

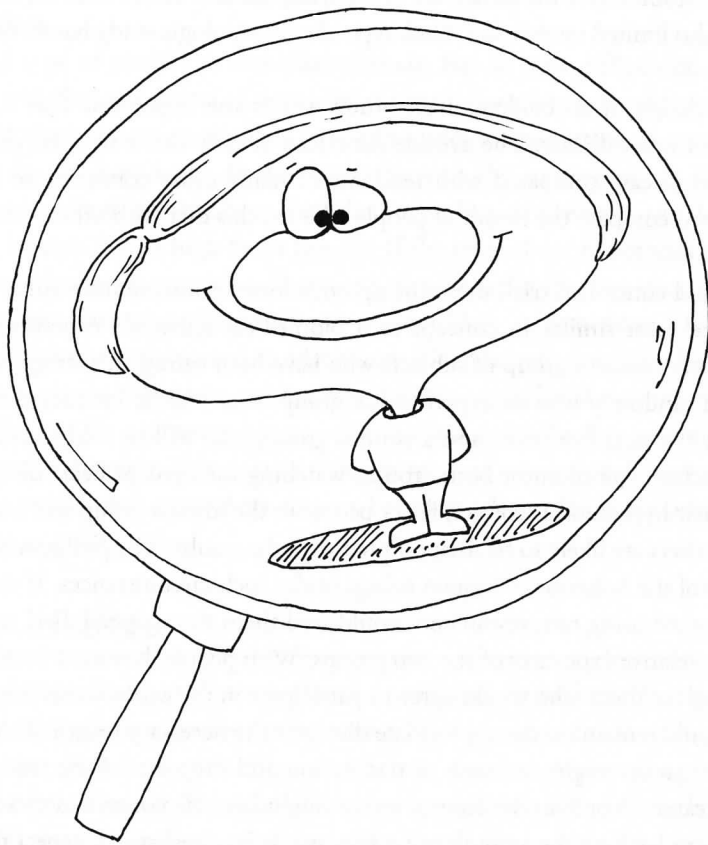
INTRODUCTION TO Public Health

Third Edition



Mary-Jane Schneider

Problems and Limits of Epidemiology



The Study of Humans

The ultimate goal of many epidemiologic studies is to determine the causes of disease. This is generally done first by observing a possible association between an exposure and an illness, second by developing a hypothesis about a cause and effect relationship, and third by testing the hypothesis through a formal epidemiologic study. While the formal study can strongly support the conclusion that a certain exposure causes a certain disease, there are many potential sources of error in drawing such a conclusion. Studies of chronic diseases, which often have multiple determinants and develop over long periods of time, are especially prone to error.

Problems with Studying Humans

All epidemiologic studies have the advantage of studying humans rather than experimental animals; but all are also limited by that fact. Each type of epidemiologic study has its own strengths and weaknesses.

Consider the design of an epidemiologic study to test the hypothesis that a low-fat diet reduces the risk of heart disease. The average American already eats a high-fat diet and has a high risk of heart disease compared with residents of many other countries, so it should be possible ethically to compare the health of people who eat this diet with others who have other dietary patterns.

The randomized controlled trial, the most rigorous form of intervention study described in Chapter 5, is the most similar in concept to a biomedical scientist's experiment with rats. Suppose researchers choose a group of subjects who have been eating an average American diet and divide them randomly into an experimental group, who will be instructed to eat a strict low-fat diet over the next five years, and a control group, who will be told to continue eating normally. Researchers will monitor both groups, watching for signs of heart disease, and they expect that, if their hypothesis is correct, fewer people in the low-fat group will become ill.

In fact, researchers are likely to be disappointed with the results. The problem is that it is impossible to control the behavior of human beings under such circumstances. If the experiment was being conducted using rats, researchers would feed them the assigned diets and could thus be certain of the relative exposures of the two groups. With people, however, even if researchers could find enough of them who would agree to participate in the experiment, it is questionable whether they would remain on the appropriate diet over the necessary length of time. People in the experimental group might succumb to temptation and drop out of the study or lie about what they have eaten. People in the control group might become concerned about their health and voluntarily cut back on the amount of fat they eat. It is unrealistic to expect to succeed at a randomized controlled trial that requires people to alter their behavior over a significant period of time, unless the subjects have a special motivation to participate—if they are suffering from a

serious disease, for example—and participation in a trial is their only chance to have access to a new, potentially more effective treatment.

