

Users' Guides to the Medical Literature

IV. How to Use an Article About Harm

Mitchell Levine, MD, MSc; Stephen Walter, PhD; Hui Lee, MD, MSc; Ted Haines, MD, MSc; Anne Holbrook, MD, PharmD, MSc; Virginia Moyer, MD, MPH; for the Evidence-Based Medicine Working Group

CLINICAL SCENARIO

You are having lunch in the hospital cafeteria when one of your colleagues raises the issue of the safety of β -adrenergic agonists in the treatment of asthma. Your colleague feels uncertain about how to respond to patients asking him about media reports of an increased risk of death associated with these medications. Another colleague mentions a key article on this topic that generated much of the publicity, but she cannot recall the details. You all agree that this is an issue that arises frequently enough in your practices that you should become familiar with the evidence contained in the article that your patients have heard about. You volunteer to search the literature for the key article

and report back to your colleagues in the next few days.

THE SEARCH

The next day you do a MEDLINE search using the following terms: asthma (MH) (MH stands for MeSH heading, indexing terms used by National Library of Medicine personnel); adrenergic beta receptor agonists (MH); adverse effects (SH) (SH stands for Subheading). You limit the search to *Abridged Index Medicus* journals knowing that you will likely find the article your colleague recalled seeing within this list of major medical journals. Your MEDLINE search (1990 through 1993) identifies 38 citations. There were nine original studies, seven review articles, and 22 letters, editorials, and commentaries. Of the nine original articles, only one is an epidemiologic study assessing the risk of death associated with inhaled β -adrenergic agonists, and you think this is the article to which your colleague referred. The study describes a 2.6-fold increased risk of death from asthma associated with the use of β -adrenergic agonist metered-dose inhalers.¹

INTRODUCTION

Clinicians often encounter patients who may be facing harmful exposures, either to medical interventions or environmental agents. Are pregnant women at increased risk of miscarriage if they work in front of video display terminals? Do vasectomies increase the risk of prostate cancer? Do hypertension management programs at work lead to increased absenteeism? When examining these questions, physicians must evaluate the validity of the data, the strength of the association between the

putative cause and the adverse outcome, and the relevance to patients in their practice (Table 1).

This article in our series of "Users' Guides to the Medical Literature" will help you evaluate an individual article assessing an issue of harm. To fully assess the cause-and-effect relationship implied in any question of harm requires consideration of all the information available. Systematic overviews (eg, meta-analyses) can provide an objective summary of all the available evidence, and we will deal with how to use an overview in a subsequent article in this series. Using such an overview requires a prior understanding of the rules of evidence for individual studies, and this article covers the basic rules for observational (nonrandomized) studies.

ARE THE RESULTS OF THE STUDY VALID?

Primary Guides

Were There Clearly Identified Comparison Groups That Were Similar With Respect to Important Determinants of Outcome Other Than the One of Interest?—In a study that identifies a harmful exposure, the choice of comparison groups has an enormous influence on the credibility of the results. Because the design of the study determines the comparison groups, we will review the basic study designs that clinicians encounter when assessing whether their patients have been or might be exposed to a potentially harmful factor (Table 2).

Randomized Controlled Trials.—A randomized controlled trial (RCT) is a true experiment in which patients are assigned, by a mechanism analogous to

From the Departments of Clinical Epidemiology and Biostatistics (Drs Levine, Walter, and Haines), Medicine (Drs Lee and Holbrook), the Occupational Health Program (Dr Haines), McMaster University, Hamilton, Ontario; and the Department of Pediatrics, University of Texas, Houston (Dr Moyer).

