

# Epidemiology Faces Its Limits

The search for subtle links between diet, lifestyle, or environmental factors and disease is an unending source of fear—but often yields little certainty

The news about health risks comes thick and fast these days, and it seems almost constitutionally contradictory. In January of last year, for instance, a Swedish study found a significant association between residential radon exposure and lung cancer. A Canadian study did not. Three months later, it was pesticide residues. The *Journal of the National Cancer Institute* published a study in April reporting—contrary to previous, less powerful studies—that the presence of DDT metabolites in the bloodstream seemed to have no effect on the risk of breast cancer. In October, it was abortions and breast cancer. Maybe yes. Maybe no. In January of this year it was electromagnetic fields (EMF) from power lines. This time a study of electric utility workers in the United States suggested a possible link between EMF and brain cancer but—contrary to a study a year ago in Canada and France—no link between EMF and leukemia.

These are not isolated examples of the conflicting nature of epidemiologic studies; they're just the latest to hit the newspapers. Over the years, such studies have come up with a mind-numbing array of potential disease-causing agents, from hair dyes (lymphomas, myelomas, and leukemia) to coffee (pancreatic cancer and heart disease) to oral contraceptives and other hormone treatments (virtually every disorder known to woman). The pendulum swings back and forth, subjecting the public to an "epidemic of anxiety," as Lewis Thomas put it over a decade ago. Indeed, last July, the *New England Journal of Medicine* (NEJM) published an editorial by editors Marcia Angell and Jerome Kassirer asking the pithy question, "What Should the Public Believe?" Health-conscious Americans, wrote Angell and Kassirer, "increasingly find themselves beset by contradictory advice. No sooner do they learn the results of one research study than they hear of one with the opposite message."

Kassirer and Angell place responsibility



Anxiety epidemic. Protesting risks that may—or may not—be real.

noninfectious diskases—smoking, for instance, which can increase the risk of developing lung cancer by as much as 3000%. Now they are left to search for subtler links between diseases and environmental causes or lifestyles. And that leads to the Catch-22 of modern epidemiology.

On the one hand, these subtle risks—say, the 30% increase in the risk of breast cancer from alcohol consumption that some studies suggest—may affect such a large segment of the population that they have potentially huge impacts on public health. On the other, many epidemiologists concede that their studies are so plagued with biases, uncertainties, and methodological weaknesses that they may be inherently incapable of accurately discerning such weak associations. As Michael Thun, the director of analytic epidemiology for the American Cancer Society, puts it, "With epidemiology you can tell a little thing from a big thing. What's very hard to do is to tell a little thing from nothing at all." Agrees Ken

on the press for its reporting of epidemiology, and even on the public "for its unrealistic expectations" of what modern medical research can do for their health. But many epidemiologists interviewed by *Science* say the problem also lies with the very nature of epidemiologic studies—in particular those that try to isolate causes of noninfectious disease, known variously as "observational" or "risk-factor" or "environmental" epidemiology.

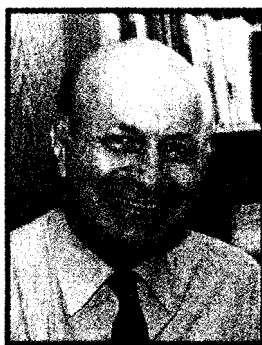
The predicament of these studies is a simple one: Over the past 50 years, epidemiologists have succeeded in identifying the more conspicuous determinants of

Rothman, editor of the journal *Epidemiology*: "We're pushing the edge of what can be done with epidemiology."

With epidemiology stretched to its limits or beyond, says Dimitrios Trichopoulos, head of the epidemiology department at the Harvard School of Public Health, studies will inevitably generate false positive and false negative results "with disturbing frequency." Most epidemiologists are aware of the problem, he adds, "and tend to avoid causal inferences on the basis of isolated studies or even groups of studies in the absence of compelling biomedical evidence. However, exceptions do occur, and their frequency appears to be increasing." As Trichopoulos explains, "Objectively the problems are not more than they used to be, but the pressure is greater on the profession, and the number who practice it is greater."

As a result, journals today are full of studies suggesting that a little risk is not nothing at all. The findings are often touted in press releases by the journals that publish them or by the researchers' institutions, and newspapers and other media often report the claims uncritically (see box on p. 166). And so the anxiety pendulum swings at an ever more dizzying rate. "We are fast becoming a nuisance to society," says Trichopoulos.

"People don't take us seriously anymore, and when they do take us seriously, we may unintentionally do more harm than good." As a solution, epidemiologists interviewed by *Science* could suggest only that the press become more skeptical of epidemiologic findings, that epidemiologists become more skeptical about their own findings—or both.



**"People don't take us seriously ... and when they do ... we may unintentionally do more harm than good."**

—Dimitrios Trichopoulos

## Sizing Up the Cancer Risks

In the history of epidemiology, only a dozen or so environmental agents have ever been repeatedly and strongly linked to human cancer, says University of Alabama epidemiologist Philip Cole. Among them are cigarette smoke, alcohol, ionizing radiation, a few drugs, a handful of occupational carcinogens, such as asbestos, and perhaps three viruses— hepatitis-R virus, human T cell leukemia virus, and human papillomavirus. But every year, epidemiologic papers are published by the journal-load, many of **them** reporting new potential causes of cancer in the environment.

Most are the product of observational epidemiology, in which researchers try to compare the lives of people suffering from a disease with those of healthy controls. Even **its** practitioners admit this effort is plagued by biases and confounding factors (see main text). As a result, most epidemiologists interviewed by *Science* said they would not take seriously a *single* study reporting a new potential cause of cancer unless it reported that exposure to the agent in question increased a person's risk by at least a factor of 3—which is to say it carries a risk ratio of 3. Even then, they say, skepticism is in order unless the study was very large and extremely well done and biological data support the hypothesized link. Sander Greenland, a university of California, Los Angeles, epidemiologist, says a study reporting a twofold increased risk might then be worth taking seriously—“but not that seriously.”