To test the dietary hypothesis, researchers might try, instead of a randomized controlled trial, a cohort study. They would choose a large group of people who are free of heart disease, ask them detailed questions about their diets, and then, over the next five years, compare the health of those who already eat a low-fat diet with those who eat an average American diet. This would not require people to change their behavior. The problem with this scenario is that people who have voluntarily chosen to eat a low-fat diet may differ in other respects from the group who eat the average diet. The low-fat group members are likely to be more health conscious in general. They may be less likely to smoke and more likely to exercise, for example. These people, therefore, would have a reduced risk of heart disease even if a low-fat diet did not have a protective effect.

The third type of study, the case-control study, has its own difficulties. In this study, researchers would choose a group of people who already have heart disease; perhaps they would go to a hospital and interview patients recovering from a heart attack. A comparable group of people who do not have heart disease would serve as the control group. Researchers would question people in both groups about their diets over the past five years and decide whether the diets should be classified as high-fat or low-fat. If the researchers' hypothesis is correct, the patients who have had a heart attack will report a diet higher in fat than the control group. This approach also has obvious problems. People are not likely to remember what they ate in the past, or they might be embarrassed to admit how self-indulgent they have been. The information researchers obtain concerning exposure in the case-control trial may not be reliable.

These difficulties do not mean that no valid conclusion can be drawn from any kind of epidemiologic study. However, they demonstrate the types of errors that different kinds of studies may be prone to and alert researchers about what to watch out for in choosing a study design and in interpreting the results.

Sources of Error

News reports of new health studies can often be confusing. Sometimes there are conflicting reports on the health effects of various substances. Coffee is reported to cause heart disease; then it is reported that there is no such effect. Oat bran is reported to prevent cancer; then it is reported to make no difference. Fish is good for your heart; fish is full of toxic chemicals that may cause harm. All these contradictions tend to make people distrustful of the news and uncertain about how to protect their health. Since most of these news reports are based on epidemiologic studies, it is useful to understand possible sources of error in such studies and how to look for the truth in the reports.

One of the most common reasons for a study to lead to a wrong conclusion is that the reported result is merely a *random variation* and that the association is merely due to chance. Chapter 7 discusses some of the factors involved in determining whether a result is statistically significant. As a general rule, epidemiologic studies of chronic diseases require large numbers of subjects to draw valid conclusions. Causes of these diseases are usually complex, and there are usually long periods between exposures to possible causes and the development of illness, making it difficult to draw conclusions about associations between exposure and disease. The cause-and-effect relationship is not obvious—as it is, for example, when a bullet in the heart causes death, or exposure of an unvaccinated child to the measles virus causes the child to develop measles in ten to twelve days. The weaker the relationship between exposure and disease, the larger the group of people that must be studied for the relationship to be evident. If the group being studied is too small, a cause-and-effect relationship is likely to be missed or a spurious relationship will show up by chance alone. One of the reasons that the Doll–Hill and Hammond–Horn results concerning smoking and lung cancer are so convincing is that they involved such large numbers of subjects (see Chapter 4).

There are a number of other possible sources of error that well-designed studies may be able to avoid. For example, the cohort study of a low-fat diet proposed previously may be invalidated by the presence of confounding variables, smoking and exercise. *Confounding variables* are factors that are associated with the exposure and that may independently affect the risk of developing the disease. Such an error may have occurred in a 1980s study that suggested coffee drinking could cause pancreatic cancer, a finding that has not been replicated in other studies. Since many heavy coffee drinkers were also smokers, there are suspicions that the cancer was caused by the smoking rather than the coffee.¹ To eliminate the errors caused by smoking as a confounding variable, researchers might conduct the study only on nonsmokers. Alternatively, there are statistical techniques for adjusting the results to compensate for confounding variables as long as the investigator is clever enough to think of possible factors that may affect the result and to take them into consideration when collecting the data and calculating the results. While the investigators in the study of coffee corrected for smoking over the five-year period before the cancer was diagnosed, the correction may have been inadequate.

An interesting example of confounding occurred in a study, published in 1999 and widely publicized, suggesting that small children were more likely to become myopic—nearsighted—if they slept in a lighted room. In a follow-up study, investigators asked the children's parents about their own vision. It turned out that myopic parents were more likely to leave lights on in their children's rooms than parents with better vision. Their children, therefore, were more likely to be nearsighted because they inherited the condition from their parents, not from the light exposure.²