A complete list of members (with affiliations) of the Evidence-Based Medicine Working Group appears in the first article of this series (*JAMA*. 1993;270:2093-2095). The following members contributed to this article: Gordon Guyatt (chair), MD, MSc; Eric Bass, MD, MPH; George Browman, MD, MSc; Deborah Cook, MD, MSc; Michael Farkouh, MD; Hertzell Gerstein, MD, MSc; Ted Haines, MD, MSc; Brian Haynes, MD, MSc, PhD; Robert Hayward, MD, MPH; Anne Holbrook, MD, PharmD, MSc; Roman Jaeschke, MD, MSc; Elizabeth Juniper, MCSP, MSc; Andreas Laupacis, MD, MSc; Hui Lee, MD, MSc; Mitchell Levine, MD, MSc; Virginia Moyer, MD, MPH; David Naylor, MD, DPh; Jim Nishikawa, MD; Andrew Oxman, MD, MSc, FACPM; Ameen Patel, MD; John Philbrick, MD; Scott Richardson, MD; Stephane Sauve, MD, MSc; David Sackett, MD, MSc; Jack Sinclair, MD; Brian Strom, MD, MPH; K.S. Trout, FRCE; Sean Tunis, MD, MSc; Stephen Walter, PhD; John Williams Jr, MD, MHS; and Mark Wilson, MD, MPH.

Reprint requests to Room 2C12, McMaster University Health Sciences Centre, 1200 Main St W, Hamilton, Ontario, Canada, L8N 3Z5 (Dr Guyatt).

Table 1.—User's Guides for an Article About Harm

Are the results of the study valid?
Primary guides:
Were there clearly identified comparison groups that were similar with respect to important determinants of outcome, other than the one of interest?
Were the outcomes and exposures measured in the same way in the groups being compared?
Was follow-up sufficiently long and complete?
Secondary guides:
Is the temporal relationship correct?
Is there a dose-response gradient?
What are the results?
How strong is the association between exposure and outcome?
How precise is the estimate of the risk?
Will the results help me in caring for my patients?
Are the results applicable to my practice?
What is the magnitude of the risk?
Should I attempt to stop the exposure?

a coin flip, to either the putative causal agent or some alternative experience (either another agent or no exposure at all). Investigators then follow the patients forward in time and assess whether they have experienced the outcome of interest. The great strength of the RCT is that we can be confident that the study groups were similar not only with respect to determinants of outcome that we know about, but also those we do not know about.

In prior articles in this series, we have shown how readers of articles about therapy can use the results of RCTs.^{2,3} Randomized controlled trials are rarely done to study possible harmful exposures, but if a well-designed RCT demonstrates an important relationship between an agent and an adverse event, clinicians can be confident of the results. For instance, the Cardiac Arrhythmia Suppression Trial is an RCT that demonstrated an association between the antiarrhythmic agents encainide, flecainide, and moricizine, and excessive mortality.^{4,5} As a result, clinicians have curtailed their use of these drugs and have become much more cautious in using other antiarrhythmic agents in the treatment of nonsustained ventricular arrhythmias.

Cohort Studies.—When it is either not feasible or not ethical to randomly assign patients to be exposed or not exposed to a putative causal agent, investigators must find an alternative to an RCT. In a cohort study, the investigator identifies exposed and nonexposed groups of patients and then follows them forward in time, monitoring the occurrence of the outcome. You can appreciate the practical need for cohort studies when subjects cannot be “assigned” to an exposure group, as occurs when one wants to evaluate the effects of an occupational exposure. For example, investigators assessed perinatal outcomes among children of men exposed to lead and organic solvents in the printing in-

Table 2.—Directions of Inquiry and Key Methodologic Strengths and Weaknesses for Different Study Designs

Design	Starting Point	Assessment	Strengths	Weaknesses
Randomized controlled trial	Exposure status	Adverse event status	Internal validity	Feasibility, generalizability
Cohort	Exposure status	Adverse event status	Feasible when randomization of exposure not possible	Susceptible to threats to internal validity
Case control	Adverse event status	Exposure status	Overcomes temporal delays, may only require small sample size	Susceptible to threats to internal validity

dustry using a cohort of all males who had been members of printers' unions in Oslo, Norway, and on the basis of job classification, they categorized fathers as to their exposure to lead and solvents. In this study, exposure was associated with an eightfold increase in preterm births, but no significant impact on birth defects.⁶

Cohort studies may also be performed when harmful outcomes are infrequent. For example, clinically apparent upper gastrointestinal hemorrhage in nonsteroidal anti-inflammatory drug (NSAID) users occurs approximately 1.5 times per 1000 person years of exposure, in comparison with 1.0 per 1000 person years in those not taking NSAIDs (assuming a stable risk over time).⁷ An RCT to study this effect would require approximately 6000 patient-years of exposure to achieve a 95% probability of observing at least one additional serious gastrointestinal hemorrhage among treated patients, and a substantially larger sample size (approximately 75 000 patient-years per group) for adequate power to test the hypothesis that NSAIDs cause the additional hemorrhages.⁸ Such an RCT would not be feasible, but a cohort study, particularly one in which the information comes from a large administrative database, would be.