Few of the entries in the following list of potential cancer risks, reported in the journals and picked up in the popular press over the past 8 years, have come close to fulfilling those criteria. Are these dangers real? As the saying goes, you be the judge.

—G.T.

High-cholesterol diet—risk ratio (rr) 1.65 for rectal cancer in men (January 1987)

Eating yogurt at least once a **month**—rr 2 for ovarian cancer (July 1989)

Smoking more than **100** cigarettes in a lifetime—rr 1.2 for breast cancer (February 1990)

High-fat diet—rr 2 for breast cancer (August 1990)

Lengthy occupational exposure to dioxin—rr 1.5 for all cancers (January 1991)

Douching once a week—rr 4 for cervical cancer (March 1991)

Regular use of high-alcohol mouthwash—rr 1.5 for mouth cancer (June 1991)

Use of phenoxy herbicides on lawns—rr 1.3 for malignant lymphoma in dogs (September 1991)

Weighing **3.6** kilograms or more at birth—rr 1.3 for breast cancer (October 1992)

Vasectomy—rr 1.6 for prostate cancer (February 1993)

Pesticide exposure, indicated by high residues in blood—rr 4 for breast cancer (April 1993); contradicted 1 year later in a larger study with one of the same authors.

Drinking more than **3.3** liters of fluid (particularly chlorinated tap water) a day—rr 2–4 for bladder cancer (July 1993)

Experiencing psychological stress in the workplace—rr 5.5 for colorectal cancer (September 1993)

Diet high in saturated fat—rr 6 for lung cancer in nonsmoking women (December 1993)

Eating more than 20 grams of processed meats (i.e., bologna) a day—rr 1.72 for colon cancer (February 1994)

Eating red meat five or more times a week—rr 2.5 for colon cancer (February 1994)

Occupational exposure to electromagnetic fields—rr 1.38 for breast cancer (June 1994)

Smoking two packs of cigarettes a day—rr 1.74 for fatal breast cancer (July 1994)

Eating red meat twice a day—rr 2 for breast cancer (July 1994)

Regular cigarette smoking—rr 1.7 for pancreatic cancer (October 1994)

Ever having used a sun lamp—rr 1.3 for melanoma (November 1994)

Abortion—rr 1.5 for breast cancer (November 1994)

Having shorter or longer than average menstrual cycles—rr 2 for breast cancer (December 1994)

Obesity in men (the heaviest 25% of those in the **study**)—rr 3 for esophageal cancer (January 1995)

Consuming olive oil only once a day or less—rr 1.25 for breast cancer (January 1995)

and environmental causes. And the fact that no single cancer affects every population at the same rate suggests that factors external to the human body cause 70% to 90% of all cancers. In other words, says Richard Peto, an Oxford University epidemiologist, “there are ways in which human beings can live whereby those cancers would not arise.” Only a few of these environmental factors are known—cigarette smoke for lung cancer, for example, or sunlight for skin cancer—and epidemiology seems to provide the best shot at identifying the others.

The most powerful tool for doing so is the randomized trial, which is the standard for studies of new drugs and other medical re-

search: Assign subjects at random to test and control groups, alter the exposure of the test group to the suspected risk factor, and follow both groups to learn the outcome. Often, both the experimenters and the subjects are “blinded”—unaware who is in the test group and who is a control. But randomized trials would be prohibitively slow and expensive for most risk factors, because they can take years or decades to show an effect and hundreds of thousands of individuals may need to be followed to detect enough cases of the disease for the results to be significant. And randomly subjecting thousands of healthy people to pollutants or other possible carcinogens raises obvious ethical problems.

Because the experimental approach is off-limits for much of epidemiology, researchers resort to observational approaches. In case-control studies, for example, they select a group of individuals afflicted with a particular disorder, then identify a control group free of the disorder and compare the two, looking for differences in lifestyle, diet, or some environmental factor. Potentially more reliable, but also much more costly, are cohort studies, in which researchers take a large population—as many as 100,000—and question the subjects in detail about their habits and environment. They then follow the entire population for years or decades to see who gets sick and who doesn't, what dis-

eases they suffer from, and what factors might be different between them. Either way, risk-factor epidemiology is "a much duller scalpel" than randomized trials, says Scott Zeger, a biostatistician at the Johns Hopkins School of Mental and Public Health.

What blunts its edge are systematic errors, known in the lingo as biases and confounding factors. "Bias and confounders are the plague upon the house of epidemiology," says Philip Cole, chair of epidemiology at the University of Alabama. They represent anything that might lead an epidemiologic study to come up with the wrong answer, to postulate the existence of a causal association that does not exist or vice versa.

Confounding factors are the hidden variables in the populations being studied, which can easily generate an association that may be real but is not what the epidemiologist thinks it is. A ubiquitous example is cigarette smoking, which can confound any study looking, for instance, at the effects of alcohol on cancer. "It just so happens," explains Trichopoulos, "that people who drink also tend to smoke," boosting their risk of cancer. As a result, epidemiologists face the possibility that any apparent cancer-alcohol link may be spurious. Smoking may also have confounded a study Trichopoulos himself co-authored linking coffee-drinking and pancreatic cancer—a finding that has not been replicated. The study, published over a decade ago, corrected for smoking, which often accompanies heavy coffee drinking—but only for smoking during the 5 years before the cancer was diagnosed. Trichopoulos now says that he and his colleagues might have done better to ask about smoking habits a full 20 years before diagnosis.

Biases are problems within study designs themselves. The process of choosing an appropriate population of controls in a case-control study, for instance, can easily lead to an apparent difference between cases and controls that has nothing to do with what caused the disease. "It's often not even theoretically clear who the right comparison group is," says Harvard epidemiologist Walter Willett. "And sometimes, even if you can design the study so that you have the theoretically correct comparison group, you usually don't get everybody willing to participate, and the people who do participate in your study will be different from the people who don't, often in health-related ways."

