Special Communication

The Case-Control Study

A Practical Review for the Clinician

Gregory F. Hayden, MD; Michael S. Kramer, MD; Ralph I. Horwitz, MD

• The retrospective case-control study is an important research strategy commonly encountered in the medical literature. A thoughtfully designed, carefully executed case-control study can be an invaluable source of clinical information, and physicians must often base important decisions about patient counseling and management on their interpretation of such studies. Unfortunately, the retrospective direction of case-control studies—looking "backwards" from an outcome event to an antecedent exposure—is accompanied by numerous methodological hazards. Careful attention must be paid to selection of appropriate study groups; definition and detection of the outcome event; definition and ascertainment of the exposure; assurance that the compared groups were equally susceptible to the outcome event at baseline; and careful statistical analysis. If systematic bias enters the research at any of these points, erroneous conclusions can result. Greater familiarity with the case-control method should enable clinicians to be more critically insightful when interpreting the results of published studies using this design format.

(*JAMA* 1982;247:326-331)

THE RANDOMIZED controlled trial (RCT) has become the methodological standard of excellence in studies of clinical therapeutics because it is designed to eliminate bias in the selection of patients, the administration of therapy, and the detection of outcomes. In many research settings, however, logistic, ethical, or economic considerations favor the use of observational or nonexperimental types of study design. One alternative methodological approach that has been widely used in studies of disease etiology is the retrospective case-control study. Although case-control studies have provided important data in a number of clinical and epidemiologic areas, the case-control design is subject to several potential sources of bias that can distort or invalidate the results. In numerous instances, multiple case-control studies performed to answer the same research question have led to conflicting and contradictory conclusions.[1]

Clinicians often encounter case-control research in their reading and must judge whether the study findings are valid and applicable to their clinical settings. They must then make important decisions about patient counseling and management on the basis of their interpretation of such studies. Some physicians argue that the task of deciding whether a study is well designed and therefore "believable" can be relegated to journal editors and that studies that have passed the demanding process of peer review can be trusted to be methodologically sound. The documented disparities among published case-control studies,[1] however, suggest otherwise, and should provide clinicians with the incentive to become familiar with the relevant methodological issues. The purpose of this review is to outline the strategies used in case-control research and to demonstrate with examples how and where bias is most likely to occur. Greater familiarity with case-control methodology should enable medical researchers to design more effective studies, and should help clinicians to be more critically insightful when interpreting the results of published studies using this design format.

General Structure of the Research

In the prototypic RCT, a group (or "cohort") of patients is exposed to an intervention (eg, a pharmaceutical agent) and, later, a particular outcome event (eg, survival) is assessed. A comparison group of unexposed patients is likewise followed prospectively to determine the "control" incidence of the outcome event. To remove patient or physician preference in the choice of therapy, a randomization (or chance) mechanism is used to assign patients to active or control treatments. The structure for this type of study may be diagrammed as in Fig 1. This basic format is shared by cohort surveys of an observational, nonexperimental nature in which exposure is not assigned randomly, but in which groups of exposed and nonexposed patients are nevertheless followed forward in time to observe a particular outcome.

With the case-control design, a different logic is employed. The researcher begins by selecting a group of patients ("cases") with a particular

From the Department of Pediatrics, University of Virginia Medical Center, Charlottesville (Dr Hayden), the Department of Pediatrics and Epidemiology, McGill University, Montreal (Dr Kramer), and the Department of Medicine, Yale University, New Haven, Conn (Dr Horwitz).

Reprint requests to the Department of Pediatrics, University of Virginia Medical Center, Charlottesville, VA 22908 (Dr Hayden).

Downloaded from www.jama.com at University of Arizona Health Sciences Library on July 11, 2009

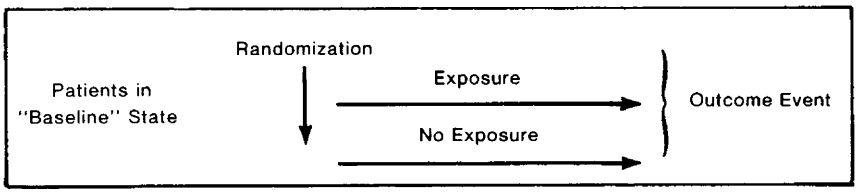

Fig 1.—Structure used in randomized control trial.

