The NEW ENGLAND JOURNAL of MEDICINE

Perspective

NOVEMBER 23, 2006

Dangerous Deception — Hiding the Evidence of Adverse Drug Effects

Jerry Avorn, M.D.

September 30 is becoming a day of infamy for drug safety. On that date in 2004, Merck announced that rofecoxib (Vioxx) doubled the risk of myocardial infarction and stroke, and the company withdrew the drug from the market after 5 years of use in more than 20 million patients. On September 30, 2006, a front-page article in the *New York Times* reported that the Food and Drug Administration (FDA) had issued a warning that the antifibrinolytic drug aprotinin, widely used to reduce perioperative bleeding in patients undergoing cardiac surgery, could cause renal failure, congestive heart failure, stroke, and death.

Some experts had been concerned about aprotinin (Trasylol) ever since its approval in 1993.[1] As Hiatt explains in his Perspective article in this issue of the *Journal* (pages 2171–2173), one of two epidemiologic studies reported early this year provided support for this concern. In an observational study involving 4374 patients who underwent coronary revascularization,[2] Mangano et al. found that patients who were given aprotinin had an incidence of postoperative renal failure requiring dialysis that was more than twice that among patients who received different agents. Among patients undergoing uncomplicated coronary-artery surgery, those given aprotinin had a 55% increase in the incidence of myocardial infarction or heart failure and a 181% increase in the incidence of stroke or encephalopathy. The authors advised against further use of the drug, since safer, cheaper alternatives are available.

After the study was published, the FDA moved to convene an advisory committee to reassess the drug's safety and assembled the relevant data. The committee met on September 21, reviewed the available evidence, and concluded that there was no need for additional warnings on the drug's official labeling.

What put aprotinin on the front page on September 30, however, was the revelation that its manufacturer, Bayer, had hired a private contract research organization to perform its own large observational study of postoperative complications in patients given the drug. The analysis, completed in time for the FDA meeting, reached conclusions similar to those of Mangano et al. It, too, adjusted for a

wide variety of clinical characteristics and showed that patients who received aprotinin had higher mortality rates and substantially more renal damage than those given other treatments. But neither Bayer nor its contractor had provided the report to the FDA or even acknowledged its existence before the meeting.

Many aspects of the aprotinin saga are familiar to observers of the drug-evaluation process: a product is approved because it is more effective than placebo, worries emerge about its safety, few or no adequately powered controlled trials are conducted to address these issues, and payers spend huge sums on the drug, despite the dearth of evidence that it is better than older, cheaper agents. The health care system has a hard time performing drug-safety analyses, in large part because it relies on the pharmaceutical industry to conduct most research on the risks and benefits of medications. It is naive to expect companies to voluntarily fund studies that could sink lucrative products, the FDA lacks the regulatory clout to require them, and despite the $220 billion we spend on drugs each year, we apparently can't find the resources to provide public support for these studies, even if the results could be of great clinical importance and save millions of dollars. Although a large randomized trial would have provided a valid means of comparing aprotinin with other treatments, no such study has been undertaken on the necessary scale.

The study by Mangano et al. was observational — its subjects were not randomly assigned to the four study groups. Instead, the investigators reviewed numerous variables for each patient and used a propensity score and multivariable methods to adjust for underlying differences among the groups. Although this approach has important limitations, observational studies often provide the only data available for evaluating critical safety issues.

A confirmatory observational study would have lent key support to the conclusions of Mangano et al. — if its findings had been aired. Bayer has admitted that its suppression of the study was "a mistake," but this is not the first time the company has behaved in this manner. When Bayer was accused of hiding data unfavorable to its cholesterol-lowering drug cerivastatin (Baycol) before it was taken off the market in 2001, litigation uncovered a memorandum from a company executive arguing against performing a study of its risk. "If the FDA asks for bad news, we have to give," read the memo, "but if we don't have it, we can't give it to them."[3]

Other companies have behaved similarly. Although Merck steadfastly denied that Vioxx increased the risk of myocardial infarction while the drug was on the market, it commissioned two epidemiologic studies of the relationship. My colleagues and I performed one of the studies, but when it confirmed an increased risk, Merck dismissed the findings and assailed the methods that it had previously accepted.[4] The second study (by the same contract research group involved in the aprotinin affair) also confirmed the association, but its results were not made public until after the drug had been withdrawn from the market.

