Letters

Radon Hazard

Richard A. Kerr reports (News & Comment, 23 Sept., p. 1594) on the Environmental Protection Agency's (EPA's) announcement of their latest radon survey in an article entitled "Radon survey seen as misleading by some scientists."

Both Lee Thomas, Administrator, EPA, and I emphasized at our press conference and in subsequent discussions with the media that EPA was reporting screening tests done usually in basements during the winter when the readings were likely to be highest. We both emphasized that before individual risks could be estimated and any appropriate action taken, additional tests in the living areas of homes were necessary. While the screening tests cannot estimate individual risk, I do not believe they are "nearly useless." With screening tests that are quick and inexpensive (72 hours and about $10), individuals can determine whether or not their homes have a potential problem. If not, as is the case with the majority, the concern can be dismissed. If the screening test is positive, then more definitive tests of longer duration that cost somewhat more (3 months to 1 year and $30 to $75) in the living area are necessary.

In addition to emphasizing the screening nature of the tests to the media, I further emphasized the synergistic relation between smoking and indoor radon risk. I urged people not to smoke and recommended that, if their home had radon in it, they prohibit smoking inside their home. This was generally picked up by the news reports, but there was no reference to it in Kerr's article.

Using data we have for humans and animals—epidemiologic studies of more than 40,000 miners, including both smokers and nonsmokers, with what comes close to providing dose-response relationships, and applying sound scientific judgment (not just rote mathematics), can estimate, although not precisely determine, the risk from radon. In most other environmental areas we have nothing close to the quality and quantity of data that exist for radon. If we cannot use these data to recommend that individuals take preventive measures—lowering the radon in their homes where it is elevated, not smoking, and prohibiting smoking in their home—where are we in estimating other environmental risks for preventive action?

In my opinion, to equate the risk from radon to that of death from falls and fires in the home at a 0.4% lifetime risk appears to trivialize the radon risk. In fact, annually nearly 6000 people die from falls in the home and 5000 die from home fires. The number of deaths from falls and fires is not inconsequential and, just like deaths from indoor radon exposure, they are preventable and not inevitable.

VERNON N. HOUK
Assistant Surgeon General, Director,
Center for Environmental Health
and Injury Control,
Centers for Disease Control,
Atlanta, GA 30333

Response: I did not state that radon screening tests are "nearly useless." I did report that "such screening surveys are nearly useless in determining the prevalence of radon health hazards." Screening surveys have their purposes, as is made clear in the article and in Houk's letter; but determining the magnitude of the national radon problem is not one of them. Unfortunately, perhaps despite Houk's best intentions, the media took the survey results as new proof of the pervasive threat of radon. Clarifying this misconception was the sole point of my article.

—RICHARD A. KERR

Methods and Molecules

Roger Lewin (Research News, 23 Sept., p. 1598) is to be applauded for calling the DNA controversy to the attention of the scientific community, but he does not make clear the position of phylogenetic systematists (cladists). Cladists hold that taxa should be united by synapomorphies, shared derived traits, rather than by raw (average, general, overall) similarities. A synapomorphy could just as well be molecular as morphological, and many molecular systematists follow a cladist approach. Sibley and his co-workers have routinely used UPGMA, a phenetic clustering method, to produce trees from their data. Cladists would object to UPGMA even if it were used on morphological information.

The choice of grouping method can easily affect results. While Sibley and Ahlquist concluded (*1*) that their data definitively established the placement of humans with chimps, one of us (J.S.F.) found (*2*) that grouping humans with gorillas gave better fit to those data. The difference in goodness of fit may not be significant, but then those data do not settle the placement of humans.

In their numerous publications in bird phylogeny, Sibley and Ahlquist give no indication of having investigated whether other groupings might fit their data better than those produced by UPGMA, nor do they provide the information necessary for others to do so. How well their conclusions are justified will not be known until their data are more widely available.