Because subjects in a cohort study select themselves (or are selected by a physician) for exposure to the putative harmful agent, there is no particular reason they should be similar to nonexposed persons with respect to other important determinants of outcome. It therefore becomes crucial for investigators to document the characteristics of the exposed and nonexposed subjects and either demonstrate their comparability or use statistical techniques to adjust for differences. In the association between NSAIDs and the increased risk of upper gastrointestinal bleeding, age is associated both with exposure to NSAIDs and with gastrointestinal bleeding, and is therefore called a possible “confounding variable.” In other words, since patients taking NSAIDs will be older, it may be difficult to tell if their increased risk of bleeding is because of their age or because of their NSAID

exposure. When such a confounding variable is unequally distributed in the exposed and nonexposed populations, investigators use statistical techniques that correct or adjust for the imbalances.

Even if investigators document the comparability of potentially confounding variables in exposed and nonexposed cohorts or use statistical techniques to adjust for differences, there may be an important imbalance in prognostic factors that the investigators don't know about or have not measured that may be responsible for differences in outcome. It may be, for instance, that illnesses that require NSAIDs, rather than the NSAIDs themselves, are responsible for the increased risk of bleeding. Thus, the strength of inference from a cohort study will always be less than that of a rigorously conducted RCT.

Case-Control Studies.—When the outcome of interest either is very rare or takes a long time to develop, cohort studies also may not be feasible. Investigators may use an alternative design in which they identify cases, patients who have already developed the outcome of interest (eg, a disease, hospitalization, death). The investigators then choose controls, persons who do not have the outcome of interest, but who are otherwise similar to the cases with respect to important determinants of outcome such as age, sex, and concurrent medical conditions. Investigators can then assess retrospectively the relative frequency of exposure to the putative harmful agent among the cases and controls. This observational design is called a case-control study.

Using a case-control design, investigators demonstrated the association between diethylstilbestrol ingestion by pregnant women and the development of vaginal adenocarcinoma in their daughters many years later.⁹ A prospective cohort study designed to test this cause-and-effect relationship would have required at least 20 years from the time when the association was first suspected until the completion of the study. Further, given the infrequency of the disease, a cohort study would have required hundreds of thousands of subjects. Using the case-control strategy, the investigators defined two groups of young

women—those who had suffered the outcome of interest (vaginal adenocarcinoma) were designated as the cases ($n=8$), and those who did not have the outcome, as the controls ($n=32$). Then, working backward in time, the exposure rates to diethylstilbestrol were determined for the two groups. Analogous to the situation with a cohort study, investigators had to ensure balance, or adjust for imbalances, in important risk factors in cases and controls (eg, intrauterine x-ray exposure). The investigators found a strong association between in utero diethylstilbestrol exposure and vaginal adenocarcinoma that was extremely unlikely to be attributable to the play of chance ($P<.00001$), without a delay of 20 years, and requiring only 40 women.

As with cohort studies, case-control studies are susceptible to unmeasured confounders. Therefore, the strength of inference that can be drawn from the results may be limited.

Case Series and Case Reports.—Case series and case reports do not provide any comparison group and are therefore unable to satisfy the requirements of the first primary guide. Although descriptive studies occasionally demonstrate dramatic findings mandating an immediate change in physician behavior (eg, thalidomide and birth defects), there are potentially undesirable consequences when actions are taken in response to weak evidence. Bendectin (a combination of doxylamine, pyridoxine, and dicyclomine used as an antiemetic in pregnancy) was withdrawn as a result of case reports suggesting it was teratogenic.¹⁰ Later, a number of comparative studies demonstrated the relative safety of the drug,¹¹ but they could not eradicate a litigious atmosphere that prompted the manufacturer to withdraw the drug from the market. Thus, many pregnant women who could have benefited were denied the symptom relief the drug could have offered. In general, clinicians should not draw conclusions about cause-and-effect relationships from case series, but recognize that the results may generate questions for regulatory agencies and clinical investigators to address.