For example, Charles Poole of Boston University has spent several years analyzing the results and methodology of a 1988 study of EMF and cancer, which found that exposure to relatively high EMF from power lines appeared to increase the risk of leukemia and brain cancer in children. David Savitz of the University of North Carolina, the study's author, selected controls for that study with a common technique known as random digit

dialing: Researchers take the phone numbers of their cases and randomly change the last four digits until they find a suitable control. Random digit dialing, however, seems to create "a pronounced bias toward the control group being deficient in persons of very low socioeconomic status," says Poole. Poor people, it seems, are either less likely to be home during the day to answer the phone, less likely to want to take part in a study, or less likely to have an answering machine and call the researchers back.

Indeed, the North Carolina researchers reported that their data showed that the risk of leukemia and brain cancer rises not just with exposure to EMF but also with higher levels of breast-feeding, maternal smoking, and traffic density, all of which are markers for poverty. This suggests, says Poole, that the study group was poorer than the controls, and that some poverty-associated factor other than EMF could have resulted in the apparent increase in cancer risk. Nonetheless, the study is still cited as supporting the hypothesis that EMF causes childhood cancer, although even Savitz concedes that the random digit dialing problem is "a legitimate source of uncertainty."

Even when such biases can be identified, their magnitude—and sometimes even their direction—can be nearly impossible to assess. David Thomas, for example, an epidemiologist at the Fred Hutchinson Cancer Research Center in Seattle, points to studies analyzing the effect of Breast Self-Examination (BSE) on breast cancer mortality rates, which, he says, have yielded some "modest suggestion that there might be a beneficial effect" from BSE. "You have to ask what motivates a woman to practice BSE," says Thomas. "Maybe she has a strong family history of breast cancer. If so, she's more likely to get breast cancer. That would be an obvious bias," which could make BSE look less useful than it is. "Or maybe a woman with a strong family history of breast cancer would be afraid to practice BSE. You have no way of predicting the direction of the bias. So it would be very difficult to interpret your results. You have to go to a randomized study to get a reliable answer."

#### Tricks of memory

Of all the biases that plague the epidemiologic study of risk factors, the most pernicious is the difficulty of assessing exposure to a particular risk factor. Rothman, for instance, calls it "a towering obstacle." When expo-

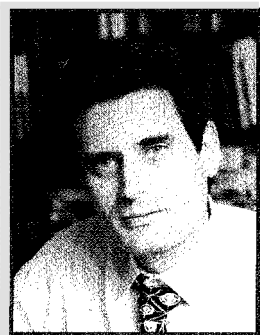
sure can be measured reliably, a subtle association may be credible—as it is in the case of early childbirth and a lower risk of breast cancer. The reason is that both cause and effect can be measured with some certainty, says Harvard epidemiologist Jamie Robins. "It's easy to know which people got breast cancer, and it's easy to know at what age they had kids," he says, adding that virtually every study on the subject comes to the same conclusion: Early childbirth reduces the risk by about 30%.

But epidemiologists are quick to list risk factors for which accurate exposure measurements are virtually impossible. Joe Fraumeni, director of the epidemiology and biostatistics program at the National Cancer Institute (NCI), points to radon: "When you're studying smoking," he says, "that's easy. Just count the number of cigarettes and duration and packs per day. But something like radon, how do you measure exposure, particularly biologically relevant exposure that has taken place in the past?" Equally uncertain are those risk factors recorded only in human memory, such as consumption of coffee or dietary fat. Ross Prentice of the University of Washington notes, for example, that underweight individuals

tend to overreport fat intake on questionnaires or in interviews and obese subjects tend to underreport it.

Such recall bias is known to be especially strong, as Willett points out, among patients diagnosed with the disease in question or among their next of kin. In studies of a possible relationship between fat intake and breast cancer, for instance, says Willett, "people may recall their past intake of fat differently if they have just been diagnosed with breast cancer than if you pluck them out of a random sample, call them up out of the blue over the phone, and ask them what their past diet was."

Recall bias, for instance, apparently accounts for the conflicting findings about oral contraceptive use and breast cancer. Many studies have looked for this association over the years, both case-control studies and cohort studies. Trichopoulos notes that case-control studies have tended to show an association between oral contraceptives and breast cancer, while cohort studies have not. Epidemiologists who have done cohort studies say the problem is in case-control studies, which are thrown off by recall bias—women who are diagnosed with breast cancer are more likely to give complete information



MARINA KENNEDY

**"We're pushing the edge of what can be done with epidemiology."**

—Ken Rothman

about contraceptive use than women who don't. Those who did case-control studies say the bias is in the cohort studies. Cohort studies have to rely on impersonal questionnaires because they are so much larger than case-control studies, and women are less likely to give complete and honest information than they are in the more intimate interviews possible in case-control studies. "The point," says Trichopoulos, "is which do we believe."

It's not just the subjects of studies who are prone to bias; epidemiologic studies can be plagued by interviewer bias as well. The interviewers are rarely blinded to cases and controls, after all, and questionnaires, the traditional measuring instrument of epidemiology, are neither peer-reviewed nor published with the eventual papers. "In the laboratory," as Yale University clinical epidemiologist Alvin Feinstein puts it, "you have all kinds of procedures for calibrating equipment and standardizing measurement procedures. In epidemiology ... it's all immensely prey to both the vicissitudes of human memory and the biases of the interview."

### Salvation from statistics?

With confounders, biases, and measurement errors virtually inevitable, many epidemiologists interviewed by *Science* say that risk-factor epidemiology is increasingly straying beyond the limits of the possible no matter how carefully the studies are done. "I have trouble imagining a system involving a human habit over a prolonged period of time that could give you reliable estimates of [risk] increases that are of the order of tens of percent," says Harvard epidemiologist Alex Walker. Even the sophisticated statistical techniques that have entered epidemiologic research over the past 20 years—tools for teasing out subtle effects, calculating the theoretical effect of biases, correcting for possible confounders, and so on—can't compensate for the limitations of the data, says biostatistician Norman Breslow of the University of Washington, Seattle.

"In the past 30 years," he says, "the methodology has changed a lot. Today people are doing much more in the way of mathematical modeling of the results of their study, fitting of regression equations, regression analysis. But the question remains: What is the fundamental quality of the data, and to what extent are there biases in the data that cannot be controlled by statistical analysis? One of the dangers of having all these fancy

mathematical techniques is people will think they have been able to control for things that are inherently not controllable."

