Some Principles of Epidemiologic Studies

Richard F. Morton, MD, and J. Richard Hebel, PhD
Baltimore, Maryland

The use of epidemiological evidence is frequently involved in clinical decision making. Usually epidemiological investigation seeks to determine the extent of the association between a suspected factor and the occurrence of a disease. When the data are collected retrospectively then rates of exposure to the factor are compared for a group of cases of the disease and a group of controls. If the study is prospective then incidence rates of the disease are compared for a group exposed to the factor and an unexposed group. Although the retrospective approach is often more feasible it is also more vulnerable to bias. The possible influence of bias and chance must be taken into consideration when interpreting the results of any study. Relative risk serves as a useful index for expressing the strength of the association between the factor and the disease.

Epidemiological studies[1-7] are being used increasingly to determine association between suspected factors and medical outcomes. Federal regulation, particularly in the form of recommendations in the *Physicians' Desk Reference (PDR)* and package inserts, is often based on the results of these investigations. The clinician is required to make therapeutic decisions based upon conclusions derived from epidemiological studies, frequently of retrospective design. This article will examine some of the premises pertinent to the validity of epidemiological evidence.

Exposure and Disease

The majority of epidemiological studies can be reduced, in their simplest form, to the relationship between exposure and disease,[8] as shown in Table 1. Exposure here is used in a general sense, including the presence of an attribute, such as hypertension, or any suspected etiological factor. Disease is used here, but it can include any outcome, such as survival.

If the approach is retrospective, the investigator starts with a + c cases. A comparison group of b + d controls is then selected and the position shown in Table 2 is reached. The participants are then assigned retrospectively to the exposure rows, so that the 2×2 table is completed (Table 1). Analysis is then performed by comparing the exposure rates between the case and control groups.

$$\text{Exposure rate among cases} = \frac{a}{a+c}$$

$$\text{Exposure rate among controls} = \frac{b}{b+d}$$

Vessey and Doll,[1] for example, did a retrospective study to investigate the relationship between the use of oral contraceptives and thromboembolic disease. Their results were as shown in Table 3. In this retrospective study,

From the Department of Epidemiology and Preventive Medicine, University of Maryland School of Medicine, Baltimore, Maryland. Requests for reprints should be addressed to Dr. Richard F. Morton, Department of Epidemiology and Preventive Medicine, University of Maryland School of Medicine, 655 West Baltimore Street, Baltimore, MD 21201.

0094-3509/79/040781-05$01.25

Table 1. Schematic Representation of Completed Data, Gathered Prospectively or Retrospectively, Showing Relationship Between Exposure and Disease

		Disease	
		Present	Absent
Exposure	Present	a	b
	Absent	c	d

Table 2. Cases and Controls Assembled at the Start of a Retrospective Study, Not Yet Differentiated into Exposed and Nonexposed Groups

		Disease	
		Present	Absent
Exposure	Present	?	?
	Absent	?	?
	Total	a + c	b + d

Table 3. Results of a Retrospective Study,[1] where Use of Oral Contraceptives Is the Exposure, and Thromboembolism Is the Disease Outcome

			Thromboembolism	
			Cases	Controls
Oral Contraceptives	Used		26	10
	Not Used		32	106
		Total	58	116

$$\text{Exposure rate among cases} = \frac{26}{58} = .448 \text{ or } 44.8 \text{ percent}$$

$$\text{Exposure rate among controls} = \frac{10}{116} = .086 \text{ or } 8.6 \text{ percent}$$

When the design is prospective, however, the investigator starts with the row total a+b (the exposed group) and the row total c+d (the nonexposed group), and the position shown in Table 4 is reached. The participants are then followed forward, and they eventually fall into the diseased or nondiseased columns, so that the 2 × 2 table is completed prospectively (Table 1). Analysis is then performed by comparing the rate of disease occurrence (incidence) between the exposed and nonexposed groups.

$$\text{Incidence in the exposed group} = \frac{a}{a+b}$$

$$\text{Incidence in the nonexposed group} = \frac{c}{c+d}$$

For example, a prospective study by Green and Sarubbi[9] examined febrile morbidity after cesarean section and the type of anesthesia used. Their results were as shown in Table 5. In this prospective study,

Table 4. The Start of a Prospective Study, Showing the Exposed and Nonexposed Groups, Not Yet Differentiated into Diseased and Non-diseased

		Disease		
		Present	Absent	Totals
Exposure	Present	?	?	a + b
	Absent	?	?	c + d

Table 5. Results of a Prospective Study[9]

		Febrile Morbidity		
		Yes	No	Total
Anesthesia Type	Regional	9	38	47
	General	45	37	82

Incidence of febrile morbidity among those exposed to regional anesthesia $= \frac{9}{47} = .192$ or 19.2%

Incidence of febrile morbidity among those exposed to general anesthesia $= \frac{45}{82} = .549$ or 54.9%