Design Issues—Summary.—It is apparent that, just as for questions of therapeutic effectiveness, clinicians should look for RCTs to resolve issues of harm. It is also apparent that they will often be disappointed in this search, and must be satisfied with studies of weaker design. Whatever the design, however, they should look for an appropriate control population before making a strong inference about a putative harmful agent.

Were the Exposures and Outcomes Measured in the Same Way in the Groups Being Compared?—In case-control studies, ascertainment of the exposure is a key issue. Patients with leukemia, when asked about prior exposure to solvents, may be more likely to recall exposure than would a control group, either because of increased patient motivation (recall bias) or greater probing by an interviewer (interviewer bias). Clinicians should attend to whether investigators used strategies, such as blinding subjects and interviewers to the hypothesis of the study, to minimize bias. For example, in a case-control study describing the association between psychotropic drug use and hip fracture, investigators established drug exposure by examining computerized claims files of the Michigan Medicaid program, a strategy that avoided both recall and interviewer bias.¹² As a result, the clinician has more confidence in the study's findings of a twofold increase in the risk of hip fracture.

Exposure opportunity should also be similar among cases and controls. There is evidence suggesting a 2.7-fold increased risk of homicide among individuals keeping a gun in their home. It would be important to know that the control group had a similar opportunity for gun possession, otherwise the true risk could be different from the study results—increased if the controls had a greater opportunity, decreased if the controls had a lesser opportunity for gun possession.¹³

In RCTs and cohort studies, ascertainment of outcome is the key issue. Investigators have reported a threefold increase in risk of malignant melanoma in individuals working with radioactive materials. One possible explanation for some of the increased risk might be that physicians, aware of a possible risk, search more diligently and therefore detect disease that might otherwise go unnoticed (or detect disease at an earlier point in time). This could result in the exposed cohort having an apparent, but spurious, increase in risk—a situation we refer to as surveillance bias.¹⁴

Was Follow-up Sufficiently Long and Complete?—An additional point relating to the measurement of outcomes is the need for adequate follow-up in RCTs and cohort studies. As discussed in a previous article in this series,² patients unavailable for follow-up threaten the validity of the results because the patients who are unavailable may have very different outcomes from those available for assessment. The longer the follow-up period required, the greater the possibility that the follow-up will be incomplete.

In a well-executed study, investigators determined the vital status of 1235 of 1261 white males (98%) employed in chrysotile asbestos textile operation between 1940 and 1975. The relative risk (RR) for lung cancer death increased monotonically from 1.4 to 18.2 with cumulative exposure among asbestos workers with at least 15 years since first exposure.¹⁵ Because the 2% missing data were unlikely to affect the results and the follow-up was sufficiently long, the study allows relatively strong inference about the increase in cancer risk with asbestos exposure.

Secondary Guides

Is the Temporal Relationship Correct?—Does exposure to the harmful agent precede the adverse outcome? The reports of increased suicidal ideation associated with the use of the antidepressant fluoxetine illustrate the importance of this question.¹⁶ Did the thoughts of suicide occur after the fluoxetine was administered, or were the patients given this drug because they were already showing signs of clinical deterioration? A meta-analysis of controlled trials of treatment for depression did not confirm the apparent association.¹⁷

Is There a Dose-Response Gradient?—We are more confident attributing an adverse outcome to a particular exposure if, as the quantity or the duration of exposure to the putative harmful agent increases, risk of the adverse outcome also increases. The risk of dying from lung cancer in male physician smokers is dose-dependent; the risk increases by 50%, 132%, and 220% for one to 14, 15 to 24, and 25 or more cigarettes smoked per day, respectively.¹⁸

WHAT ARE THE RESULTS?