Breslow adds that epidemiologists will commonly report that they have unveiled a possible causal association between a risk factor and a disease because the association is "statistically significant," meaning that the error bars—the limits of a 95% confidence interval—do not include the null result, which is the absence of an effect. But, as Breslow explains, such statistical "confidence" means considerably less than it seems to. The calculation of confidence limits only takes into consideration random variation in the data. It ignores the systematic errors, the biases and confounders, that will almost invariably overwhelm the statistical variation.

University of California, Los Angeles (UCLA) epidemiologist Sander Greenland says most of his colleagues fail to understand this simple point. "What people want to do when they see a 95% confidence [interval]," he says, "is say 'I bet there's a 95% chance the true value is in there.' Even if they deny it, you see them behaving and discussing their study result as though that's exactly what it means. There are certain conditions under which it's not far from the truth, but those conditions are generally not satisfied in an epidemiologic study."



UNIVERSITY OF WASHINGTON

**"People [may] think they have been able to control for things that are inherently not controllable."**

—Norman Breslow



KIMBERLY GRANT

**"Authors and investigators are worried that there's a bias against negative studies."**

—Marcia Angell

co-authored a study erroneously suggesting that women who took the anti-hypertension medication reserpine had up to a fourfold increase in their risk of breast cancer, suggests that no single epidemiologic study is persuasive by itself unless the lower limit of its 95% confidence level falls above a threefold increased risk. Other researchers, such as Harvard's Trichopoulos, opt for a fourfold risk increase as the lower limit. Trichopoulos's ill-fated paper on coffee consumption and pancreatic cancer had reported a 2.5-fold increased risk.

"As a general rule of thumb," says Angell of the *New England Journal*, "we are looking for a relative risk of three or more [before accept-

ing a paper for publication], particularly if it is biologically implausible or if it's a brand-new finding." Robert Temple, director of drug evaluation at the Food and Drug Administration, puts it bluntly: "My basic rule is if the relative risk isn't at least three or four, forget it." But as John Bailar, an epidemiologist at McGill University and former statistical consultant for the *NEJM*, points out, there is no reliable way of identifying the dividing line. "If you see a 10-fold relative risk and it's replicated and it's a good study with biological backup, like we have with cigarettes and lung cancer, you can draw a strong inference," he says. "If it's a 1.5 relative risk, and it's only one study and even a very good one, you scratch your chin and say maybe."

Some epidemiologists say that an association with an increased risk of tens of percent might be believed if it shows up consistently in many different studies. That's the rationale for meta-analysis—a technique for combining many ambiguous studies to see whether they tend in the same direction (*Science*, 3 August 1990, p. 476). But when *Science* asked epidemiologists to identify weak associations that are now considered convincing because they show up repeatedly, opinions were divided—consistently.

Take the question of alcohol and breast cancer. More than 50 studies have been done, and more than 30 have reported that women who drink alcohol have a 50% increased risk of breast cancer. Willett, whose Nurse's Health Study was among those that showed a positive association, calls it "highly probable" that alcohol increases the risk of breast cancer. Among other compelling factors, he says, the finding has been "reproduced in many countries with many investigators controlling for lots of confounding

variables, and the association keeps coming up." But Greenland isn't so sure. "I'd bet right now there isn't a consensus. I do know just from talking to people that some hold it's a risk factor and others deny it." Another Boston-based epidemiologist, who prefers to remain anonymous, says nobody is convinced of the breast cancer-alcohol connection "except Walt Willett."

Another example is long-term oral contraceptive use and breast cancer, a link that has been studied for a quarter of a century. Thomas of the Fred Hutchinson Cancer Research Center says he did a meta-analysis in 1991 and found a dozen studies showing a believable association in younger women who were long-time users of oral contraceptives. "The bottom line," he says, "is it's taken us over 20 years of studies before some consistency starts to emerge. Now it's fairly clear there's a modest risk." But Noel Weiss of the University of Washington says he did a similar review of the data that left him unconvinced. "We don't know yet," he says. "There is a small increased risk associated [with oral contraceptive use], but what that represents is unclear." Mary Charleson, a Cornell Medical Center epidemiologist, calls the association "questionable." Marcia Angell calls it "still controversial."

Consistency has a catch, after all, explains David Sackett of Oxford University: It is persuasive only if the studies use different architectures, methodologies, and subject groups and still come up with the same results. If the studies have the same design and "if there's an inherent bias," he explains, "it wouldn't make any difference how many times it's replicated. Bias times 12 is still bias." What's more, the epidemiologists interviewed by *Science* point out that an apparently consistent body of published reports showing a positive association between a risk factor and a disease may leave out other, negative findings that never saw the light of day.

"Authors and investigators are worried that there's a bias against negative studies," and that they will not be able to get them published in the better journals, if at all, says Angell of the *NEJM*. "And so they'll try very hard to convert what is essentially a negative study into a positive study by hanging on to very, very small risks or seizing on one positive aspect of a study that is by and large negative." Or, as one National Institute of Environmental Health Sciences researcher puts it, asking for anonymity, "Investigators who find an effect get support, and investigators who don't find an effect don't get support. When times are tough it becomes extremely difficult for investigators to be objective."

When asked why they so willingly publish inconclusive research, epidemiologists say they have an obligation to make the data public and justify the years of work. They also argue that if the link is real, the public health

effect may be so dramatic that it would be irresponsible not to publish it. The University of North Carolina's Savitz, for instance, who recently claimed a possible link between EMF exposure and a tens of percent increase in the risk of breast cancer, says: "This is minute. ... But you could make an argument that even if this evidence is 1000-fold less than for [an EMF-leukemia link], it is still more important, because the disease is 1000-fold more prevalent."

