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 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 noninfectious diseases—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.” Agree Ken 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 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 believe more skeptical about their own findings—or both.

An observational science
What drives the epidemiologic quest for risk factors is the strong circumstantial evidence that what we eat, drink, breathe, and so on are major factors in many devastating illnesses. Rates of heart disease, for example have changed much faster over recent decades than can be explained by genetic changes, implicating diet.
Sizing Up the Cancer Risks

A steady diet of epidemiology, only a sliver of the environmental risks we face, has steadily and strongly linked to human cancers. And this is why the University of Alabama epidemiologist Philip Cole lists cigarette smoke, alcohol, obesity, radiation, and a handful of occupational carcinogens, such as asbestos, among three villains—human papillomavirus, human T cell leukemia virus, and human papillomavirus. But even a year, epidemiologists publish by the journal load, many of them addressing the potential cause of cancer in the environment.

The results of observational studies, in which scientists try to compare the lives of people suffering from a disease with those of healthy controls, vary in their practitioners efforts. This effort is plagued by biases and confounding factors (see text). As a result, most epidemiologists interviewed by said they would not take seriously a single study reporting a potential cause of cancer unless it reported a person’s risk by at least a factor of 3. Even then, they say, the result is in order unless the study was very large and extensively done and biological data support the hypothesis. An epidemiologist, a University of California, Los Angeles, says a study reporting a twofold increased risk is then worth taking seriously—but not that seriously.

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

**G.T.**

- Eating yogurt at least once a month---RR 2 for ovarian cancer August 1989.
- Eating more than 100 cigarettes in a lifetime—RR 1.2 for breast cancer (February 1990).
- Non-occupational exposure to dioxins—RR 1.5 for all cancers January 1991.
- Douching once a week—RR 4 for cervical cancer (March 1991).

**Use of phenol ether herbicides on lawns—RR 1.3 for malignant hemophagocytic 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 (partially 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 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 procedure for studies of new drugs and other medical interventions. The randomized trial involves assigning subjects to random groups 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 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 selecting thousands of healthy people to pollutants or other possible carcinogens raises obvious ethical problems.

Because the experimental approach is often limited for much of epidemiology, researchers often resort to observational approaches. In case-control studies, for example, they select a group of individuals affected 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. Then they follow the entire population for years or decades to see who gets sick and who doesn't, what dis-
Press Coverage: Leaving Out the Big Picture

In the past 2 years, thorough readers of the Los Angeles Times would have learned about an extraordinary range of potential cancer causes. Among these putative hazards of modern life are hot dogs, breast implants, dioxin, stress, asbestos, allergy drugs, gas leaks, living in Orange County, tubal ligation, sunscreen, Asian food, pesticides, vasectomy, liquor, working in restaurants, Retin-A, vegetables, dietary fat, delayed child-bearing, impurities in meat, and lesbianism. This litany of fear was accompanied by a similar, although shorter, series of reports on dietary habits and lifestyles that may reduce cancer risk. Parallel coverage appeared in other newspapers and magazines and on television. To many scientists, though, the media would do well to curb its appetite for such news.

The problem, many researchers say, is that journalists often misunderstand the context of the research. Because of the limitations of risk-factor epidemiology, most individual studies cannot produce authoritative findings (see main text). "Articles published in medical journals are often misconstrued by the lay press to be more definite than they really are," says Larry Freedman, a biostatistician at the National Cancer Institute. "Broccoli prevents cancer, garlic prevents cancer—all these things do appear in the literature. But epidemiologists understand very well that these studies are far from definitive. It's only when a body of evidence exists over many, many studies that epidemiologists should really get serious about giving the public advice."

Instead of presenting surveys of the big, evolving picture, he and others say, the media tend to report each new study in isolation, as a new breakthrough. Such reporting is encouraged by press releases put out by journals and researchers' institutions. But whoever is to blame, says Noel Weiss, an epidemiologist at the University of Washington in Seattle, the result is "just too many false alarms. When we do have a serious message, I fear it won't be heeded because of the large number of false messages."

One example is an item from Time magazine's "Health Watch," which tersely summarizes recent research. Published last January, the item read, in its entirety, "Olive oil seems to do more than make food taste good. Research indicates that women who consume olive oil more than once a day reduce their risk of breast cancer 25% compared with women who don't." Time didn't mention that the risk reduction is smaller than many epidemiologists think can be reliably detected in an observational study. Nor did it point out that the study—apparently a study of 1750 Spanish women reported several weeks earlier in the International Journal of Cancer—is in conflict with many other studies suggesting that dietary fats may raise rather than lower the risk of breast cancer. Although the overall fat-breast cancer link is disputed, and olive oil may pose less cardiovascular risk than other forms of fat, few epidemiologists would interpret these findings as indicating that women should "consume olive oil more than once a day."