Bias, or systematic error, may be introduced into a study in a number of ways. *Selection bias* is a particular problem in choosing subjects for a case-control study. For example, if the cases of heart disease are chosen from hospitalized patients recovering from heart attacks, and the controls include hospitalized patients being treated for a digestive disorder that causes extreme discomfort from eating fatty foods, the study may suggest an exaggerated effect of dietary fat on heart disease. The results would probably be different if the controls were patients recovering from the effects of motor vehicle crashes, whose diet might be more like the average American's. Selection bias may also occur when there is a systematic difference between people who choose—or are chosen—to participate in a study and those who do not. For example, in a 1988 case-control study that found exposure to high electromagnetic fields (EMF) from power lines increased the risk of childhood cancer, the controls were chosen by a process of telephone random digit dialing until a child was located who matched a case by age and sex. Cases and controls were compared, and cases were found to have had a higher exposure to EMF. However, the cases also were found to be of lower socioeconomic status; they were more likely to live in areas of high traffic density, and their mothers were more likely to smoke. The random-digit dialing had created a bias: because poor families were less likely to have a telephone, or less likely to have an answering machine and to return calls, the control group was more affluent and consequently was less exposed to confounding poverty-associated factors.¹

An extreme example of selection bias—one that no well-trained epidemiologist would make—was seen in the report of the author Shere Hite on male and female relationships. Out of 100,000 questionnaires on women's attitudes about men and sex that Hite distributed, only 4500 replies were received. Hite reported that 84 percent of the women in the study were dissatisfied with their intimate relationships, results that were widely publicized. The low response rate suggests that selection bias was operating and that the most dissatisfied women were responding preferentially to the survey.³

Cohort studies, which tend to extend over many years, are likely to suffer from a form of bias caused by people dropping out or being untraceable when results are being sought. If people who get sick drop out at a different rate from those who remain healthy, the results will be compromised. Subjects who are lost to follow-up may be more likely than those who are traceable to have entered an institution or to have moved in with family, indicating a serious health problem. A high dropout rate casts doubt on the results of any epidemiologic study.

Reporting bias or recall bias is a common problem in case-control studies. It occurs if the study group and the control group systematically report differently even if the exposure was the same. Subjects' reports of their dietary intake are notoriously unreliable. For example, underweight individuals consistently overreport their fat intake, while obese individuals underreport

it.¹ Similarly, studies attempting to relate certain diseases to alcohol consumption may suffer from reporting bias because people who drink heavily tend to underreport their consumption. Case-control studies that attempt to determine causes of birth defects are especially subject to recall bias, since the mother of a child born with a malformation is likely to have thought a great deal about what might have caused the problem, while mothers of healthy children would be less likely to notice an unusual exposure.

Proving Cause and Effect

For the most part, epidemiologic studies, no matter how well designed to avoid error, cannot prove cause and effect. In fact, that is why epidemiologists usually speak of risk factors rather than causes. However, there are several factors that can be combined to make the cause-and-effect relationship almost certain.

First, as discussed previously, a study with a large number of subjects is more likely to yield a valid result than a small study. Second, the stronger the association measured between exposure and disease—the higher the relative risk or odds ratio—the more likely that there is a true cause-and-effect relationship. For example, the Reye's syndrome case-control study described in Chapter 5 found a 42.7 odds ratio from exposure to aspirin during a viral infection. The British case-control study linking birth control pills to breast cancer found only a 2.3 odds ratio, while the Nurses' Health Study—a cohort study—found at most a 1.5 relative risk of breast cancer from oral contraceptives. The much stronger association found in the Reye's syndrome study makes it highly probable that aspirin causes the syndrome in children, while the breast cancer results could possibly be due to some error or alternative explanation. Nevertheless, exposure to hormones is generally accepted as a risk factor for breast cancer, as discussed in the next section.

Third, a dose-response relationship between exposure and risk of disease is evidence supporting exposure as a cause of the disease. Some of the earliest evidence that long-term exposure to low levels of x-rays had adverse health consequences came from a study comparing the mortality rates of physicians exposed to different amounts of radiation. Radiologists had the lowest life expectancy of the three groups of specialists studied. Ophthalmologists and otolaryngologists, who have little exposure to radiation, had the highest life expectancy. Internists, whose exposure was intermediate, had intermediate life expectancy, confirming a dose-response effect—the higher the dose of radiation, the greater the effect on lifespan.⁴

Fourth, epidemiologic evidence is more convincing if there is a known biological explanation for an association between an exposure and a disease. Studies suggesting that EMFs cause leukemia and other forms of cancer have been looked on with skepticism because of the lack of a known mechanism by which such low energy fields could have a biological effect. The question is unresolved. However, a number of other exposures have been identified by epidemio-

logic studies as causes of disease before a biological explanation was found. For example, strong epidemiologic evidence that cigarette smoking was a major cause of heart disease existed long before there was any biological explanation, and the mechanism is still not well understood.