JAMES S. FARRIS
Department of Ecology and Evolution,
State University of New York,
Stony Brook, NY 11794
ARNOLD G. KLUGE
Museum of Zoology and
Department of Biology,
University of Michigan,
Ann Arbor, MI 48109

REFERENCES

1. C. G. Sibley *et al.*, *J. Mol. Evol.* **20**, 2 (1984).
2. J. S. Farris, *Cladistics* **1**, 67 (1985).

Psychiatric Diagnosis

In their article "The expert witness in psychology and psychiatry" (1 July, p. 31), David Faust and Jay Ziskin discuss the many problems involved in having psychologists and psychiatrists present "expert testimony" in court. We take exception to several statements in their article and in their subsequent response (Letters, 2 Sept., p. 1143) to correspondence about recent efforts to improve the classification of mental disorders and the reliability with which psychiatric diagnoses can be made. Referring to the recent revisions of the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (DSM-III in 1980 and DSM-III-R in 1987), they say that "This process of revision little resembles the refinement of categories or cumulative gains common to advanced scientific fields." Unfortunately, this remark is not followed by a discussion of what they believe is necessary before revisions in a general purpose classification system in medicine is appropriate.

The reader is not told that, with the advancement of DSM-III, organized psychiatry, for the first time, faced up to the long-standing problem of diagnostic unreliability and devoted significant resources to a process that would improve diagnostic agreement. Despite the critique of methodologic flaws in the DSM-III field trials, which have been addressed elsewhere (*1*), no one can reasonably deny that diagnostic reliability in psychiatry has been significantly improved by the innovation of specified diagnostic criteria in DSM-III and DSM-III-R.

Diagnostic unreliability in psychiatry continues to be a problem, but the reader should realize that the rest of medicine also has problems with reliability. For example, Feinstein (*2*) has noted that the reliability for abnormal opthalmologic findings is poor

(Kappa = .33) and that for reading exercise cardiograms for potential bypass surgery equally poor (Kappa = .30). These are not isolated findings in other fields of medicine.

Unfortunately, diagnosis in psychiatry, and in the rest of medicine, depends on complex clinical variables. There is no "quick fix" to completely eliminate variability in clinicians' judgements. The best that can be done is to face the problem squarely, as psychiatry is now doing.

ROBERT L. SPITZER*
JANET B. W. WILLIAMS†
Department of Psychiatry,
College of Physicians and Surgeons,
Columbia University,
New York, NY 10032
HAROLD ALAN PINCUS
Deputy Medical Director,
Director, Office of Research,
American Psychiatric Association,
1400 K Street, NW,
Washington, DC 20005

REFERENCES

1. J. B. W. Williams and R. L. Spitzer, *Harvard Med. School Mental Health Lett.*, in press.
2. A. R. Feinstein, *Clinical Epidemiology: The Architecture of Clinical Research* (Saunders, Philadelphia, PA, 1985).

*Chair, committees that developed DSM-III and DSM-III-R.
†Text Editor, DSM-III and DSM-III-R.

Response: We agree with Spitzer *et al.* that diagnostic unreliability in psychiatry is longstanding, organized psychiatry is devoting considerable effort to the problem for the first time, there is no quick fix, reliability remains problematic, and it is best to face these problems squarely. More expert witnesses might follow Spitzer *et al.*'s lead by freely admitting these points and thus, too, facing the problem squarely.

Recommendations for revision had no bearing on the main topic of our article, the current status of the expert witness in psychology and psychiatry. The reader might consult an article by one of us (D.F.) (*1*) which details one major requirement for adequate revision—scientific advance that provides the needed knowledge base. We also offer a specific recommendation here: For most specific categories in DSM-III and throughout DSM-III-R, formal reliability trials were not conducted and reported before revisions were introduced. This should be done the next time.

Grove's comparison (*2*) of numerous reliability studies along a common metric (Kappa) raises serious doubts about claims for significant improvement. Where data were available across systems and common categories, the DSM-III showed only a slight overall advantage over earlier systems; whether DSM-III-R fares differently awaits adequate study. Even if one posits significant improvement, according to Spitzer *et al.* reliability remains problematic, and thus it has not improved enough for use in the legal forum. For example, the heterogeneity of general diagnostic categories often makes time of little or no possible use in legal determinations. For specific categories, the citations in our article and book (*3*), and even the DSM-III field trials, back our statement that "a number of . . . studies showed that rate of disagreement . . . often equals or exceeds rate of agreement." Finally, citing studies that demonstrate problems with medical diagnosis does not lessen problems with psychiatric diagnosis.