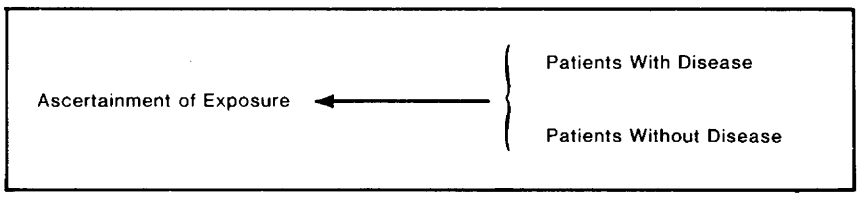

Fig 2.—Structure used in case-control study.

disease. (Other types of outcome event can serve as the basis for group selection, but the narrower term "disease" will be used in this discussion, since this is the most frequent situation.) Next, the researcher ascertains the rate of prior exposure to the alleged etiologic agent in the case group. The prior exposure rate is also determined in a group of patients without the disease ("controls"). The structure for this type of study may be diagrammed as in Fig 2.

Since this logic is directed "backwards" from effect to cause, a direction opposite to that used in the RCT or cohort design, this type of study has sometimes been referred to as a "trohoc" ("cohort" spelled backwards).[2] The major advantages of this format include savings in cost and time. Since the assembly and long-term follow-up of a large cohort is expensive and logistically difficult, the case-control design is often employed for investigation of diseases that occur infrequently or that develop years after exposure. The case-control format is also useful in the study of exposures (eg, water hardness or alcohol consumption during pregnancy) that cannot be randomized for logistic or ethical reasons.

Although the case-control format is therefore desirable, or even invaluable, in certain research settings, it must be employed with caution to avoid potential sources of error and bias. The methodological problems encountered in case-control studies will be considered in the same "reverse" order as they are faced by the clinical researcher: selection of appropriate case and control groups; definition and detection of the disease under study; definition and ascertainment of the exposure; baseline susceptibility of the compared groups to developing the disease; and statistical considerations.

Selection of Appropriate Case and Control Groups

One early decision faced by the researcher is how best to select the case and control groups. An important source of bias may occur if certain kinds of patients are arbitrarily excluded from either the case or control groups. If the excluded patients have particularly high or low rates of exposure, the exposure rate among the patients who remain as the cases or controls will be falsely lowered or raised, respectively. For example, in a recent study of the possible role of chronic low-level lead exposure in causing mental retardation, blood lead levels among mentally retarded cases were compared with those among hospitalized controls.[3] Children with diseases known to be associated with lead were excluded from the control group, but not from the case group. This exclusion would tend to decrease the average blood lead level among the remaining controls. Blood lead levels were statistically significantly higher among the case group, but this difference may perhaps be attributed to the constraint in selecting controls. A similar problem may have occurred in the earliest studies evaluating a possible association between reserpine and breast cancer.[4,5] Patients with thyrotoxicosis, renal disease, or cardiovascular disease were excluded from the control, non-breast-cancer groups in these studies. Since such patients are particularly likely to be receiving reserpine, their exclusion would probably decrease the exposure rate among the remaining controls and would falsely elevate the apparent carcinogenic risk of reserpine exposure. As Brown[6] has stated:

> Is it surprising that the breast cancer group had a higher proportion of patients who had received reserpine than did the control group, which had been selected to exclude conditions for which reserpine might have been used? This is like the fable from White showing that selecting a perfectly "random sample" of children to estimate the distribution of family size would reveal that there were no childless families!

The opposite problem may have occurred when children with spina bifida served as controls in a study attempting to link prenatal exposure to female sex hormones with congenital limb-reduction defects.[7] Since neural tube defects have themselves been associated with gestational exposure to female sex hormones,[8] the rate of hormonal exposure in the control group may have been falsely elevated, and the apparent risk of limb reduction defects falsely reduced.

Case-control studies using convenient patient samples from hospital populations are subject to an additional potential source of error referred to as "Berkson's bias." The relationship between the disease and exposure under consideration may be distorted in hospital-based studies if persons who are both exposed and diseased are more likely to be admitted to the hospital than are other groups. Brown has provided a detailed example of how Berkson's bias could account for misleading results in a hypothetical study investigating a possible association between low birth weight and cerebral palsy,[6] and Roberts et al[9] have recently demonstrated the bias empirically for the first time. Unfortunately, this potential source of error can be removed only if the case and control groups are selected from the same nonhospital community roster—a tactic that is rarely possible.