One of the more pervasive arguments for publishing weak effects, Rothman adds, is that any real effect may be stronger than the reported one. Any mismeasurement of exposure, so the argument goes, will only serve to reduce the observed size of the association. Once researchers learn how to measure exposure correctly, in other words, the actual association will turn out to be bigger—and thus more critical to public health. That was the case in studies of steelworkers and lung cancer decades ago, says Robins. Early studies saw only a weak association, but once researchers homed in on coke-oven workers, the group most exposed to the carcinogens, the relative risk shot up. None of the epidemiologists who spoke to *Science* could recall any more recent parallels, however.

#### An unholy alliance

There would be few drawbacks to publishing weak, uncertain associations if epidemiologists operated in a vacuum, wrote Brian MacMahon, professor emeritus of epidemiology at Harvard, in an April 1994 editorial in the *Journal of the National Cancer Institute*. But they do not, he said. "And, however cautiously the investigator may report his conclusions and stress the need for further evaluation," he added, "much of the press will pay little heed to such cautions. ... By the time the information reaches the public mind, via print or screen, the tentative suggestion is likely to be interpreted as a fact."

This is what one epidemiologist calls the "unholy alliance" between epidemiology, the journals, and the lay press. The first one or two papers about a suspected association "spring into the general public consciousness in way that does not happen in any other field of scientific endeavor," says Harvard's Walker. And once a possible link is in the public eye, it can be virtually impossible to discredit. As far as scientists were concerned, for instance, a 1981 epidemiologic study put to rest a suggestion that saccharine can cause

bladder cancer—one of the few cases in which epidemiology had managed to put an end to a suspected association. Yet 14 years later, television advertisements for Nutra-Sweet, which contains the artificial sweetener aspartame, still tout it as the sweetener that does not have saccharine.

Epidemiologists themselves are at a loss as to how to curb the "anxiety of the week" syndrome. Many, like Rothman, simply argue that risk factor epidemiology is a young science that will take time to mature. Others, like Robins, suggest that barring a major breakthrough in the methodological tools of epidemiology, maturity will be hard to come by. The pressures to publish inconclusive results and the eagerness of the press to publicize them, he and others say, mean that the anxiety pendulum, like Foucault's, will continue to swing indefinitely (see box on p. 165).

The FDA's Temple does make one positive suggestion: Although risk-factor epidemiology will never be as sharp a tool as randomized clinical trials, epidemiologists could still benefit by adopting some of the scientific practices of those studies. "The great thing about a clinical control trial," he says, "is that, within limits, you don't have to believe anybody or trust anybody. The planning for a clinical control trial is prospective; they've written the protocol before they've done the study, and any deviation that you introduce later is completely visible." While agencies like the NCI do insist on seeing study protocols in risk-factor epidemi-

ology prospectively, this is still not standard procedure throughout the field. Without it, says Temple, "you always wonder how many ways they cut the data. It's very hard to be reassured, because there are no rules for doing it."

In the meantime, UCLA's Greenland has one piece of advice to offer what he calls his "most sensible, level-headed, estimable colleagues." Remember, he says, "there is nothing sinful about going out and getting evidence, like asking people how much do you drink and checking breast cancer records. There's nothing sinful about seeing if that evidence correlates. There's nothing sinful about checking for confounding variables. The sin comes in believing a causal hypothesis is true because your study came up with a positive result, or believing the opposite because your study was negative."

—Gary Taubes



TERRY O'DONNELL

**"The sin comes in believing a causal hypothesis is true because your study came up with a positive result."**

—Sander Greenland

# SCIENCE

Published by the **American Association for the Advancement of Science (AAAS)**, *Science* serves its readers as a forum for the presentation and discussion of important issues related to the advancement of science, including the presentation of minority or conflicting points of view, rather than by publishing only material on which a consensus has been reached. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by the AAAS or the institutions with which the authors are affiliated.

The American Association for the Advancement of Science was founded in 1848 and incorporated in 1874. Its objectives are to further the work of scientists, to facilitate cooperation among them, to foster scientific freedom and responsibility, to improve the effectiveness of science in the promotion of human welfare, to advance education in science, and to increase public understanding and appreciation of the importance and promise of the methods of science in human progress.

## Membership/Circulation

**Director:** Michael Spinella  
**Deputy Director:** Marlene Zendell  
**Member Services:** Rebecca Dickerson, *Manager*; Mary Curry, *Supervisor*; Pat Butler, Helen Williams, Laurie Baker, *Representatives*  
**Marketing:** Dee Valencia, *Manager*; Jane Pennington, *Europe Manager*; Hilary Baar, *Associate*; Angela Mumeka, *Coordinator*  
**Research:** Renuka Chander, *Manager*  
**Business and Finance:** Kevin Bullock, Nina Araujo de Kobes, *Coordinators*  
**Computer Specialist:** Chris Hageman  
**Science Member Services**  
Washington, DC: 202-326-6417  
**Other AAAS Programs:** 202-326-6400

## Advertising and Finance

**Associate Publisher:** Beth Rosner  
**Advertising Sales Manager:** Susan A. Meredith  
**Recruitment Advertising Manager:** Janis Crowley  
**Business Manager:** Deborah Rivera-Wienhold  
**Finance:** Randy Yi, *Senior Analyst*; Shawn Williams, *Analyst*  
**Marketing:** John Meyers, *Manager*; Allison Pritchard, *Associate*  
**Traffic:** Carol Maddox, *Manager*; Christine Pierpoint, *Associate*  
**Recruitment:** Terri Seiter Azie, *Assistant Manager*; Pamela Sams, *Production Associate*; Celeste Miller, Bethany Ritchey, Rachael Wilson, *Sales*; Debbie Cummings, *European Sales*  
**Reprints:** Corrine Harris  
**Permissions:** Lincoln Richman  
**Exhibits Coordinator:** Arlene Ennis  
**Administrative Assistant:** Nyana Gollopp de King  
**PRODUCT ADVERTISING SALES: East Coast/E. Canada:** Richard Teeling, 201-904-9774, FAX 201-904-9701 • **Midwest/Southeast:** Elizabeth Mosko, 312-665-1150, FAX 312-665-2129 • **West Coast/W. Canada:** Neil Boylan, 415-673-9265, FAX 415-673-9267 • **UK, Scandinavia, France, Italy, Belgium, Netherlands:** Andrew Davies, (44) 1-457-838-519, FAX (44) 1-457-838-898 • **Germany/Switzerland/Austria:** Tracey Peers, (44) 1-270-760-108, FAX (44) 1-270-759-597 • **Japan:** Mashy Yoshikawa, (3) 3235-5961, FAX (3) 3235-5852  
**RECRUITMENT ADVERTISING SALES: US:** 202-326-6555, FAX 202-682-0816 • **Europe:** Gordon Clark, (44) 1-81539-5211, FAX (44) 1223-302068 • **Australia/New Zealand:** Keith Sandell, (61) 02-922-2977, FAX (61) 02-922-1100  
Send materials to *Science* Advertising, 1333 H Street, NW, Washington, DC 20005.