The problem is not limited to observational studies. A few years ago, it was discovered that some companies had funded multiple clinical trials of their selective serotonin-reuptake inhibitor antidepressants but reported the results of only the favorable trials — distorting the evidence base physicians use in choosing drugs. But the issue is thornier for epidemiologic analyses. Companies can conduct them secretly, even in-house, with the use of a pur-

chased proprietary database, making the results even easier to conceal.

Carefully performed observational studies may provide the best information available about side effects, but propensity scores and other multivariable techniques applied to epidemiologic research cannot always control for all the inevitable selection bias, making the transparency of methods and raw data even more important than in randomized trials. Rather than yielding "virtual randomized trials," the methods available for controlling confounding in observational research can sometimes look better than they work.[5] Thus, these studies can inform our understanding only after their methods have been scrutinized closely, fairly, and objectively — but only if the data are available.

On September 30, 1982, six people in the Chicago area died after taking acetaminophen (Tylenol) that had been laced with cyanide. The tragedy riveted the country's attention for months. We should be able to muster at least a fraction of that concern to address more clinically relevant adverse drug effects that could sicken or kill thousands of patients. How can we capture such interest in less sensational problems of medication safety? A good start would be to make a national commitment to publicly supported studies of drug risks so that no company could take possession of critical findings for its own purposes. The results of that research could be discussed openly at an annual conference on the risks and benefits of drugs. To keep everyone's attention focused on medication safety, perhaps the conference could be held every year on September 30.

An interview with Dr. Avorn can be heard at www.nejm.org.

Dr. Avorn is a professor of medicine at Harvard Medical School and chief of the Division of Pharmacoepidemiology and Pharmacoeconomics at Brigham and Women's Hospital — both in Boston.

1. Statement on aprotinin. Rockville, MD: Food and Drug Administration, December 30, 1993. (Accessed November 2, 2006, at http://www.fda.gov/bbs/topics/NEWS/NEW00453.html.)
2. Mangano DT, Tudor IC, Dietzel C. The risk associated with aprotinin in cardiac surgery. N Engl J Med 2006;354:353-65.
3. Berenson A. Trial lawyers are now focusing on lawsuits against drug makers. New York Times. May 18, 2003.
4. Burton TM. Merck takes author's name off Vioxx study. Wall Street Journal. May 18, 2004.
5. Sturmer T, Schneeweiss S, Brookhart MA, Rothman KJ, Avorn J, Glynn RJ. Analytic strategies to adjust confounding using exposure propensity scores and disease risk scores. Am J Epidemiol 2005;161:891-8.

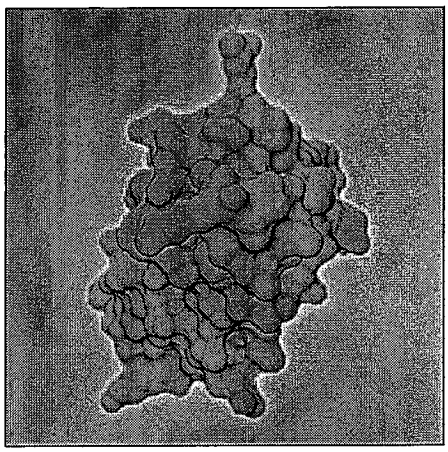
Observational Studies of Drug Safety — Aprotinin and the Absence of Transparency

William R. Hiatt, M.D.

The full safety profile of a new drug is rarely known at the time of approval by the Food and Drug Administration (FDA). Most drug-development programs designed for treatments of symptomatic indications are underpowered to detect any increased risk of rare drug reactions or change in background event rates attributable to the drug. Large, post-marketing, randomized, controlled trials provide robust data on drug safety but may be subject to multiple sources of bias. Observational studies of a drug's effects in clinical practice can offer additional information on risks. The recent discussions of aprotinin (Trasylol, Bayer) by the Cardiovascular and Renal Drugs Advisory Committee of the FDA, which I chair, provide insight into the strengths and weaknesses of using observational data to assess drug safety and highlight the importance of using a transparent and open process when reviewing such data.