How Strong Is the Association Between Exposure and Outcome?—We have described the most common way of expressing an association between exposure and outcome, the RR, in detail in an earlier article in this series.³ In brief, the RR is the risk (or incidence) of the adverse effect in the exposed group divided by the risk of the adverse effect in the unexposed group. Values greater than 1 represent an increase in risk associated with the exposure, while values less than 1 represent a reduction in risk. To illustrate, in a cohort study assessing in-hospital mortality following noncardiac surgery in male veterans, 23 of 289 patients with a history of hypertension died, compared with three of 185 patients without hypertension. The RR of death for hypertensive men was 4.9.¹⁹ The RR tells us that death occurs almost five times more often in the hypertensive patients than in normotensive patients.

Table 3.—Estimate of Relative Risks and Odds Ratios for Exposed and Unexposed Patients

Patient	Adverse Event (Case)	No Adverse Event (Control)
Exposed	a	b
Not exposed	c	d

*Relative risk = $[a/(a+b)]/[c/(c+d)]$.
Odds ratio = $(a/c)/(b/d)$.

The estimate of RR depends on having samples of exposed and unexposed patients, where the proportion of the patients with the outcome of interest can be determined. The RR is therefore not applicable to case-control studies in which the number of cases and controls, and therefore the proportion of individuals with the outcome, is chosen by the investigator. For case-control studies, instead of using a ratio of risks, we use a ratio of odds: the odds of a case patient being exposed divided by the odds of a control patient being exposed. Using a simple 2×2 table, RRs and odds ratios (ORs) can be represented as depicted in Table 3.

When the outcome of interest is rare in the population from which the sample of cases was drawn, which is often the reason for using a case-control design in the first place, the OR closely approximates the RR.

When considering both study design and strength of association, we may be ready to interpret a small increase in risk as representing a true harmful effect when the study design is strong (such as an RCT). A much higher increase in risk might be required of weaker designs (such as cohort or case-control studies) as subtle findings are more likely to be because of subtle flaws in design. Very large values for RRs or ORs represent strong associations that are less likely to be caused by confounding or bias.

How Precise Is the Estimate of the Risk?—In a previous article in this series we have shown how the clinician can evaluate the precision of the estimate of treatment effect by examining the confidence interval (CI) around that estimate.³ The clinician can take the same approach with articles assessing risk. In a study in which the investigators have shown an association between an exposure and an adverse outcome, the lower limit of the estimate of RR associated with the adverse exposure provides a minimal estimate of the strength of the association. In a study where the investigators fail to demonstrate an association (a “negative” study), the upper boundary of the CI around the RR tells the clinician just how big an adverse effect may still be present, despite the failure to show a statistically significant association.

WHAT ARE THE IMPLICATIONS FOR MY PRACTICE?

Are the Results Applicable to My Practice?—If you are convinced that the results of the study are valid for the population that was studied, you then have to decide whether you can extrapolate the results to patients in your own practice. Are your patients similar to those described in the study with respect to morbidity, age, race, or other potentially important factors? Are there clinically important differences in the treatments or exposures between your patients and the patients studied? For example, the risk of thrombophlebitis associated with oral contraceptives described in the 1970s may not be applicable to the patient of the 1990s because of the lower estrogen doses currently in use. Similarly, increases in uterine cancer secondary to postmenopausal estrogen probably don't apply to women who are also taking concomitant progestins tailored to produce monthly withdrawal bleeding.

What Is the Magnitude of the Risk?—The RR and the OR do not tell us how frequently the problem occurs, only that the observed effect occurs more or less often in the exposed group compared with the unexposed group. Thus, the reader needs a method for assessing clinical importance. In our discussion of therapy we described how the clinician can calculate the number of patients she must treat to prevent an adverse event.³ When the issue is harm, the clinician can use data from an RCT or cohort study to make an analogous calculation to determine how many people must be exposed to the harmful agent to cause an adverse outcome. From the Cardiac Arrhythmia Suppression Trial, over an average of 10 months of follow-up, mortality was 3.0% and 7.7% for placebo and encainide/flecainide patients, respectively. The absolute risk increase was 4.7%, the reciprocal of which tells us that, on average, for every 21 patients we treat with encainide or flecainide for about a year, we will cause one excess death.⁴ This contrasts with NSAIDs and upper gastrointestinal bleeding. Of 2000 unexposed patients, two will suffer a hemorrhage each year. Of 2000 patients taking NSAIDs, three will suffer a hemorrhage each year. Thus, if we treat 2000 patients with NSAIDs, we can expect a single additional bleeding event.⁶

Should I Attempt to Stop the Exposure?—After evaluating the evidence that an exposure is harmful, determining subsequent actions may not be simple. There are at least three aspects the physician must consider in making a clinical decision.²⁰

One is the strength of inference; how strong was the study or studies that demonstrated harm in the first place? Second, what is the magnitude of the risk to patients if exposure to the harmful agent continues? Third, what are the adverse consequences of reducing or eliminating exposure to the harmful agent?