**Information for Contributors** appears on pages 112–114 of the 6 January 1995 issue. Editorial correspondence, including requests for permission to reprint and reprint orders, should be sent to 1333 H Street, NW, Washington, DC 20005.  
**Science World Wide Web address:** <http://www.aaas.org>  
**Other Internet addresses:** [science\\_editors@aaas.org](mailto:science_editors@aaas.org) (for general editorial queries); [science\\_letters@aaas.org](mailto:science_letters@aaas.org) (for letters to the editor); [science\\_reviews@aaas.org](mailto:science_reviews@aaas.org) (for returning manuscript reviews); [membership@aaas.org](mailto:membership@aaas.org) (for member services); [science\\_classifieds@aaas.org](mailto:science_classifieds@aaas.org) (for submitting classified advertisements); [science\\_advertising@aaas.org](mailto:science_advertising@aaas.org) (for product advertising)

# LETTERS

## Cause and effect?

Like the statistical methods it uses, the field of epidemiology is relatively young. The study of disease incidence, distribution, and control in populations has helped squelch incipient plagues, tag new pathogens, flag toxic substances, and clear paths to healthy living. But epidemiology may have tested the limits of its powers, as indicated by “growing pains” of controversy. When does a research result warrant a headline? Or a new regulation? What constitutes a real risk to health? Such questions—of value and judgment as well as technique—are discussed here by some of the field’s top practitioners in response to a 14 July Special News Report.

## GRAPHIC OMITTED PER PUBLISHER

## The Discipline of Epidemiology

In the Special News Report “Epidemiology faces its limits” (14 July, p. 164), Gary Taubes assembles a series of quotations from ourselves and others about potential methodologic pitfalls in epidemiologic studies that might leave readers with the misimpression that evidence based on epidemiologic findings is not usually credible.

A problem does exist with general media reports about single scientific studies. Such reports often herald new results without describing the scientific context, which can create unnecessary fear and confusion. However, this is more an abuse of epidemiologic evidence than a problem with epidemiologic research. Taubes seems to perpetuate this confusion by listing several media reports of published findings and telling the reader “you be the judge” (p. 156) when proper judging is impossible without substantial additional information. In any scientific field, findings of individual studies are usually not considered seriously until confirmed by others. Also, in epidemiology, as in any other scientific field, more powerful studies need to be conducted to evaluate smaller effects, where sources of bias may be especially problematic. Often, doing so will require large and long-term prospective studies with repeated measures of exposure based on both questionnaires and biological measurements; a substantial number of such studies have commenced over the last 15 years.

Taubes did not emphasize that what we do know about the prevention of cancer and cardiovascular disease has derived largely from epidemiologic findings. This knowledge includes not just the many ad-

verse effects of cigarette smoking, but also the relation of overweight to many diseases, the benefits of increased physical activity for cardiovascular disease, the effects of many occupational exposures (such as benzene and asbestos), the relation of exogenous postmenopausal estrogens to cancer of the uterus, the relation of sunlight to all forms of skin cancer, the relation of ionizing radiation to many cancers, the adverse effects of many pharmacologic agents (for example, DES and thalidomide), and the protective effects of high intake of fruits and vegetables against many cancers.

Epidemiology has also provided important reassurance that many aspects of daily life are *not* major risk factors. For example, the relation between coffee consumption and coronary heart disease may not be completely settled, but the danger is minimal: The uncertainty is whether as much as five cups per day is a weak risk factor or not a risk factor at all (1). Fear of saccharin carcinogenicity engendered by studies in rats was quelled by epidemiologic research. Furthermore, epidemiologic studies have provided clear evidence that the incidence of several other forms of cancer, including ovarian cancer, is lessened as a consequence of using birth control pills.

If we wish to continue our progress in understanding the importance of lifestyle and environmental risk factors, we have little choice but to monitor the occurrence of illness of persons who have and have not been exposed to such factors. As Bruce Ames, a molecular biologist at the University of California, has noted (2), advances in other biological sciences can greatly add to the power of epidemiologic studies, but cannot replace them.

**Walter Willett**

Department of Nutrition,  
Harvard School of Public Health,  
Harvard University,  
Boston, MA 02115, USA

**Sander Greenland**

Department of Epidemiology,  
School of Public Health,  
University of California,  
Los Angeles, CA 90095, USA

**Brian MacMahon**

Department of Epidemiology,  
Harvard School of Public Health

**Dimitrios Trichopoulos**

Department of Epidemiology,  
Harvard School of Public Health

**Ken Rothman**

Epidemiology,  
Newton Lower Falls, MA 02162-1450

**David Thomas**

Program in Epidemiology,  
Fred Hutchinson Cancer Research Center,  
Seattle, WA 98104, USA

**Michael Thun**

Department of Epidemiology,  
American Cancer Society,  
Atlanta, GA 30329-4521, USA

**Noel Weiss**

Department of Epidemiology,  
School of Public Health and  
Community Medicine,  
University of Washington,  
Seattle, WA 98195, USA

**References**

1. S. Greenland, *Epidemiology* **4**, 366 (1993).
2. B. N. Ames, *Science* **221**, 1256 (1983).

Taubes's report is insightful and useful for epidemiologists and nonepidemiologists alike. However, I have two objections, one of them of personal nature, the other more general.