Another related problem arises because diseases that either are transient or lead to early death may be underrepresented in the case group. If

Downloaded from www.jama.com at University of Arizona Health Sciences Library on July 11, 2009

the case group is not representative of the intended population (ie, all persons with the particular disease), this may produce a substantial distortion in the study results, referred to as "Neyman's bias." This potential source of bias is least likely to occur in studies of chronic, nonfatal conditions.

Since some types of bias may arise when using only one particular control group for purposes of comparison, many authorities recommend the use of multiple and different control groups.[10-12] Consistency in the study findings despite using more than one control group tends to strengthen confidence in the validity of the results, whereas disparate findings suggest biased selection of one of the control groups. In a recent study of risk factors in subacute sclerosing panencephalitis (SSPE), for example, children with SSPE were matched for age, race, and gender with each of two controls—one who was a playmate of the patient, and another who was admitted to the same hospital in which the patient was confined at the time the diagnosis of SSPE was made.[13] Comparison of the case group with each of the two control groups revealed a statistically significant positive association between measles illness and SSPE, and a negative association between measles vaccination and SSPE. Another recent study of a possible association between aflatoxin and Reye's syndrome employed a similar choice of controls—one drawn from the patient's neighborhood and the other from current inpatients at the hospital to which the patient was admitted.[14] With respect to the presence of aflatoxin in serum or urine or both, there was no significant difference between the Reye's syndrome patients and either group of controls. Again, the consistency of the results using both control groups strengthens the likelihood of a correct conclusion. Inconsistency of results between separate control groups suggests the possibility of biased group selection. The explanation of such a discrepancy can be enlightening, however, as in a recent study of the proposed association between tonsillectomy and Hodgkin's disease.[15] Patients with Hodgkin's disease served as the cases; spouses and siblings of the cases served as the controls. The risk estimate of developing Hodgkin's disease among tonsillectomized vs nontonsillectomized persons was 3.1 using spouses as the controls and only 1.4 using siblings as the controls. This result suggests that a childhood event (such as a toxic exposure) that was common to both the cases and their siblings, and that was therefore "overcontrolled-for" by the use of the sibling controls, may be a cause of Hodgkin's disease.[10]

Regardless of how the investigator chooses the case and control groups, the method of selection should be established before the research data are obtained and analyzed. It may be possible to influence the study outcome after the fact by altering the composition of the case or control groups. In another study of the proposed association between reserpine and breast cancer, for example, the investigators compared rates of previous exposure to *Rauwolfia* derivatives among patients with breast cancer vs patients with all other neoplasms.[16] This analysis revealed no significant difference. The investigators then discarded the original control group and created a new one consisting of patients with selected neoplasms only. Reanalysis of the data using this limited control group revealed a statistically significant difference in the rates of reserpine use, suggesting an oncogenic risk of reserpine exposure. Such post hoc manipulation of the study groups to achieve a desired result is not scientifically permissible. Multiple reanalyses of study results based either on subgroups of cases and controls or on different definitions of exposure (a process that, when carried to its extreme, has been referred to as "dredging" the data), may be useful, however, in suggesting new hypotheses that can be evaluated in subsequent investigations.

Definition and Detection of the Disease Under Study

Once the investigator has decided in a general way how to select the case and control groups, the disease under study should be defined in specific, unambiguous terms by operational diagnostic criteria. The investigator must then check that similar diagnostic procedures and criteria have been used among cases and controls. Cases and controls should have had equal likelihoods of being checked for the occurrence of the disease, and the diagnostic procedures should have been performed and interpreted equally in both groups. This requirement is necessary to ensure that members of the control group are free of the disease, and is especially important when the disease may occur in an asymptomatic form (eg, gallstones or a "silent" undiagnosed cancer). Unequal diagnostic surveillance may have distorted the results, for example, in a recent study of risk factors associated with thrombocytopenia in the high-risk infant.[17] Thrombocytopenia was detected on the basis of either a routine weekly platelet count or a blood smear done for some medical reason. Hyperbilirubinemia and phototherapy were noted significantly more often in the thrombocytopenic babies than the controls, suggesting to the investigators a causative role of the hyperbilirubinemia. It seems plausible, however, that the association is artifactual and results from unequal diagnostic testing. Babies with hyperbilirubinemia are likely to have had a blood smear performed as part of their medical evaluation, and therefore had at least one "extra" chance for thrombocytopenia to be detected. Babies with hyperbilirubinemia may therefore have been overrepresented in the case group.