The most important indication that an epidemiologic result is valid is that it is consistent with other investigations. If several independently designed and conducted studies lead to the same conclusion, it is unlikely that the conclusion resulted from bias or other error. If the reports are conflicting, however, people must be wary of accepting any of the results.

Epidemiologic Studies of Hormone Replacement Therapy—Confusing Results

When women reach menopause at age fifty or so, their natural production of the hormone estrogen drops significantly. Many women at this stage of life begin to have menopausal symptoms that can be troubling: hot flashes that disturb their well-being during the day and their sleep at night and vaginal dryness that causes discomfort and interferes with sexual activity. Prescription of estrogen supplements relieves these symptoms, and this treatment became popular in the 1960s. Estrogen was promoted to help keep women “feminine forever” as promised in a best-selling book of that title by Robert Wilson, published in 1966.⁵ Large numbers of postmenopausal women took the hormone in the hope that it would keep them looking and feeling younger, improve their memory, and stave off other effects of aging. When evidence appeared in the 1970s that women taking estrogen had an increased risk of uterine cancer, the problem was averted by adding another hormone, progesterone, to the prescription. Progesterone countered the effect of the estrogen on the uterus without appearing to diminish its positive effects on other organs. There was good reason to believe that these female hormones protect women. Rates of cardiovascular disease are well-known to be much lower among women than men until middle age, increasing after menopause to match the rates among men. And older women are much more likely to suffer from osteoporosis, thinning of the bones that leads to fractures. It was reasonable to think that hormone supplements might also protect women against these problems.

Numerous epidemiologic studies over the years supported the protective role of estrogen for bones and hearts. Most notably, the Nurses’ Health Study, the large cohort study, ongoing since 1976, found that women taking hormone therapy had a 61 percent lower risk of heart disease and a 75 percent lower risk of hip fractures.⁶ These studies found small increases in breast cancer risk, but the trade-off seemed worthwhile for many women. In 1999, approximately 38 percent of postmenopausal women in the United States were using hormone-replacement therapy (HRT).⁷

Then in July 2002 the news broke that HRT was not as beneficial as it had seemed. The previous positive evidence had all come from observational studies. Meanwhile, a huge clinical trial, called the Women's Health Initiative (WHI), had been under way since 1991. The researchers announced in 2002 that the WHI had been stopped early on the basis that the risks had been found to outweigh the benefits.⁸ Women randomly assigned to take a combination pill of estrogen plus progesterone were found to have a higher risk of breast cancer than women taking a placebo, which was not surprising. The surprise was that women taking the pill were also found to have a higher risk of heart attack, stroke, and blood clots. The women in the experimental group had fewer hip fractures and fewer cases of colorectal cancer than the control group, but this protective effect was not enough to outweigh the risks.

The news from the WHI study seemed to contradict the overwhelming evidence from cohort studies that HRT protected women against heart disease. However, the WHI was a clinical trial, the gold standard of epidemiologic studies, and thus was much less likely to be subject to bias. Many women stopped taking HRT when the news came out, and the drug's sales fell by 50 percent within six months.⁹

Since reports of the study were published, epidemiologists have been struggling to understand why the two studies produced such conflicting results. There are still many unanswered questions, but one important factor seems to be selection bias. Women in the observational studies who chose to take hormones were healthier to begin with and had healthier habits than the women who did not take the hormones. Many other factors appear to be involved, including biologic differences between the women in the two types of studies (women in the Nurses' Health Study were younger and thinner than the women in the WHI); there is also a bias stemming from the fact that cohort studies tend to miss adverse events that occur very soon after a therapy is begun, and the cardiovascular risk from HRT is highest during the first year after beginning therapy.^{10,11} Evidence supporting some of the WHI conclusions emerged in 2006 when routinely collected cancer data revealed that breast cancer incidence in the United States had dropped significantly in 2003 and 2004, apparently the result of so many women discontinuing use of HRT.⁹ Current recommendations call for HRT to be used only short-term for postmenopausal symptoms.

Ethics in Epidemiology

Most epidemiologic studies are observational and have little potential for harm. There are exceptions, however, especially in the conduct of intervention studies. Nowadays, strict ethical limitations apply in any study involving humans. These rules were developed in reaction to abuses such as those by Dr. Joseph Mengele, who conducted medical experiments on concentration camp prisoners during World War II. Ethical abuses have not been limited to Nazi war

criminals, however. At one time, medical researchers in the United States were not overly concerned with the rights of the experimental subjects, who were often poor patients or captive populations such as prisoners or inmates of mental institutions. That changed in 1972, when news of the Tuskegee syphilis study shocked the nation.