Post-marketing observational studies permit the evaluation of drug safety in a large number of patients in a real-world setting, where practice patterns, including off-label use, can be assessed. Such studies have limited ability to determine causation, but they can detect signals that may suggest a safety concern.

In an observational study, decisions to use specific drugs are made by local physicians, according to their perceptions of the risks and benefits for particular patients. Such treatment allocation results in an imbalance in demographic and risk factors between patients given the drug of interest and those given an alternative or no treatment. An imbalance in various clinical factors will directly influence safety outcomes when patients at higher risk receive the drug. Therefore, appro-



The Aprotinin Molecule.

Source of the aprotinin molecule data (Protein Data Bank number 2PTC): Research Collaboratory for Structural Bioinformatics Protein Data Bank (www.pdb.org).

priate statistical methods must be used to adjust for the nonrandom assignment to treatment groups.[1]

One way to address imbalances between groups is to use a propensity score that incorporates confounders and other covariates into a model predicting the probability of assignment to a particular treatment. This score can be used for adjustment or for matching patients who have similar probabilities of receiving a treatment. Differences in outcomes between treated and untreated patients (or patients treated with a comparison drug) with equal propensity scores provide a less biased estimate of treatment effect.

Aprotinin was approved by the FDA in 1993 as a means of reducing perioperative blood loss and the need for blood transfusion in patients undergoing coronary-artery bypass grafting. Neither the clinical trial database that led to approval nor the numerous randomized, controlled trials conducted after approval identified an association between aprotinin and any short-term increase in the risk of death or nonfatal cardiovascular events or any serious renal toxic effects (except for a transient increase in the serum creatinine concentration). However, in early 2006, two observational studies were published that raised serious concerns about the drug's safety.[2,3]

One study, by Mangano et al.,[2] evaluated 4374 patients undergoing coronary-artery bypass surgery. End points were prospectively defined, and data on a large number of clinical variables were collected for each patient. The decision to use aprotinin, aminocaproic acid, or tranexamic acid to inhibit fibrinolysis or to withhold antifibrinolytic therapy was made by the treating physician. Since group assignment was not random, it is not surprising that patients assigned to the group receiving aprotinin were at much higher risk for adverse cardiovascular outcomes than were the other patients in the study, who were treated with alternative therapies. When a propensity score was used in a logistic-regression model to adjust for these baseline differences, aprotinin was associated with a nonsignificant increase in the risk of death (odds ratio, 1.59; 95% confidence interval [CI], 0.76 to 3.34) and with significant increases in the risks of renal events (odds ratio, 2.34; 95% CI, 1.27 to 4.31), cardiovascular events (odds ratio, 1.42; 95% CI, 1.09 to 1.86), and cerebrovascular events (odds ratio, 2.15; 95% CI, 1.14 to 4.06). The authors concluded that "the observed association between aprotinin and serious end-organ damage indicates that continued use is not prudent."

The second study, by Karkouti et al.,[3] compared the risks associated with aprotinin with those associated with tranexamic acid in 898 patients undergoing high-risk cardiac surgery. The authors used a propensity score to match patients who had been given the different treatments — an approach that permitted the identification of a population of patients who were subsequently fully matched according to a large number of covariates. Calculations from the study data, performed by the Colorado Prevention Center, indicated no significant difference between the two treatments in the overall risk of myocardial infarction (odds ratio, 1.20; 95% CI, 0.52 to 2.75), stroke (odds ratio, 1.15; 95% CI, 0.56 to 2.40), or death (odds ratio, 0.91; 95% CI, 0.56 to 1.46). However, among patients who had abnormal renal function at baseline, those given aprotinin had a significantly increased risk of postoperative renal dysfunction (odds ratio, 1.69; 95% CI, 1.07 to 2.69). Because this study included fewer patients than the study by Mangano et al., it had less power to detect a safety problem.

These observational studies did not permit a definitive conclusion about cardiovascular or renal risk, but they did raise concern. The FDA appropriately issued a public health advisory about the potential risks posed by aprotinin and urged physicians to monitor their patients carefully for renal, cardiac, and cerebral toxic effects.[4] The agency also convened a meeting of the Cardiovascular and Re-

nal Drugs Advisory Committee on September 21, 2006, to evaluate the evidence on the cardiovascular and renal toxic effects of aprotinin. Our committee reviewed the published studies and the global safety and efficacy database submitted by Bayer, which included 45 randomized, controlled trials involving 2249 patients who received aprotinin and 2164 who received placebo.