A similar problem due to unequal diagnostic surveillance may have resulted in an overestimation of risk in the proposed relationship between estrogens and endometrial cancer.[18] When the controls consisted of women thought not to have endometrial cancer on historical or clinical grounds, there appeared to be a significant association between hormonal use and endometrial cancer. When controls were selected from among women who had undergone diagnostic examination equal to that among the cases, however, there was a substantial decrease in the magnitude of risk for hormonal exposure. Additional, more recent studies[19-21] have discussed these results and disputed their interpretation, but the issue remains unsettled.

Another type of bias can occur if knowledge of the exposure has been considered in arriving at the diagno-

Downloaded from www.jama.com at University of Arizona Health Sciences Library on July 11, 2009

sis of the disease. This problem could arise in studies of estrogens and endometrial cancer, for example, if pathologists were more likely to diagnose a malignant neoplasm in patients with known hormonal exposure. This potential source of error has been termed "diagnostic-review bias,"[22] and can be avoided by blind interpretation of the data used to establish the diagnosis.

Definition and Ascertainment of Exposure

Once the study groups have been specifically identified, the investigator must retrospectively determine whether the cases and controls were exposed to the agent or disease under study. First, "exposure" should be defined in a precise, unambiguous fashion so that others can attempt to reproduce the findings. For example, if exposure is to a pharmaceutical agent, dosage and duration of therapy should be specified. The investigator must then collect data about exposure in an unbiased fashion. To preclude the soliciting and recording of information in a manner favorable to the hypothesis of the investigator, the data collector should be unaware of the hypothesis being tested or the identity of the person as a case or control. In one study of a possible association between the use of oral contraceptives during pregnancy and Down's syndrome, for example, two data collectors were used for each interview.[23] The first set up the interview, which was then conducted in "blind" fashion by the other data collector in a separate location.

Even when using unbiased blinded techniques to collect data, there may be inaccuracies in historical interview data. For example, in the previously cited study of risk factors for SSPE, 14% of parents gave a history of measles vaccination that conflicted with data discovered in an audit of the children's medical records.[13] In another study, women were interviewed during the fifth month of pregnancy and again immediately after delivery about their drug consumption during the first trimester.[24] Maternal responses in the first interview correlated well with medical and prescription records and were judged reasonably accurate. Responses in the later interview, however, were often in error.

Since such historical data are often variable and unreliable, an important type of bias can arise if cases or controls have greater ability or incentive to recall exposures. For example, in studies of various risk factors and birth defects, mothers of defective children are likely to have submitted their memory of pregnancy to rigorous and repeated reexamination for a possible cause of the malformation. Such mothers may therefore have better recall of events during pregnancy than mothers of normal children. This anamnestic inequality could in turn lead to an erroneous conclusion that a particular exposure was more common among mothers of malformed babies. The investigator should attempt to equalize recall by using multiple ways of asking the same question about exposure, by checking the answers with an alternative source of data, or by using other special strategies such as choosing the control group from among persons with another disease that would similarly enhance recall of exposure. For example, in one study of prenatal and postnatal complications and childhood autism, data were collected from both family interviews and medical records.[25] The investigators compared for the cases and the controls the proportion of family-reported complications that could be confirmed in the medical records. In this particular instance, the family-provided data agreed well with the medical records, and there was no evidence that the families had preferentially recalled medical events for either cases or controls.

Issues of timing can also be extremely important concerning the ascertainment of exposure, and can either augment or detract from the strength of a study's conclusions. In one case-control investigation of a nursery outbreak of pneumoperitoneum, for example, exposure to a particular nurse's aide who took overly deep rectal temperatures was significantly associated with illness.[26] This finding was strengthened by the observation of a relationship between the time infants were first exposed to this nurse's aide and the time they became ill (ie, the later infants were exposed, the later they became ill). In other instances, this type of temporal analysis can raise serious doubts about a study's conclusion. In one case-control analysis of risk factors for spina bifida, for example, a significant association was noted between this defect and gestational exposure to female sex hormones. Further analysis, however, revealed that the average interval between conception and exposure was six weeks. The neural tube is formed and closed by 4 weeks' gestational age. If the primary mechanism of spina bifida is failure of fusion, rather than rupture of the tube after closure, the teratogenic exposure must presumably occur before this time.[27] The timing of exposures in the cited study therefore makes the observed association less biologically plausible.