Clinical decision making is simple when both the likelihood of harm and its magnitude are great. Because the evidence of increased mortality from encainide and flecainide came from an RCT, we can be confident of the causal connection. Since treating only 21 people will result in an excess death, it is no wonder that clinicians quickly curtailed their use of these antiarrhythmic agents when the study results became available.

The clinical decision is also made easier when an acceptable alternative for avoiding the risk is available. For example, β-blockers prescribed for the treatment of hypertension can result in a symptomatic increase in airways resistance in patients with asthma or chronic airflow limitation, mandating the use of an alternative drug, such as a thiazide diuretic, in susceptible patients.²¹ Even if the evidence is relatively weak, the availability of an alternative can result in a clear decision. The early case-control studies demonstrating the association between aspirin use and Reye's syndrome were relatively weak and left considerable doubt about the causal relationship. Although the strength of inference was not great, the availability of a safe, inexpensive, and well-tolerated alternative, acetaminophen, justified use of this alternative agent in children at risk of Reye's syndrome.²²

In contrast to the early studies regarding aspirin and Reye's syndrome, multiple well-designed cohort and case-control studies have consistently demonstrated an association between NSAIDs and upper gastrointestinal bleeding, and our inference about harm has therefore been relatively strong. However, the risk of an upper gastrointestinal hemorrhage is quite low, and we don't have safer, equally efficacious anti-inflammatory alternatives available. We are therefore probably right in continuing to prescribe NSAIDs for the appropriate clinical conditions.

RESOLUTION OF THE SCENARIO

The study you retrieved on the risks of inhaled β-adrenergic therapy used a case-control design relying on computer record linkages between health insurance data and a drug plan.¹ The database for the study included 95% of the population of the province of Saskatch-

ewan in western Canada. The investigators matched 129 cases of fatal or near-fatal asthma with 655 controls who were also asthmatics. The investigators attempted to control for potential confounders, such as disease severity. Their measures of disease severity included the number of hospitalizations in the previous 24 months and an index of the aggregate use of medications. They found an association between the routine use of large doses of β -adrenergic agonist metered-dose inhalers and death from asthma (OR, 2.6 per canister per month; 95% CI, 1.7 to 3.9).

The study satisfied the validity criteria in Table 1 quite well. The investigators chose an appropriate control population and corrected for measurable potential differences in important prognostic factors in the treatment and control groups; exposure and outcome were measured the same way in treatment and control groups; the temporal

relationship is correct; and they found a dose-response gradient. However, the study used a case-control design rather than an RCT, and we remain uncertain whether differences in unmeasured prognostic variables between the treatment and control groups explain the results. In other words, it is still possible that the patients who used more β -agonists were sicker, and this (rather than their increased use of the drug) explains the increased risk of death.

The magnitude of the association is moderate, and although the baseline risk of death from asthma (44 deaths in 12 301 asthmatic patients receiving medication, 0.3%) is low enough that we would have to treat large numbers of patients before the drugs were responsible for a death, reducing preventable deaths is extremely important. The fact that the data came from a population-based study suggests the results are widely generalizable.

Thus, as an individual study on the subject, you find the results of an "association" between inhaled β -adrenergic agonist use and death both believable and relevant to your practice. Because it is not an RCT, you are less certain about a true causal relationship underlying the observed association. Full assessment of the likelihood of a causal relationship would require a systematic review of all the evidence in the literature. You tell your inquiring patients that there is an increased risk of death in heavy users of inhaled β -adrenergic agonists, but that you cannot be certain whether it is because of the drug or possibly the consequence of having severe disease. Intermittent use of inhaled β -agonist therapy in patients with reversible airflow obstruction provides an attractive alternative to more intensive administration, and many clinicians have responded to the results of this and other studies by choosing this alternative approach.