Spitzer *et al.* do not mention some of the more pointed criticisms of DMS-III and DSM-III-R raised in the literature, for example, that the list of disorders exceeds reasonable limits (incoordination characterized by such things as poor handwriting is included in the list of mental disorders) and that the diagnostic categories were shaped as much or more by political processes and personal values as by science (many categories were "established" by committee vote).

Most important, diagnostic reliability is but one piece of a far larger problem—the validity of psychologists' and psychiatrists' opinions on legal matters. A higher Kappa does not establish, for example, that a diagnosis can aid in determining mental state at the time of a crime; the DSM-III-R itself raises cautions about application to legal issues. Perhaps when organized psychiatry completes its noteworthy struggle with reliability, it might turn its attention to developing the scientific knowledge needed to answer legal questions with sufficient reliability *and* validity.

DAVID FAUST
Department of Psychiatry and Pediatrics,
Rhode Island Hospital,
593 Eddy Street,
Providence, RI 02903, and
Department of Psychiatry and
Human Behavior,
Brown University,
Program in Medicine,
Providence, RI 02903
JAY ZISKIN
Post Office Box 9489,
Marina del Rey, CA 92095

REFERENCES

1. D. Faust, *Am. J. Psychiatr.* **143**, 962 (1986).
2. W. M. Grove, in *Issues in Diagnostic Research*, C. G. Last and M. Hersen, Eds. (Plenum, New York, 1987), pp. 99–119.
3. J. Ziskin and D. Faust, *Coping with Psychiatric and Psychological Testimony* (Law and Psychology Press, Venice, CA, 1988).

Malformations in Chicken Embryos in the Northeast

Recent observations of chick embryos from hatcheries in Pennsylvania, Maryland, and Connecticut raise the possibility that something has been wrong with the eggs and embryos in these areas.

From early February until early October the following observations were made of chick embryos at two early periods of development—after 24 to 30 and 50 to 60 hours, respectively, of incubation. The younger embryos had a high incidence of blisters beneath the epiblast in the area pellucida. The later occurring, more prevalent, anomaly was observed in the tissue that forms somites, the segmental plate. This tissue began to break down after 50 hours of incubation, and for the next 10 hours the site that should have been occupied by the segmental plate appeared as a large, irregular blister. In some flocks the incidence of these segmental plate blisters was as high as 95 to 100%. Surprisingly, most embryos appeared normal 10 hours later, with the usual somitic array. It seemed as if the segmental plate or somites, or both, were reformed by restorative or catch-up growth. All embryos, however, were not normal. A high incidence of visible malformations was seen, including spina bifida, anencephaly, and herniated cardiac coelom.

We examined hundreds of embryos from ten different flocks from five hatcheries. No flock or hatchery was without anomalies. The problem has been serious in 1988, but the same observations were made in 1985. At this earlier time the incidence of segmental plate blisters was much lower and did not interfere with our research.

We would like to alert researchers that these anomalies may be widespread geographically and may have been interfering with results of experiments performed on young embryos. The anomalies were coincident with, but not necessarily related to, reports that chicken eggs in the northeastern states appear to be carriers of *Salmonella enteritidis*. Tests performed on the embryos by Robert Eckroade at the New Bolton Center of the University of Pennsylvania were negative for *S. enteritidis*.

JAMES W. LASH
Department of Anatomy,
School of Medicine,
University of Pennsylvania,
Philadelphia, PA 19104–6058

Erratum: In Deborah Barnes' article "Joint Soviet–U.S. attack on heart muscle dogma" (Research News, 14 Oct., page 193), credit for a photograph was inadvertently omitted. John Oberpriller of the University of North Dakota graciously supplied the photograph of a newt ventricular myocyte dividing in culture.