The advantage of the retrospective design is that it may be used to study a rare disease[6,10]; for a series of cases of a rare disease may be retrospectively collected from a group of large hospitals, and compared with controls free of the disease. The prospective design, by contrast, entails assembling a study cohort free of the disease under consideration, and following them forward in time so as to observe the development of the disease among some of their members. With a rare disease, however, the number of cases would be very small. Further, the retrospective study yields a result in a relatively short time. This may be important, if the outcome is a suspected pathology resulting from drug exposure.[1-6,10,11]

When the phocomelia epidemic was detected in 1962,[11] the retrospective approach was used to identify the cause rapidly. Cases of phocomelia were compared with a control group of healthy babies, and the exposure of the two groups contrasted with respect to a variety of prenatal factors, including thalidomide. It was found that the affected babies had a far greater proportion of mothers who had taken thalidomide early in pregnancy than the normal babies.

The price to be paid for these two advantages, namely, the ability to study the etiology of a rare disease, and speed, is that the probability of bias is greater in a retrospective study than in the prospective approach.

Bias

Bias is systematic error, resulting in over or under estimation of the strength of the association. The validity of any study depends on the accuracy with which the subjects are assigned to the four categories, a, b, c, and d (Table 1). Misclassification may occur because of over or underdiagnosis. If a disease entity is well defined, such as cancer of the lung, the diagnosis being uniform and established and the majority of cases coming to medical attention, and there is little selectivity by physicians in hospitalizing the patient, misclassification is minimal. Thromboembolic phenomena present a contrast. The disease is difficult to diagnose; it may present as a complication of another medical or surgical condition, and criteria are not uniform.[12] Also the presence or absence of exposure may influence the management. If a physician is faced with a young woman, known to be on oral contraceptives, presenting with leg pain, she may

be hospitalized more readily than another similar patient not on oral contraceptives. The control group is represented by the participants in the column b + d. It is necessary to ensure the disease is absent in this group, which may be simple in the case of lung cancer, but less certainty prevails with thromboembolic disease.

The sample of cases studied may not represent the entire spectrum of the disease. Hospitalized cases may exclude mild cases and those who die prior to admission. Coronary heart disease is such an example.

The design in Table 1 also assumes that the exposure is present or absent, both categories, of course, being mutually exclusive. The exposure may be difficult to define and measure, eg, type A personality, or difficult to recall, as in drug use. Further, the exposure may be intermittently present, as in the case of oral contraceptive use. Thus, the row a + b may over or under represent the exposure in the sample. If the information sought is unchanging and usually available, such as blood group, bias may be minimal. In a more usual case the needed information is not available and is sought by interview or questionnaire. The recall of events in the distant past may be inaccurate, or the information supplied by the informant may be biased.

Selective recall may occur among the cases, as they necessarily know they have the disease, and may already associate it with exposure. The interviewer, being aware of the identity of cases and controls, may unconsciously probe more among the cases, seeking a positive association. To minimize bias from this source, the interviewer ideally should be unaware of which participant is a case or control, but this is difficult to achieve in retrospective studies.

Bias may occur in the selection of controls, particularly if hospitalized patients are used, as is frequently the case. It is essential that the control group be as like the diseased group as possible, yet also similar to the general population in distribution of the exposure, if the results are to be extended to the general population. The hospitalized controls may contain an unrepresentative proportion of a particular attribute, hypertensives and smokers for example. If controls with serious systemic diseases, such as cancer, are used,[3] presence of this disease may affect exposure to the suspected etiological factor. Patients with gynecologic cancers are less likely to be precribed estrogen than those free of such cancers.

In an attempt both to reduce and measure the possible bias introduced by using hospital controls, a second series of controls may be assembled from the general population.[13] If the results are similar for both control groups in terms of exposure frequency, then bias from use of hospital controls is minimal.

Matching

The comparison between cases and controls may reveal a difference in exposure rates, and the development of disease may be ascribed to this difference, provided the two groups are otherwise comparable. In order to attain such comparability they are frequently matched for characteristics known to influence the distribution of the disease. Age is a powerful determinant of disease, so that, in order to eliminate this effect from the comparison, the controls are matched to the cases for age. Socioeconomic status, because it influences environmental hazards and life-styles, is frequently a disease determinant, and therefore is used to match cases with controls. It should be emphasized that when a variable is used for matching, its etiologic role cannot be investigated because cases and controls are then automatically similar with respect to that characteristic.

A study by Douglas[14] showed that children weighing 2,500 g or less at birth had significantly lower school achievement than control children matched for sex and age. The implication of this result was that the child's low birth weight was the cause of the lower school achievement. When the children were further matched according to socioeconomic factors, however, the difference in school achievement largely disappeared.