Taubes writes that I have expressed the view that only a fourfold risk should be taken seriously. This is correct, but only when the finding stands in a biological vacuum or has little or no biomedical credibility. We all take seriously small relative risks when there is a credible hypothesis in the background. Nobody disputes that the prevalence of boys at birth is higher than that of girls (an excess of 3%), that men have a 30% higher rate of death compared to women of the same age, or that fatality in a car accident is higher when the car is smaller.

The more general issue is that Taubes has omitted a consideration that is of paramount importance in any scientific argument. Epidemiology should be evaluated in comparison to other disciplines that serve the same objective, that is, to identify the causes of human disease and facilitate their prevention. Among these disciplines, only

epidemiology can document causation without concern about dose-extrapolation or species variability and with built-in accounting for potential modifiers.

It could be said for epidemiology, with respect to disease etiology and prevention, what is frequently said about democracy as a system of government: They both have many problems and weaknesses, but they still represent the best available approach for the achievement of their respective objectives.

**Dimitrios Trichopoulos**

Department of Epidemiology,  
Harvard School of Public Health,  
Harvard University,  
677 Huntington Avenue,  
Boston, MA 02115, USA

Taubes's excellent article about the proliferation of health-related messages to the public, and in particular the role of the popular press in their promulgation, misses one factor driving this process. Research institutions are eager to have the results of health risk factor studies performed in their laboratories appear in prominent newspapers and news magazines. This is so because individual philanthropists like almost nothing better than to support institutions whose research efforts have appeared on

# I'm looking to purify peptides, proteins and oligos simply, accurately and fast.

Whether you're purifying a peptide, protein or oligonucleotide, one tool assures you of receiving your results simply. The tool is pre-packed RESOURCE™ columns from Swedish-based Pharmacia Biotech. You'll find every RESOURCE leads to efficient and convenient working methods for a wide range of applications.

Efficiency is yours as RESOURCE allows for excellent resolution at high flow rates with high capacity at moderate back pressures. And convenience is yours as you can work with ion exchange, RPC or HIC.

What gives RESOURCE its combination of simplicity, accuracy and speed? The answer is the media. Every RESOURCE is pre-packed with SOURCE™ media—so while pre-packed RESOURCE will save you the time of doing your own packing, monosized beads of SOURCE assure you of reproducible results and batch to batch consistency.

With this blending of media with uncompromising pore structure and a tool with convenient handling and working procedures, pre-packed RESOURCE offers a unified approach for complete small scale purification. To find out more about RESOURCE, just call us at 1 (800) 526 3593 in the United States or +46 18 16 5011 from the rest of the world. It's that simple.



page one of, say, the *New York Times*. With the deceleration in government funds available for research and the concomitant increased dependence on private, and especially individual, funding sources, there will likely be an acceleration of these sorts of articles appearing in the popular press. It would generate far less confusion if they were just left in the scientific literature.

**Jerry Rapp**

Department of Biological Sciences,  
State University of New York,  
New York, NY 10010-3677, USA

The limits of epidemiology for environmental studies are well covered by Taubes. Genetic epidemiology is quite a different story. Clustering of cancer in families has led to the recognition of tumor suppressor genes by Alfred G. Knudson Jr. through study of retinoblastoma in childhood (1). These genes have since been found in other cancers of children and some of the commonest cancers of adults. Epidemiologic identification of the diverse familial cancers that cluster in Li-Fraumeni syndrome led to laboratory research that has furthered understanding the role of the *p53* gene in carcinogenesis (2). New clues to the origins of neoplasia are also coming from laboratory studies based on cancer clusters in heritable

disorders, such as ataxia-telangiectasia (3). Genetic epidemiology should not suffer guilt by association with the downside of its environmental counterpart.

**Robert W. Miller**

Genetic Epidemiology Branch,  
National Cancer Institute, EPN 400,  
Bethesda, MD 20892, USA

#### References

1. A. G. Knudson Jr., *Proc. Natl. Acad. Sci. U.S.A.* **68**, 820 (1971).
2. D. Malkin *et al.*, *Science* **250**, 1233 (1990).
3. K. Savitsky *et al.*, *ibid.* **268**, 1749 (1995).

When critics of epidemiology pay homage at the altar of the randomized clinical trial, such trials are made to sound only moderately troublesome compared to observational studies, when in fact they are often absolutely impractical or absolutely unethical. Examples include randomizing women to method of birth control and individuals to diet.

For such research, observational studies are the only recourse if you want to work with humans. The future and power of epidemiology rest not with simply self-reported data, but with combining such information with molecular data on susceptibility. In this way, risk measurements reflect characteristics of both host and environment

and make targeting prevention strategies rational. The challenge will be to use these host factors, such as genetic data, in a socially acceptable and nonpunitive fashion. Then epidemiology will provide truly meaningful and relevant estimates of risk.

**Alfred J. Saah**

Department of Epidemiology,  
School of Public Health,  
Johns Hopkins University,  
Baltimore, MD 21205, USA

Most of the epidemiology of multifactorial diseases fails a test of method, due to absent experimental randomization and unachievable control of biases and confounders. In general, it also fails the ultimate test of predictivity, as large randomized experiments designed to verify major observational inferences have been thoroughly disappointing (1). Now, a resounding admission of impotence threatens our survival and demands remedial measures.

As other professionals have done, epidemiologists could establish a code of good practice, spelling out optimal standards of hypothesis formulation, study design, and conduct. Structural uncertainties should limit heuristic causal inferences to relative risk or odds ratio values above 3 or 4, as Trichopoulos (quoted in the article by

Here's a handful  
of simple RESOURCES  
from Sweden.



**Pharmacia  
Biotech**

Uppsala, Sweden. (And the rest of the world)

Circle No. 58 on Readers' Service Card  
Electronic Marketplace: <http://www.aaas.org>



Taubes) and others before him have concluded (2). Although still short of assuring verification, this last provision would link with *de minimis* considerations of ongoing regulatory reform.