Syphilis was a dreaded disease for hundreds of years, inspiring some of the same moral revulsion as AIDS has sometimes done more recently. Spread by sexual contact, syphilis had an unpredictable course that, over a variable number of years, could lead to a range of grim symptoms, including blindness, heart disease, dementia, and paralysis. It was sometimes treated with an arsenic-containing drug called salvarsan, which had been shown to cure syphilis in rabbits but which was not always effective in human patients and sometimes killed them. Some scientists suspected that the disease was not as uniformly dire as its reputation suggested and that the treatment might be worse than the disease. This conclusion was supported by the results of a Norwegian study of untreated syphilis done during the early part of the 20th century, which found that up to 75 percent of the patients were symptom-free after more than twenty years of the disease.¹²

In 1932, the U.S. Public Health Service and scientists from Tuskegee Institute began a similar study of about 400 black men in Macon County, Alabama, where syphilis was rampant: 40 percent of the population suffered from the disease. The purpose was to observe the course of the disease in these men, who were not to receive treatment. In part because it was not common practice at the time, and in part because the subjects were poor, black, and uneducated, the investigators did not try to explain what they intended to do or ask the subjects' permission. The men were told they had "bad blood" and were enticed to participate with free "treatments" and physical examinations, free hot lunches, and free burials. In the 1940s, penicillin was discovered and became standard treatment for syphilis, but the Tuskegee subjects did not receive the antibiotic until after the story broke in 1972.¹²

There is some question about whether the men were physically harmed by the withholding of antibiotic treatment. The course of the disease is complicated, and the surviving subjects were in a late, noninfectious stage by the time penicillin was discovered, perhaps too late to help most of them. However, this study raised a number of ethical issues, the major one being that the men were deceived. They were not told what syphilis was or that they were part of a study, and they were led to believe that they were receiving treatment. Furthermore, one of the tests that was done on the subjects was a spinal tap, a painful procedure that uses a needle to withdraw spinal fluid, which has the potential of causing harmful side effects, including—rarely—paralysis. This treatment would not likely have been tolerated by white, middle class Americans, and many critics have concluded that the study was racist. In fact, revelations about the study led many African-Americans to distrust medical research. The misconception still lingers that the men were deliberately infected with syphilis.¹³

The outcry that followed the publicity about the Tuskegee study in 1972 led directly to the establishment of rules for the conduct of human experimentation. All institutions that receive federal funds must follow these rules. The rules require that every research subject must be informed of the purpose of a study and its risks and benefits. The subjects must freely consent to participate. In addition, any such study must be approved in advance by an institutional review board, a committee that includes representatives of the community as well as other scientists, who must agree that the study is well designed, that its benefits outweigh its risks, and that the subjects are truly given the opportunity for informed consent. Clinical trials are halted if the treatment group is clearly showing better or worse results than the control group. This was done, for example, in the portion of the Physicians' Health Study that looked at aspirin's effectiveness in preventing heart attacks when it became clear that subjects taking aspirin were suffering fewer heart attacks than those in the placebo group.¹⁴ It was also done in the WHI study of HRT, described previously.

Even with the current strict ethical guidelines, there are a number of controversial issues surrounding clinical trials, including whether such trials should be conducted at all, who should participate, whether informed consent is truly possible, and whether unproven treatments should be available outside of clinical trials.

All of these controversies came to the foreground several years ago in connection with the AIDS epidemic. People with AIDS knew they had a fatal disease that had no known cure, and they were desperate. Many of them were very politically active. People with AIDS argued that they did not have time to wait for clinical trials to test the efficacy of every promising new drug. They wanted immediate access to any new drug that showed promise in the laboratory, because they would prefer to try something—anything that had the slightest chance of working—rather than face certain death. On the other side of the argument is the history of useless therapies that have been employed for years or decades because no one had ever done a scientific test of whether they worked. This is a true ethical dilemma pitting the individual against society. Can we deny today's AIDS patient a treatment that "can't hurt" and might help so that future patients will have access to treatment whose effectiveness is proven? The pressure for untested therapies for AIDS has now been eased somewhat by the development of new drugs that have been found effective in clinical trials.

The use of bleeding by 18th century physicians as a treatment for almost any illness is well known. The argument for this therapy appears foolish to us today, but the absence of curative power was not obvious to the people of the time. Similarly, tonsillectomies were performed on more than half of all children in the 1930s through the 1950s in the belief that the operation prevented rheumatic fever and other complications of strep throat. In fact, there was no evi-

dence for this benefit. It is now believed that a tonsillectomy may make strep infection more difficult to diagnose and treat.¹⁵ Unfortunately, it is difficult to do a randomized controlled study on a treatment that is already in wide use, because people do not want to risk being randomized to a placebo treatment if they suspect the active therapy is effective.