The Bayer safety analysis was confined to the unadjudicated adverse events reported within 7 days after the administration of aprotinin, which included 520 deaths or serious cardiovascular or renal events. The results did not reveal any increased risk of fatal or nonfatal cardiovascular events. According to the Colorado Prevention Center's calculations based on Bayer's data, the odds ratio for death was 1.14 (95% CI, 0.80 to 1.62); for myocardial infarction, 1.17 (95% CI, 0.93 to 1.49); for stroke, 0.71 (95% CI, 0.42 to 1.18); and for renal failure, 1.15 (95% CI, 0.74 to 1.78). Two studies in which myocardial infarctions were adjudicated by an independent review committee showed a nonsignificant increase in the risk of myocardial infarction, with a point estimate of the odds ratio as high as 2.24 (95% CI, 0.56 to 9.00).

How can the results of these trials and the study by Karkouti et al. appear to be so different from the findings of Mangano et al.? Part of the problem may derive from the inability of observational studies to identify and measure all relevant covariates that may influence the outcome; similar studies using a few different covariates can sometimes come to opposite conclusions.

FDA advisory committee meetings usually include an independent review of data — either those provided by the sponsor or those used by the authors of relevant published studies. This independent and transparent process becomes even more critical in the evaluation of observational studies, whose conclusions can be highly influenced by the statistical methods used and the inability to identify all confounding factors. Unfortunately, Mangano did not give the FDA or the committee full access to his data, which would have allowed the agency to perform an independent review and analysis to validate his group's findings. Although there are many legitimate concerns with regard to the sharing of data (for instance, confidentiality and the need for informed consent), the lack of independent review greatly limited the committee's ability to draw conclusions from the study. Having reviewed all the data available, the committee decided that there was insufficient evidence to require an additional warning on aprotinin's labeling and agreed that the clinical data supported an acceptable safety and efficacy profile for aprotinin.

Days after the committee meeting, the FDA was made aware of additional observational data from the sponsor that had not been presented at the meeting. Bayer evidently had commissioned an observational study involving 67,000 patients who were given aprotinin.[5] According to the initial FDA review of data from that study, aprotinin may be associated with "increased risk for death, kidney failure, congestive heart failure and stroke." The failure of Bayer to disclose all its data on aprotinin seriously undermined the advisory committee process and hindered the safety review.

Although observational studies cannot be definitive (and so should evoke measured responses), they can provide important new safety information that can direct further scientific and regulatory actions — if their findings can be confirmed. Since further analysis of the new data provided by Bayer is ongoing at the FDA, conclusions about the overall safety of aprotinin cannot be drawn at this time. Still, the aprotinin story demonstrates that full disclosure and a transparent process are essential in evaluating the findings of all studies germane to drug safety and the public health.

Dr. Hiatt is a professor of medicine at the University of Colorado School of Medicine and president of the Colorado Prevention Center — both in Denver — and is the current chair of the Cardiovascular and Renal Drugs Advisory Committee of the FDA. Opinions expressed in this article are those of the author and do not necessarily reflect the opinions of the FDA or the advisory committee.

1. Rubin DB. Estimating causal effects from large data sets using propensity scores. Ann Intern Med 1997;127:757-63.
2. Mangano DT, Tudor IC, Dietzel C. The risk associated with aprotinin in cardiac surgery. N Engl J Med 2006;354:353-65.
3. Karkouti K, Beattie WS, Dattilo KM, et al. A propensity score case-control comparison of aprotinin and tranexamic acid in high-transfusion-risk cardiac surgery. Transfusion 2006;46:327-38.
4. Food and Drug Administration. Public health advisory: aprotinin injection (marketed as Trasylol). Updated September 29, 2006. (Accessed November 2, 2006, at http://www.fda.gov/cder/drug/advisory/aprotinin.htm.)
5. *Idem.* FDA statement regarding new Trasylol data. September 29, 2006. (Accessed November 2, 2006, at http://www.fda.gov/bbs/topics/NEWS/2006/NEW01472.html.)