Monographs in Epidemiology and Biostatistics
edited by Abraham M. Lilienfeld

1 THE EPIDEMIOLOGY OF DEMENTIA
James A. Mortimer and Leonard M. Schuman
1981

2 CASE CONTROL STUDIES
Design, Conduct, Analysis
James J. Schlesselman
1982

3 EPIDEMIOLOGY OF MUSCULOSKELETAL DISORDERS
Jennifer L. Kelsey
1982

4 URBANIZATION AND CANCER MORTALITY
The United States Experience, 1950–1975
Michael R. Greenberg
1983

5 AN INTRODUCTION TO EPIDEMIOLOGIC METHODS
Harold A. Kahn
1983

6 THE LEUKEMIAS
Epidemiologic Aspects
Martha S. Linet
1984

7 SCREENING IN CHRONIC DISEASE
Alan S. Morrison
1985

8 CLINICAL TRIALS
Design, Conduct, and Analysis
Curtis L. Meinert
1986

9 VACCINATING AGAINST BRAIN DYSFUNCTION SYNDROMES
The Campaign Against Rubella and Measles
Ernest M. Gruenberg
1986

10 METHODS IN OBSERVATIONAL EPIDEMIOLOGY
Jennifer L. Kelsey, W. Douglas Thompson, Alfred S. Evans
1986

11 CLINICAL EPIDEMIOLOGY: THE STUDY OF THE OUTCOME OF ILLNESS
Noel S. Weiss
1986

Monographs in Epidemiology and Biostatistics
Volume 10

Methods in Observational Epidemiology

Jennifer L. Kelsey
COLUMBIA UNIVERSITY FACULTY OF MEDICINE
SCHOOL OF PUBLIC HEALTH, DIVISION OF EPIDEMIOLOGY

W. Douglas Thompson
YALE UNIVERSITY SCHOOL OF MEDICINE
DEPARTMENT OF EPIDEMIOLOGY AND PUBLIC HEALTH

Alfred S. Evans
YALE UNIVERSITY SCHOOL OF MEDICINE
DEPARTMENT OF EPIDEMIOLOGY AND PUBLIC HEALTH

New York Oxford
OXFORD UNIVERSITY PRESS
1986

Chapter 10, are used. Chapter 10 also gives procedures for estimating sample sizes needed to detect significant effects, using different types of study designs.

CONFOUNDING, INDIRECT AND DIRECT CAUSES OF DISEASE, AND EFFECT MODIFICATION

Three other major concerns of epidemiologists when designing, analyzing, and interpreting studies are controlling for *confounding variables*, identifying *indirect and direct causes* of disease, and detecting *effect modification*. *Confounding* can perhaps best be introduced by an example. Early epidemiologic studies suggested that women with breast cancer were less likely to have breast-fed their infants and had breast-fed their infants for shorter periods of time than control women of the same age, thus suggesting that lactation protects against breast cancer (17, 43). However, longer breast feeding is associated with having larger numbers of children, and it was subsequently found that breast feeding was not less common among cases than controls, once it was taken into account that the controls had had more children than the cases (20, 21). In this instance, parity was a confounder for the apparent negative association between breast feeding and breast cancer. Next, it was realized that women who have few children tend to have their first child at a relatively late age, and that in most geographic areas the number of children a woman has does not affect her risk for breast cancer, once account is taken of her age at first full-term pregnancy (22). In other words, unless age at first full-term pregnancy, the confounding variable, is taken into account either in the study design or in the statistical analysis, misleading conclusions may be reached about the relationship between lactation and breast cancer or between parity and breast cancer.

Now that an example of confounding has been presented, a definition will be given: A *confounder* is a variable that (a) is causally related to the disease under study (or, as often occurs in practice, serves as a proxy measure for unknown or unmeasured causes), and (b) is associated with the exposure under study in the study population, but is not a consequence of this exposure. It follows from (a) that within each level of the exposure under study, the confounder is related to risk for disease, or in probabilistic terms, the confounder is related to the disease conditional on exposure. Accordingly, age at first delivery is a confounder for the association between par-

ity and breast cancer because it is causally related to breast cancer (or a proxy for an as yet unknown cause), and is associated with parity, but is not caused by parity. At each level of parity, older age at first delivery is positively associated with risk for breast cancer. For instance, at parity 1 the relative risks for breast cancer according to age at first delivery (rates of breast cancer in those at a given age at first delivery divided by the rate in nulliparous women) have been estimated (22) to be

Age at first delivery				
<20	20–24	25–29	30–34	≥35
0.6	0.6	0.8	1.0	1.0

Similarly, increasing relative risks for breast cancer with older age at first delivery are seen for parity 2, parity 3, and higher levels of parity. On the other hand, for a given age at first delivery, no effect of number of births can be detected. If age at first birth is taken into account in the analysis (by methods described in Chapters 5–7), the relative risks associated with various levels of parity (rates of breast cancer among women of a given parity divided by rate of breast cancer among nulliparous women) are

Number of births					
0	1	2	3	4–8	≥9
1.0	1.2	1.0	1.0	0.9	0.9

In other words, the apparent relationship between low parity and risk for breast cancer occurred only because low parity is associated with late age at first delivery. Parity is not associated with risk for breast cancer, once age at first delivery is taken into account.