References

1. Spitzer WO, Suissa S, Ernst P, et al. The use of beta-agonists and the risk of death and near death from asthma. *N Engl J Med.* 1992;326:501-506.
2. Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, II: how to use an article about therapy or prevention, A: are the results of the study valid? *JAMA.* 1993;270:2598-2601.
3. Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, II: how to use an article about therapy or prevention, B: what were the results and will they help me in caring for my patients? *JAMA.* 1994;271:59-63.
4. Echt DS, Liebson PR, Mitchell LB, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo: the Cardiac Arrhythmia Suppression Trial. *N Engl J Med.* 1991;324:781-788.
5. The Cardiac Arrhythmia Suppression Trial II Investigators. Effect of the antiarrhythmic agent moricizine on survival after myocardial infarction. *N Engl J Med.* 1992;327:227-233.
6. Kristenson P, Irgens LM, Daltveit AK, Andersen A. Perinatal outcomes among children of men exposed to lead and solvents in the printing industry. *Am J Epidemiol.* 1993;137:134-143.
7. Carson JL, Strom BL, Soper KA, West SL, Morse ML. The association of nonsteroidal anti-inflammatory drugs with upper gastrointestinal tract bleed-

- ing. *Arch Intern Med.* 1987;147:85-88.
8. Walter SD. Determination of significant relative risks and optimal sampling procedures in prospective and retrospective comparative studies of various sizes. *Am J Epidemiol.* 1977;105:387-397.
9. Herbst AL, Ulfelder H, Poskanzer DC. Adenocarcinoma of the vagina: association of maternal stilbestrol therapy with tumor appearance in young women. *N Engl J Med.* 1971;284:878-881.
10. Soverchia G, Perr PF. Two cases of malformation of a limb in infants of mothers treated with an antiemetic in a very early phase of pregnancy. *Pediatr Med Chir.* 1981;3:97-99.
11. Holmes LB. Teratogen update: bendectin. *Teratology.* 1983;27:277-281.
12. Ray WA, Griffin MR, Schaffner W, Baugh DK, Melton LJ III. Psychotropic drug use and the risk of hip fracture. *N Engl J Med.* 1987;316:363-369.
13. Kellermann AL, Rivara FP, Rushforth NB, et al. Gun ownership as a risk factor for homicide in the home. *N Engl J Med.* 1993;329:1084-1091.
14. Hiatt RA, Fireman B. The possible effect of increased surveillance on the incidence of malignant melanoma. *Prev Med.* 1986;15:652-660.
15. Dement JM, Harris RL, Symons MJ, Shy CM. Exposures and mortality among chrysotile asbestos workers: part II, mortality. *Am J Med.* 1983;1:421-433.
16. Teicher MH, Glod C, Cole JO. Emergence of

- intense suicidal preoccupation during fluoxetine treatment. *Am J Psychiatry.* 1990;147:207-210.
17. Beasley CM, Dornseif BE, Bosomworth JC, et al. Fluoxetine and suicide: a meta-analysis of controlled trials of treatment for depression. *BMJ.* 1991;303:685-692.
18. Doll R, Hill AB. Mortality in relation to smoking: ten years' observation of British doctors. *BMJ.* 1964;II:1399-1410, 1460-1467.
19. Browner WS, Li J, Mangano DT, for the Study of Perioperative Ischemia Research Group. In-hospital and long-term mortality in male veterans following noncardiac surgery. *JAMA.* 1992;268:228-232.
20. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical Epidemiology: A Basic Science for Clinical Medicine.* 2nd ed. Boston, Mass: Little Brown & Co Inc; 1991.
21. Ogilvie RI, Burgess ED, Cusson JR, Feldman RD, Leiter LA, Myers MG. Report of the Canadian Hypertension Society Consensus Conference, 3: pharmacologic treatment of essential hypertension. *Can Med Assoc J.* 1993;149:575-584.
22. Soumerai SB, Ross-Degnan D, Kahn JS. Effects of professional and media warnings about the association between aspirin use in children and Reye's syndrome. *Milbank Q.* 1992;70:155-167.