Epidemiologist have no choice but to warrant their credibility. We owe it to society and to the young entering the profession, who need to know honestly whether they can make a difference. Too much of epidemiology has become predictable advocacy without secure philosophical foundations. A code of good epidemiologic practice would be a beginning, perhaps after some soul-searching about the morality of provoking public anxieties and policies based on essentially unverifiable conjectures.

**Gio Batta Gori**  
Health Policy Center,  
Bethesda, MD 20816, USA

#### References

1. M. Susser, *Epidemiol. Rev.* 7, 147 (1985); L. Werko, *Acta Med. Scand.* 221, 323 (1987); MRFIT Group, *J. Am. Med. Assoc.* 263, 1795 (1990).
2. N. E. Breslow and N. E. Day, *Statistical Methods in Cancer Research, vol. 1, The Analysis of Case-Control Studies* (International Agency for Research on Cancer. Publ. No. 32, Lyon, France, 1980); K. J. Rothman, in *Cancer Epidemiology and Prevention*, D. Schottenfeld and J. F. Fraumeni, Eds. (Saunders, Philadelphia, PA, 1982), pp. 15-22; E. L. Wynder, *Prev. Med.* 16, 139 (1987).

#### Glass Ceiling: Bump, Bump

We were struck by the excess of males among those quoted in Taubes's news article of 14 July: 25 men versus 2 women. The Society for Epidemiologic Research, the primary professional organization of epidemiologists in the United States, has a membership, as of 1993, of 1194 men and 1009 women. The latter include senior faculty, department chairs, and a dean of a school of public health. Prominent female epidemiologists are located in most of the institutions where those who were interviewed work. Many of the studies cited in the news report had women as first authors. Women epidemiologists deserve more of a voice in *Science*.

**Irva Hertz-Picciotto**  
Department of Epidemiology,  
University of North Carolina,  
Chapel Hill, NC 27599-7400, USA

**Maureen Hatch**  
Department of Community Medicine,  
Mount Sinai Medical Center,  
New York, NY 10029-6574, USA

As evolutionary biologists, we were excited and interested to see the "Frontiers in biology: Ecology" special section in the 21 July issue (pp. 313-360). As women scientists, we were disappointed that in the first two

articles only one of the more than 30 ecologists mentioned or quoted was a woman. From this representation it is difficult to tell that ecology is a field where, in 1992, 36% of the graduating Ph.D.'s were women and where four of the last nine Ecological Society of America presidents were women. We know "a good woman is hard to find," but really. . .

**Sharon Emerson**  
**Phyllis D. Coley**  
Department of Biology,  
University of Utah,  
Salt Lake City, UT 84112, USA

#### Plasma Physics and Fusion Research

James Glanz's otherwise excellent article about the National Research Council's (NRC's) panel report on the state of plasma physics and fusion research (News, 14 July, p. 153) does not treat what may be the most intractable part of the history of the subject: the degree to which the perspectives, procedures, and dominant personalities of the field have been selected by the Department of Energy (DOE) and its ancestors (the Energy Research and Development Administration and the Atomic Energy Commission), on the basis of an agenda that was not solidly rooted in anything scientific. Our style did not emerge from any traditional academic process, but rather a political and economic one.

Every branch of physics older than plasma physics developed its habits, interests, and formative experiences in the rough-and-tumble atmospheres of university seminars for several years before they became of interest to newspapers or government agencies. Collegial ideas about how and on what to work were allowed to develop to some extent independently of the funding required to support them. From day one, with only the briefest of interludes in the 1960s, plasma physics has had its priorities arranged by managers in the government who, while well meaning, were essentially unacquainted with the subject at a working research level. The subject, under their tutelage, began to speak with one voice in public about 20 years earlier than it was appropriate to do so. In selling Congress and the *New York Times* on the tokamak as the cure for energy shortages in the early 1970s, the field committed itself to a way of life in which its public image and its annual funding struggles in Congress assumed more importance than any scientific issue that could ever come up. To a large degree, we are still functioning in this mode.

A technical point, largely unappreciated, is the extent to which plasmas at the tem-

peratures we now operate at are experimentally undiagnosed. Information about spatial and temporal profiles of such internal plasma variables as the magnetic field, the current density, the velocity field, and the electric field is largely lacking. Stories about the internal dynamical behavior of confined plasmas are easy to make up, hard to dispute, and at this stage virtually impossible to demonstrate. It is largely unappreciated that the DOE in its wisdom went around for years turning off every plasma experiment that was cool enough to diagnose, on the grounds that those temperatures "were not of thermonuclear interest." Only lately has it been possible to hear respectable doubts expressed that this was a wise thing to have done. Many groups perished then and were not heard from again.

If the NRC or anybody else can turn the situation around, then more power to them. But it would be a mistake to think that it is obvious how to do this. Even very good people who have spent a lifetime adapting themselves to unwise agency policies not of their making and being rewarded for it are not likely, in middle age, suddenly to start biting the coins and questioning the wisdom of what they have been doing for the last few decades. What plasma physics needs more than anything is a long period of benign neglect, during which it is modestly but reliably funded, insulated from agency-directed campaigns and from congressional feasts and famines, and allowed to go through the scientific maturation that has heretofore been denied it. When we are ready to build a fusion reactor, you will know it; it won't be a matter of lobbying or image-making.

**David Montgomery**  
Department of Physics and Astronomy,  
Dartmouth College,  
Hanover, NH 03755-3528, USA

#### Restoring Prince William Sound

I am concerned that the article "Marine center is lightning rod in dispute over restoration" by Lisa Busch (News & Comment, 14 July, p. 159) will leave readers with the impression that the decision of the Exxon Valdez Oil Spill Trustee Council to support the Alaska Sealife Center is divisive, widely opposed, and leaves residents of Prince William Sound with incomplete restoration. The article does not mention that the Trustee Council has spent tens of millions of dollars to improve other aspects of pink salmon and Pacific herring management in Prince William Sound, including more than \$9 million to support the Sound Ecosystem Assessment, based at the Prince William Sound Science Center in Cordova, which is investigating the causes of annual