Such was the case in the 1990s with bone marrow transplant as a treatment for advanced breast cancer. With conventional chemotherapy, a patient had a 40 percent to 45 percent chance of living for five more years. A procedure that removed a woman's bone marrow, administered a much higher dose of chemotherapy than usual, and then replaced the bone marrow, in theory gave her a better chance of surviving the cancer. However, the procedure was itself arduous and risky, subjecting the woman to a 5 percent chance of dying of complications of the treatment. It was also expensive, costing up to \$200,000. The National Cancer Institute sponsored three large national trials of bone marrow transplants for breast cancer. The trials required women to be randomly assigned to the transplant group or to conventional therapy. Many were reluctant to participate in a trial because they wanted the most aggressive treatment, perceiving that this offered them their last best hope for survival. There were questions whether it was ethical to deny women the chance to choose the procedure, forcing them into a trial. On the other hand, might the practice of offering the transplant outside of a trial be unethical because surgeons and hospitals have a conflict of interest, perhaps influencing patients to choose a treatment from which they—the surgeons and hospitals—stand to profit financially? Insurance companies were forced through lawsuits and political pressure to pay for these expensive and arduous procedures without evidence that they saved lives.¹⁶

Fortunately, enough women ultimately enrolled in clinical trials in the United States and in other countries to test the hypothesis. Negative results began to appear in 1999, and an analysis of results from several studies published in 2004 strongly suggested that the intensive procedure did not lead to better survival for women who underwent it and, in fact, led to more treatment-related deaths and adverse side effects than suffered by the controls. As the authors of *False Hope: Bone Marrow Transplant for Breast Cancer*, point out, 23,000 to 40,000 American women with breast cancer had the procedure done outside of clinical trials, while only 1000 were recruited to participate in the clinical trials.¹⁷ "Although there was no deliberate effort to deceive women," they write, "the combined effect of salesmanship by physicians, lawyers, legislators, entrepreneurs, and the press led one of our respondents to say, 'We were all sold a bill of goods.'"^{17(p.286)} If the clinical trials had not been completed, bone marrow transplant might have become the standard treatment, although, like bleeding and tonsillectomies, it appears to do more harm than good.

Conflicts of Interest in Drug Trials

Epidemiologic studies are complicated enough, with many opportunities to make honest errors in interpreting them (as described earlier in this chapter), but when millions of dollars are at stake, which is the case with clinical trials of new prescription drugs, it is increasingly obvious that conflicts of interest often affect reported results. Randomized controlled trials are required by the U.S. Food and Drug Administration (FDA) before any new drug can be approved for use in the United States. Pharmaceutical companies conduct these studies to establish the safety and efficacy of a drug and submit the results to the FDA in search of the agency's approval. Often, the results of these studies are also submitted for publication to medical journals; such a publication in a reputable journal adds to the credibility of a drug's effectiveness.

Because randomized controlled trials are considered the best way to test drugs, and because FDA scientists review the results of the companies' studies, FDA approval was generally considered evidence that a drug was indeed safe and effective. However, in the late 1990s and early 2000s a rash of publicity about harm caused by FDA-approved drugs raised questions about the clinical trials that supported their approval. Some of these drugs were removed from the market after news of their harmful side effects came out; others were required to post "black-box warnings" on their packaging, indicating that they should be prescribed with caution. Drugs that were potentially hazardous included the arthritis drugs Vioxx and Bextra and the diabetes drug Avandia, which raised the risk of heart attacks; the cholesterol-lowering drug Baycol, which caused muscle damage; the antibiotic Cipro, which was sometimes associated with the rupture of tendons; the asthma drugs Serevent and Advair, which may lead to severe, sometimes fatal, asthma attacks; and the psychotropic drug Paxil, which increases the risk of suicidal behavior in children and young people.¹⁸

Harmful side effects may be missed in a clinical trial because they are rare and the number of subjects studied was too small for them to be noted. However, the case of Vioxx demonstrated that a company may purposely suppress negative information about a drug during the approval process. In fact, there is now evidence that companies may purposely bias their studies in ways that make them appear safer and more effective than they are.

Vioxx was the first of a new class of drugs called COX-2 inhibitors to be introduced in the late 1990s. These drugs are a class of nonsteroidal anti-inflammatory drugs (NSAIDs), used for pain relief—especially arthritis pain—and are designed to be less irritating to the digestive system than the established, over-the-counter NSAIDs, such as aspirin, ibuprofen, and naproxen.

