Intern. Med.* **146**, 1695 (1986); C. K. Wells *et al.*, *Am. J. Epidemiol.* **128**, 1016 (1988).
3. R. I. Horwitz *et al.*, *Am. J. Med.* **81**, 503 (1986); M. J. McFarlane *et al.*, *ibid.*, p. 855.
4. B. W. Forsyth *et al.*, *J. Am. Med. Assoc.*, in press.

Erratum: In line 7 of the caption for table 1 (p. 1682) of the report "Association of transfer RNA acceptor identity with a helical irregularity" by William H. McClain *et al.* (23 Dec., p. 1679), "≥20%" should have been "≤20%."