Baseline Susceptibility of Cases and Controls

An important source of bias can occur if, in their preexposure "baseline" state, the members of the case and control groups were not equally likely to have development of the disease under study. With the RCT design, randomization of exposure helps to ensure that the exposed and unexposed patients are equally susceptible to the outcome event. In case-control studies, however, exposure is not assigned randomly and may reflect important differences in demographic or clinical variables that might affect the development of the disease. Variables that are linked to exposure and also predispose to the disease are often referred to in epidemiologic parlance as "confounding factors." The potential impact of such factors must be assessed and, when necessary, controlled for. In a recent case-control study investigating the proposed protective effect of breast feeding on the development of atopic eczema, for example, the "raw" data analysis revealed a statistically significant increase in the risk of childhood eczema for breast-fed infants.[28] This risk was entirely eliminated, however, when the confounding factors of age, race, and ethnicity were controlled by matching. Although such demographic features are often examined and taken into account, clinical variables are frequently ignored. Equal clinical susceptibility is of great importance, however, as illustrated by an investigation of the

Downloaded from www.jama.com at University of Arizona Health Sciences Library on July 11, 2009

proposed relationship between birth defects and gestational exposure to various medications.[29] In this study, a substantial difference was noted between the case and control groups in the rate of maternal exposure to female sex hormones. On further analysis, however, many more exposed women had positive family histories of certain birth defects. Such women are at increased risk of giving birth to children with malformations, and when they were removed from the analysis, the observed difference in exposure rates was substantially decreased.

The investigator can use several different strategies to ensure that, at baseline, the members of the case and control groups are similar in demographic and clinical features. This goal can be accomplished by matching the group members on the pertinent variables or, at a minimum, by showing group evidence of comparability with respect to these variables. An alternative strategy is to employ stratification or multivariate methods in the data analysis to control for whatever differences may exist. So, for example, in a recent study intended to determine the source of lead for inner-city children with moderately elevated blood lead levels, cases and controls were matched for age, area of residence, and social class.[30] Despite this matching, the groups were observed to differ in racial and economic composition. These latter two factors were therefore controlled for in the statistical analysis.

The reasons for exposure are often relevant, since they may reflect differing baseline susceptibility of the cases and controls to the disease under study. One source of such bias can occur if any early manifestation of the disease serves as a stimulus for exposure to the alleged etiologic agent. Suppose, for example, that an abnormal embryo can cause maternal genital bleeding and that such bleeding can, in turn, result in the prescription of "supportive" hormone therapy. If a proportion of such pregnancies continue to term, an association may be observed between hormonal exposure and birth defects. To conclude that the hormones caused the malformations would be erroneous, however, because of the differences in susceptibility to malformations before the hormones were administered. This type of bias, in which an early manifestation of the disease serves as a stimulus for exposure, has been called "protopathic bias."[31] To control for this potential error, the investigator must determine the reason(s) for exposure and make adjustments accordingly.

Statistical Considerations

For both the RCT and case-control designs, it is often possible to display the results in a conventional 2×2 table.

		Disease		
		Yes	No	**Total**
Exposure	Yes	a	b	**a+b**
	No	c	d	**c+d**
	Total	**a+c**	**b+d**	N

In the RCT, the risk of developing the outcome event in the exposed group is directly compared to the risk in the nonexposed group, and a "risk ratio" calculated as

$$(a/[a+b])/(c/[c+d])$$

In a case-control study, a true risk or risk ratio cannot be calculated since the researcher does not begin with groups of exposed and nonexposed patients. The sums (a+b) and (c+d) are contrived values formed by the arbitrary selection of the case and control groups. The only meaningful sums in the case-control arrangement of the data are (a+c), the members in the selected disease group, and (b+d), the members in the selected control group. As a substitute for a risk ratio, one may compare the likelihood (or odds) of exposure vs nonexposure among the case group to the likelihood (or odds) of exposure vs nonexposure among the control group.[18] This "odds ratio" can be calculated as

$$\left(\frac{a}{a+c}\Big/\frac{c}{a+c}\right)\Big/\left(\frac{b}{b+d}\Big/\frac{d}{b+d}\right)$$