Soon after Vioxx was approved by the FDA, the *New England Journal of Medicine* published a report of a clinical trial conducted by drug company scientists that had found a 50 percent reduction of serious gastrointestinal side effects in patients taking Vioxx compared with those taking naproxen.¹⁹ The same article reported that Vioxx caused a five-fold increase in the risk of

heart attacks and strokes, but the drug company, Merck, claimed that this was because naproxen protected the heart, as aspirin was known to do. Meanwhile, Pfizer introduced its own COX-2 inhibitors, Celebrex and Bextra. There were high hopes for these drugs, which were also being studied for prevention of colon cancer and Alzheimer's disease. However, the evidence mounted that all the COX-2 inhibitors increased the risk of heart attacks. A later study found that naproxen was not protective of the heart, although it was not harmful either.²⁰ In 2004, Merck removed Vioxx from the market; Bextra was withdrawn in 2005. Celebrex, and several newer COX-2 inhibitors, are still being sold, although they are required to carry warnings of cardiovascular risk.

These events raised many questions, however, about the way the clinical trials were conducted and reported. *The New England Journal of Medicine* in 2005 published an "Expression of Concern" accusing the Merck authors of providing misleading information in the 1999 article.²¹ Information that came out during lawsuits by patients who had been harmed by Vioxx revealed that the scientists knew of three heart attacks and other cardiovascular problems among the subjects taking the drug but had not included them in the data submitted to the journal.

It turns out that there are many tricks used by the pharmaceutical industry to prejudice the conclusions of clinical trials. Marcia Angell, a former editor of the *New England Journal of Medicine* describes them in her 2004 book, *The Truth About the Drug Companies: How They Deceive Us and What To Do About It*.²² She lists seven strategies the industry uses to bias research. One of the most common is to test a new drug in a clinical trial against a placebo. This seems reasonable, but the results may be misleading if there are older, well-established drugs already in use for the same condition. The new drug will inevitably be more expensive than the older ones—a benefit to the company—but there is no benefit for patients unless the new drug works better, something the trial does not test.

Drug companies use financial influence to ensure that physician-researchers come up with results favorable to the companies. In the extreme case, companies sometimes design clinical trials and seek academic scientists to carry them out, paying the scientists for their work; then the company analyzes and interprets the results and decides what should be published. Even when the scientists conduct their own research, they may be paid as consultants to companies whose products they are studying, or they may become paid members of advisory boards or speakers' bureaus, or they may own stock in the company. These arrangements tend to bias the researchers in favor of the companies' products. One survey found that industry-sponsored research was nearly four times as likely to be favorable to the company's product than NIH-sponsored research.^{22, p.106}

Until recently, when a company sponsored a study, it often had the last word on whether the results could be published at all. This led to strong publication bias: trials with positive results were published, while those with negative results were never revealed. In fact, this tendency was reinforced by the preference of medical journals, which tend not to be interested in publishing articles about treatments that don't work. Beginning in 2005, many reputable journals have adopted a policy of refusing to publish reports of clinical trials unless they had been registered at the beginning in a database of clinical trials, meaning that negative results could not be hidden. The 2007 Food and Drug Administration Revitalization Act now requires registration of all such trials in a public database sponsored by the National Library of Medicine, clinicaltrials.gov.²³

Conclusion

Epidemiologic studies are susceptible to many sources of error. Confounding factors may influence the results, suggesting an association where none exists. Bias may be introduced in the selection of cases or controls, in the reporting of exposures or outcomes, or in the disproportionate loss to follow-up of exposed and unexposed groups. Nevertheless, epidemiology is the basic science of public health. It is the only science of disease that focuses on human experience.

Epidemiology cannot prove cause and effect. However, certain characteristics of well-designed studies can make them very convincing. Studies with large numbers of subjects are more likely to be valid than smaller studies. A strong measure of association between exposure and disease, in the form of a high relative risk or odds ratio, is likely to indicate a true cause-and-effect relationship. A dose-response relationship that shows increasing risks from higher exposures adds to the validity of a study. A known biological explanation for an association between an exposure and a disease makes epidemiologic evidence more convincing than in situations when there is no known mechanism.