Interpretation of Results

A difference, either in exposure rates, in the case of the retrospective study, or in incidence rates, for the prospective study, is evidence of an association between the exposure and the disease. Caution must be used here, however, because the difference may be due to: (1) a real effect of the exposure; (2) a bias in the study; (3) chance; or some combination of these. If (2) and (3) cannot both be ruled out as possible explanations, the difference permits no clear interpretation.

Statistical tests may be used to determine the

probability that chance could explain the difference observed. The P value given by a statistical test expresses the probability of a difference as large or larger than that observed occurring by chance alone. Hence, a small P value attached to the difference indicates that it is not likely to be the result of mere chance. On the other hand, if P>.05, then the role of chance as a possible determinant for the difference should not be discounted. For this reason such results are labeled as not statistically significant (NS).

A result which is not statistically significant, that is one with P>.05, does not exclude the existence of an association. Unless the association is strong the sample size required to demonstrate statistical significance may be huge for an epidemiologic study. Goldstein,[15] for example, shows that to have a 90 percent chance to attain a significant result, in a study of the influence of smoking in pregnancy on perinatal mortality, a sample of 23,000 is needed.

Finally, the study design must be scrutinized to assess the influence of possible biases. As previously noted, there is more opportunity for bias in a retrospective study than a prospective one. Thus, in a retrospective study, it is often difficult to eliminate bias as the explanation for a difference detected in exposure rates.

Relative Risk

The strength of the association between the factor and the disease is measured by the relative risk (RR) where

$$RR = \frac{\text{Incidence rate for those exposed to the factor.}}{\text{Incidence rate for those not exposed to the factor.}}$$

Although RR can be computed directly from a prospective study, it must be approximated when the study is retrospective. From a retrospective study, RR≈ad/bc where a, b, c, and d are taken from the 2 × 2 table in Table 1. The approximation will be close provided that the frequency of the disease is low and the exposure rates in the segments of the general population with and without the disease are similar to those found for the corresponding study groups.[16]

Relative risk enables the investigator to assess the importance of a particular factor as an etiologic agent. It is useful to the clinician in determining when a patient is at increased risk for a disease because of some particular exposure. However, the relative risk does not indicate the probability that someone with the factor will develop the disease. It has been shown[1,2] that oral contraceptive users have a 4.5 relative risk of developing venous thromboembolism compared with nonusers. But, in an oral contraceptive user under the age of 20, the probability of her developing thromboembolism is extremely small. Physicians must always apply the relative risk in conjunction with their knowledge of the natural history of the disease. It is especially important to bear this in mind when the relative risk has been determined from a retrospective study. This is because this design does not yield incidence rates for either the exposed or the nonexposed groups. Thus, the relative risk estimate for those exposed is merely a multiple of an unknown incidence rate among those not exposed.

References

1. Vessey MP, Doll R: Investigation of relation between use of oral contraceptives and thromboembolic disease. Br Med J 2:199, 1968
2. Sartwell PE, Masi AT, Arthes FG, et al: Thromboembolism and oral contraceptives: An epidemiologic case-control study. Am J Epidemiol 90:365, 1969
3. Smith DC, Prentice R, Thompson DJ, et al: Association of exogenous estrogen and endometrial carcinoma. N Engl J Med 293:1164, 1975
4. Ziel HK, Finkle WD: Increased risk of endometrial carcinoma among users of conjugated estrogens. N Engl J Med 293:1167, 1975
5. Mack TM, Pike MC, Henderson BE, et al: Estrogens and endometrial cancer in a retirement community. N Engl J Med 294:1262, 1976
6. 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 284:878, 1971
7. Bibbo M, Haenszel WM, Wied GL, et al: A twenty-five year follow-up study of women exposed to diethylstilbestrol during pregnancy. N Engl J Med 298:763, 1978
8. Sartwell PE: Epidemiology. In Sartwell PE (ed): Preventive Medicine and Public Health. New York, Appleton-Century-Crofts, 1959, chap 1
9. Green SI, Sarubbi FA: Risk factors associated with post cesarean section febrile morbidity. Obstet Gynecol 49:686, 1977
10. Edmondson HA, Henderson B, Benton B: Liver cell adenomas associated with use of oral contraceptives. N Engl J Med 294:470, 1976
11. Kohler HG, Fisher AM, Dunn PM: Thalidomide and congenital abnormalities. Lancet 1:326, 1962
12. Harris WH, Salzman EW, Athanasoulis CA, et al: Aspirin prophylaxis of venous thromboembolism after total hip replacement. N Engl J Med 297:1246, 1977
13. Collaborative Group for the Study of Stroke in Young Women: Oral contraception and increased risk of cerebral ischemia or thrombosis. N Engl J Med 288:871, 1973
14. Douglas JWB: 'Premature' children at primary schools. Br Med J 1:1008, 1960
15. Goldstein H: Smoking in pregnancy: Some notes on the statistical controversy. Br J Soc Prevent Med 31:13, 1977
16. Lilienfeld AM: Foundations of Epidemiology. New York, Oxford University Press, 1976, p 179