This expression may be simplified as (a/c)/(b/d), or ad/bc. As demonstrated algebraically by Cornfield,[32] this odds ratio may serve as a good approximation of the risk ratio as long as the disease under study is relatively uncommon. Statistical significance can be determined either by the use of the χ^2 test, or by placing a 95% confidence interval around the odds ratio.[33] If this 95% confidence interval includes the value of 1.0, the "null hypothesis" of no difference between the groups cannot be rejected, and the risk is not statistically greater or less than 1.0. So, for example, in the aforementioned study of childhood autism, six (5%) of 118 cases had been slow to cry at delivery, compared with only four (2%) of 246 controls.[25] The odds ratio in this instance is 3.2, suggesting a moderately large increase in risk. Because of the small numbers of affected patients, however, the 95% confidence interval around the odds ratio is large, ranging from 0.95 to 11.1. Since this interval includes the value of 1.0, the risk estimate of 3.2, despite its magnitude, is not statistically significantly greater than 1. This conclusion is confirmed by χ^2 analysis (χ^2=3.57; .10>P>.05).

Just as with the RCT design, one cannot always equate statistical significance (or lack thereof) with clinical significance. In a large case-control study, for example, a clinically insignificant difference in rates of exposure may attain statistical significance. In contrast, a clinically significant difference in exposure rates may not attain statistical significance in a small study owing to the limited sample size or low rate of exposure. The finding of no statistically significant differences does not prove that the exposure rates are in fact the same. One must calculate type II (or β) error in these situations to determine whether, in a study of that particular size, one could reasonably expect to detect whatever difference in exposure rates is considered clinically meaningful.[34] To continue with the previous example, even though the odds ratio of 3.2 is not statistically significantly greater than 1, this does not necessarily mean that there is no risk involved. If one assumed that the actual prevalence of being slow to cry is 0.02 among controls, and that this prevalence is tripled to 0.06 among cases, then there is a 30% chance (P_β=.3) of failing to show a significant difference in this study using 118 cases and 246 controls. In this instance, it might therefore be unwise to accept the null hypothesis of no intergroup difference. In fact, if the sample size had been only 10% larger, and if the rates of being slow to cry among the cases and controls had remained the same,

Downloaded from www.jama.com at University of Arizona Health Sciences Library on July 11, 2009

the difference would have been significant at the 5% level.

Conclusion

The retrospective case-control study is an important research strategy that is commonly encountered in the medical literature. A thoughtfully designed, carefully executed case-control study can provide important clinical information. Unfortunately, the backwards logic of case-control studies is accompanied by several methodological hazards, and numerous case-control studies have arrived at conflicting or incorrect conclusions.[1] Many of these errors can be directly attributed to methodological deficiencies in the basic study design. When reading a report of a case-control investigation, it therefore behooves the clinician to ask two fundamental questions: (1) In this particular research situation, what are the most likely and potentially most damaging sources of bias? and (2) What precautions, if any, has the investigator taken to minimize and control for these possible errors? In making this methodological assessment, the reader must determine whether adequate attention was paid to the selection of appropriate study groups; the definition and detection of the outcome event; the definition and ascertainment of exposure; assurance that the compared groups were equally susceptible to the outcome event at baseline; and careful statistical analysis. Systematic bias at any of these points may result in erroneous conclusions and misleading clinical inferences. Greater familiarity with the case-control method should enable the clinician to be more critically insightful when reading published studies and to make better-educated judgments about the validity of data obtained with this design format.

Alvan R. Feinstein, MD, stimulated our interest in clinical epidemiology and made many helpful suggestions. Judy Mauck and Marjorie Hampton prepared the manuscript.

References

1. Horwitz RI, Feinstein AR: Methodologic standards and contradictory results in case-control research. *Am J Med* 1979;66:556-564.

2. Feinstein AR: Clinical biostatistics: 20. The epidemiologic trohoc, the ablative risk ratio and 'retrospective' research. *Clin Pharmacol Ther* 1973;14:291-307.

3. Beattie AD, Moore MR, Goldberg A, et al: Role of chronic low-level lead exposure in the etiology of mental retardation. *Lancet* 1975; 1:589-592.

4. Reserpine use in relation to breast cancer, Boston Collaborative Drug Surveillance Program. *Lancet* 1974;2:669-671.

5. Heinonen OP, Shapiro S, Tuminen L, et al: Reserpine use in relation to breast cancer. *Lancet* 1974;2:675-677.

6. Brown GW: Berkson fallacy revisited: Spurious conclusions from patient surveys. *Am J Dis Child* 1976;130:56-60.

7. Hellstrom B, Lindsten J, Nilsson K: Prenatal sex-hormone exposure and congenital limb-reduction defects. *Lancet* 1976;2:372-373.