In their proclivity for "news," newspaper and television reporters not only single out weak studies; they may focus on the one positive result in a sea of negative data. That was the case with coverage of two big studies on occupational exposure to electromagnetic fields (EMF) that appeared recently in the American Journal of Epidemiology. The first study, of 223,000 French and Canadian electric utility workers, found no link between EMF and 25 of the 27 varieties of cancer in the study; the exceptions, two rare types of leukemia, had a weak and inconsistent positive association with EMF. Yet the Wall Street Journal reported the study in late spring under the headline, "Magnetic Fields Linked to Leukemia."

Early this year the American Journal of Epidemiology published the second study, on 139,000 workers at five U.S. utilities. It found no association between exposure to EMF and 17 of 18 types of cancer, including the leukemias linked to EMF in the first study. The sole exceptions were eye and brain cancers—conditions that had shown no link to EMF in the first study. Yet the headline of the Wall Street Journal article that reported the second study was "Link Between EMF, Brain Cancer Is Suggested by Study at 5 Utilities." Says Jerry Bishop, who wrote one of the Wall Street Journal articles, "People are not interested in what diseases [a risk factor] don't cause, but what it might cause. ... We've had this argument with scientists many times over the past few years."

In October, the New York Times provided another example, when it reported on a study in the Journal of the National Cancer Institute (JNCI) from the Fred Hutchinson Cancer Research Center in Seattle that suggested induced abortion might increase the risk of breast cancer by 50%. Although the article noted that 40 previous studies of abortion and breast cancer had found no such correlation, the headline read "New Study Links Abortions and Increase in Breast Cancer Risk." Inevitably, public attention was directed to a risk that is unlikely to be real.

If there is "blame" for such coverage, argues Lawrence Altman, author of the Times article, much of it belongs to scientific journals. "The JNCI sent out a big release touting that study as if it were the biggest thing since whatever," he says. "I don't recall them telling us that it was only one of 40 studies and probably had little meaning."

In Altman's view, epidemiologists who complain about press coverage are trying to have it both ways. "Scientists supposedly want us not to go outside the scientific process, but wait until findings have appeared in a peer-reviewed professional journal. When we do that, they apparently complain that we didn't go outside the scientific process and say that a published report is meaningful."

"Journals do overemphasize individual studies, but they are often inspired to do that by medical journals," agrees Ross Prentice of Weiss's colleagues at the University of Washington. "The problem is, some of the press releases that journals and universities send out are legitimate. It's a wonder sometimes that the reporting isn't better."

Charles C. Mann

Charles C. Mann is the co-author, with Mark L. Plummer, of Noah's Choice: The Future of Endangered Species.

New Study Links Abortions and Increase in Breast Cancer Risk

Magnetic Fields Linked to Leukemia

Link Between EMF, Brain Cancer Is Suggested by Study at 5 Utilities

But More Canadian Study Fails to Show Exposure Is Case of the Cancer

But Industry-Flavored Work Fails to Show EMF Risk Is Illusory

...
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 "most suggestive 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 exposure 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's 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 underreport fat intake on questionnaires or in interviews and obese subjects tend to overreport 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...
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 self-reported questionnaires because they are far more numerous 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 interview bias as well. The interviewers are rarely blinded 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 prone 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 possible causal associations 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 results 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.”

What to believe?

So what does it take to make a study worth taking seriously? Over the years, epidemiologists have offered up a variety of criteria, the most important of which are a very strong association between disease and risk factor and a highly plausible biological mechanism. The epidemiologists interviewed by Science say they prefer to see both before believing the latest study, or even the latest group of studies. Many respected epidemiologists have published erroneous results in the past and say it is so easy to be fooled that it is almost impossible to believe less-than-stunning results. Sir Richard Doll of Oxford University, who once 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 epidemiological 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 accepting it as a 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 statistician consultant for the NEJM, points out, there is no reliable way of identifying the division 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, an 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 of 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. Willott, 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 confounders.
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 Charlestone, 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’s 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 seeing 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 ten 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 Mahon, 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 ways that do 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 NutraSweet, which contains the artificial sweetener aspartame, still tout it as the sweetener that does not have saccharin.

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 epidemiology 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 that if 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

“...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.’”

—Sander Greenland

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

SCIENCE • VOL. 269 • 14 JULY 1995