While observational studies have little potential for harming people, many ethical questions have been raised about clinical trials. In response to well-publicized abuses of the past, clinical trials and many other epidemiologic studies are required to be approved by committees, called institutional review boards, which ensure that the subjects' rights are protected. Other ethical concerns have been raised about the availability of treatments that have not been tested in clinical trials. On the other hand, conflicts of interest in the clinical trials for testing safety and efficacy of new drugs, which are required of pharmaceutical companies, have raised questions about the integrity of the research. Drug companies, which have vast amounts of money at stake in the outcomes of these trials, have found ways to manipulate the research to make drugs look better than they are.

Despite its flaws, epidemiology is still of necessity the basic science of public health. Epidemiologic data, when confirmed by repeated, well-designed studies and supported by the results of biomedical experiments in the laboratory, provide the best certainty as to the causes and cures of human disease.

References

1. G. Taubes, "Epidemiology Faces Its Limits," *Science* 269 (1995): 164–169.
2. A. Aschengrau and G. R. Seage III, *Essentials of Epidemiology in Public Health* (Sudbury, MA: Jones and Bartlett, 2003), 286.
3. S. Hite, *Women and Love: A Cultural Revolution in Progress* (New York: Alfred A. Knopf, 1987), described in V. Cohn and L. Cope, *News and Numbers: A Guide to Reporting Statistical Claims and Controversies in Health and Other Fields*, 2nd ed., (Ames: Iowa State University Press, 2001), 154–156.
4. R. Seltser and P. E. Sartwell, "The Influence of Occupational Exposure to Radiation on the Mortality of American Radiologists and Other Medical Specialists," *American Journal of Epidemiology* 81 (1965): 2–22.
5. R. A. Wilson, *Feminine Forever* (New York: M. Evans and Company, 1966).
6. G. Kolata, "Hormone Studies: What Went Wrong?" *The New York Times*, April 22, 2003.
7. J. E. Manson and K. A. Martin, "Postmenopausal Hormone-Replacement Therapy," *New England Journal of Medicine*, 345 (2001): 34–40.
8. G. Kolata, "Citing Risks, U.S. Will Halt Study of Drugs for Hormones," *The New York Times*, July 9, 2002.
9. G. Kolata, "Reversing Trend, Big Drop Is Seen in Breast Cancer," *The New York Times*, December 15, 2006.
10. F. Grodstein et al., "Understanding the Divergent Data on Postmenopausal Hormone Therapy," *New England Journal of Medicine* 348 (2003): 645–650.
11. M. Stampfer, "Commentary: Hormones and Heart Disease: Do Trials and Observational Studies Address Different Questions?" *International Journal of Epidemiology* 33 (2004): 454–455.
12. G. E. Pence, *Classic Cases in Medical Ethics: Accounts of the Cases That Shaped and Define Medical Ethics*, 5th ed. (Boston: McGraw-Hill, 2008).
13. S. M. Reverby, "More than Fact and Fiction: Cultural Memory and the Tuskegee Syphilis Study," *Hastings Center Report* (2001) September–October: 22–28.
14. Steering Committee of the Physicians' Health Study Research Group, Final Report on the Aspirin Component of the Ongoing Physicians' Health Study, *New England Journal of Medicine* 321 (1989): 129–135.
15. E. Braunwald et al., eds., *Harrison's Principles of Internal Medicine* (New York: McGraw-Hill, 1987).
16. G. Kolata, "Women Resist Trials to Test Marrow Transplants," *The New York Times*, February 15, 1995.
17. R. Rettig et al., *False Hope: Bone Marrow Transplantation for Breast Cancer* (Oxford, UK: Oxford University Press, 2007).

18. U.S. Food and Drug Administration, "Public Health Advisories (Drugs)," www.fda.gov/Drugs/DrugSafety/PublicHealthAdvisories/default.htm, last updated January 29, 2010, accessed January 31, 2010.
19. C. Bombardier et al., "Comparison of Upper Gastrointestinal Toxicity of Rofecoxib and Naproxen in Patients with Rheumatoid Arthritis," *New England Journal of Medicine* 343 (2000): 1520–1528.
20. D. J. Graham, "COX-2 Inhibitors, Other NSAIDs, and Cardiovascular Risk: The Seduction of Common Sense," *Journal of the American Medical Association* 296 (2006): 1653–1656.
21. G. D. Curfman et al., "Expression of Concern: Bombardier et al., 'Comparison of Gastrointestinal Toxicity of Rofecoxib and Naproxen in Patients with Rheumatoid Arthritis,'" *New England Journal of Medicine* 353 (2005): 1520–1528.
22. M. Angell, *The Truth About the Drug Companies: How They Deceive Us and What To Do About It* (New York: Random House, 2004).
23. J. M. Drazen et al., "Open Clinical Trials," *New England Journal of Medicine* 357 (2007): 1756–1757.