8. Gal I: Risks and benefits of the use of hormonal pregnancy test tablets. *Nature* 1972; 240:241-242.

9. Roberts RS, Spitzer WO, Sackett DL, et al: An empirical demonstration of Berkson's bias. *J Chronic Dis* 1978;31:119-128.

10. Cole P: The evolving case-control study. *J Chronic Dis* 1979;32:15-27.

11. Sartwell PE: Retrospective studies: A review for the clinician. *Ann Intern Med* 1974; 81:381-386.

12. Ibrahim MA, Spitzer WD: The case-control study: The problem and the prospect. *J Chronic Dis* 1979;32:139-144.

13. Halsey NA, Modlin JF, Jabbour JT, et al: Risk factors in subacute sclerosing panencephalitis: A case-control study. *Am J Epidemiol* 1980;111:415-424.

14. Nelson DB, Kimbrough R, Landrigan PS, et al: Aflatoxin and Reye's syndrome: A case-control study. *Pediatrics* 1980;66:865-869.

15. Gutensohn N, Li F, Johnson R, et al: Hodgkin's disease, tonsillectomy and family size. *N Engl J Med* 1975;292:22-25.

16. Armstrong B, Stevens N, Doll R: Retrospective study of the association between use of *Rauwolfia* derivatives and breast cancer in English women. *Lancet* 1974;2:672-675.

17. Mehta P, Vasa R, Neumann L, et al: Thrombocytopenia in the high-risk infant. *J Pediatr* 1980;97:791-794.

18. Horwitz RI, Feinstein AR: Alternative analytic methods for case-control studies of estrogens and endometrial cancer. *N Engl J Med* 1978;299:1089-1094.

19. Antunes CM, Stolley PD, Rosenshein NB, et al: Endometrial cancer and estrogen use: Report of a large case-control study. *N Engl J Med* 1979;300:9-13.

20. Hulka BS, Grimson RC, Greenberg BG, et al: 'Alternative' controls in a case control study of endometrial cancer and exogenous estrogen. *Am J Epidemiol* 1980;112:376-387.

21. Shapiro S, Kaufman DW, Slone D, et al: Recent and past use of conjugated estrogen in relation to adenocarcinoma of the endometrium. *N Engl J Med* 1980;303:485-489.

22. Ransohoff DF, Feinstein AR: Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med* 1978;299:926-930.

23. Janerich DR, Flink EM, Keogh MD: Down's syndrome and oral contraceptive usage. *Br J Obstet Gynaecol* 1976;83:617-620.

24. Klemetti A, Saxen L: Prospective versus retrospective approach in the search for environmental causes of malformation. *Am J Public Health* 1967;57:2971-2975.

25. Deykin EY, MacMahon B: Pregnancy, delivery, and neonatal complications among autistic children. *Am J Dis Child* 1980;134:860-864.

26. Horwitz M, Bennett JV: Nursery outbreak of peritonitis with pneumoperitoneum probably caused by thermometer-induced rectal perforation. *Am J Epidemiol* 1976;104:632-644.

27. Laurence KM: Reply to Gal. *Nature* 1972; 240:242.

28. Kramer MS, Moroz B: Do breast feeding and delayed solids protect against subsequent atopic eczema? *J Pediatr* 1981;98:546-550.

29. Greenburg G, Inman WH, Weatherall JA, et al: Maternal drug histories and congenital abnormalities. *Br Med J* 1977;2:853-856.

30. Charney E, Sayre J, Coulter M: Increased lead absorption in inner city children: Where does the lead come from? *Pediatrics* 1980;65:226-231.

31. Horwitz RI, Feinstein AR: The problem of 'protopathic bias' in case-control studies. *Am J Med* 1980;68:255-258.

32. Cornfield J: A method for estimating comparative rates for clinical data. *JNCI* 1950-1951;11:1269-1275.

33. Fleiss JL: Confidence intervals for the odds ratio in case-control studies: The state of the art. *J Chronic Dis* 1979;32:69-77.

34. Feinstein AR: Clinical biostatistics 34. The other side of 'statistical significance': Alpha, beta, delta, and the calculation of sample size. *Clin Pharmacol Ther* 1975;18:491-505.

Downloaded from www.jama.com at University of Arizona Health Sciences Library on July 11, 2009