In many instances in which it is believed that a causal association exists, the possibility must be kept in mind that an unidentified confounder is in fact responsible for the association. For instance, many studies have demonstrated a positive association between the presence of herpes simplex virus type 2 (HSV 2) and cervical cancer. However, one of the main reasons this association has not been accepted as causal by many people is that another as yet unidentified sexually transmitted agent may exist that (a) *is* causally associated with cervical cancer and (b) is associated with HSV 2 exposure only because HSV 2 and the unidentified agent are associated with multiple sexual partners and with early age at first intercourse. Certain human papilloma viruses may be such agents.

In addition to identifying confounding variables, which, if

Table 1–2. Number and percentage of firstborn among infants with congenital dislocation of the hip (CDH) and among control infants, hypothetical data

	Cases (CDH)	Controls	Total
Number firstborn	100	80	180
Number not firstborn	100	120	220
Total	200	200	400
Percentage firstborn	50	40	

ignored, may cause the investigator to believe that an association exists when it does not or to believe that an association does not exist when in fact it does, an investigator also is concerned in the study design and analysis with *separating out indirect causes from more direct causes.* For instance, it is known that infants with congenital dislocation of the hip are more likely to have been firstborn than are other infants. This is shown using data from a hypothetical case-control study in Table 1–2. It is also recognized that infants with congenital dislocation of the hip are more likely to have been born by breech delivery than later-born infants (Table 1–3) and that first-born infants are more likely to have been born by breech delivery than other infants. Accordingly, the investigator wants to know (a) if infants with congenital dislocation of the hip are more likely to have been born by breech regardless of whether they are firstborn, and (b) if infants with congenital dislocation of the hip are more likely to be firstborn regardless of whether they are born by breech. Table 1–4 shows that infants with congenital dislocation of the hip are more likely to have been born by breech regardless of whether

Table 1–3. Number and percentage of births by breech delivery among infants with congenital dislocation of the hip (CDH) and among control infants, hypothetical data

	Cases (CDH)	Controls	Total
Number born by breech	60	20	80
Number not born by breech	140	180	320
Total	200	200	400
Percentage born by breech	30	10	

Table 1–4. Number and percentage of births by breech delivery among infants with congenital dislocation of the hip (CDH) and among control infants according to whether they were firstborn, hypothetical data

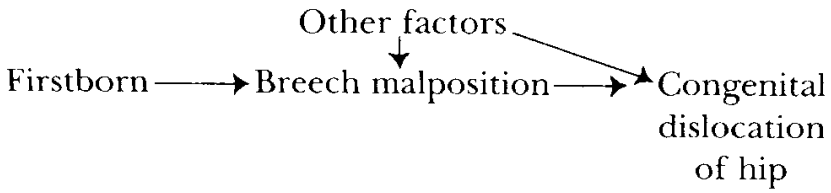
	Firstborn			Not firstborn		
	Cases (CDH)	Controls	Total	Cases (CDH)	Controls	Total
Number born by breech	48	16	64	12	4	16
Number not born by breech	52	64	116	88	116	204
Total	100	80	180	100	120	220
Percentage born by breech	48	20		12	3	

they are firstborn, whereas Table 1–5 indicates that infants with congenital dislocation of the hip are not more likely to have been firstborn, once the association between breech delivery and congenital dislocation of the hip is taken into account. Breech malposition is not a confounder for the association between being firstborn and congenital dislocation of the hip because although breech malposition is causally related to congenital dislocation of the hip, being firstborn contributes to the risk of breech malposition. One concludes that being firstborn is indeed a risk factor for congenital dislocation of the hip, but only because it increases the risk for breech

Table 1–5. Number and percentage of firstborn among infants with congenital dislocation of the hip (CDH) and among control infants according to whether the delivery was breech, hypothetical data

	Breech delivery			Not breech delivery		
	Cases (CDH)	Controls	Total	Cases (CDH)	Controls	Total
Number firstborn	48	16	64	52	64	116
Number not firstborn	12	4	16	88	116	204
Total	60	20	80	140	180	320
Percentage firstborn	80	80		37	36	

malposition, which is the more direct cause of congenital dislocation of the hip. The sequence can be illustrated as follows:



Obviously, other factors, many of which are not yet identified, also influence the likelihood of breech malposition. Sorting out such associations is a major concern in data analysis and interpretation.

In some instances it is not possible to determine whether a given variable is in fact a confounder, is a link in the causal chain, or is by chance associated with the risk factor and disease in a particular study. Generally, an understanding of the biologic relationships among the exposure, possible confounder, and disease is necessary to be certain of the true role of a given variable.

Effect modification, sometimes referred to as statistical interaction, also needs to be considered when epidemiologic studies are designed, analyzed, and interpreted. Effect modification occurs when the magnitude of the chosen measure of association between a causal agent and a disease differs according to the level of a third variable (or according to the levels of two or more variables). For instance, when all age groups are considered together, females have two to three times the risk for hip fractures as males (2). However, this overall ratio disguises the fact that at young ages males are at higher risk than females, whereas at older ages females are at considerably higher risk than males. The association between gender and hip fracture is thus modified by age, which in this instance is a surrogate measure of the high prevalence of osteoporosis in older women and of the propensity to severe trauma in young males. The association between the Epstein–Barr virus (EBV) and African Burkitt's lymphoma exists mainly in the presence of malaria (28), which is generally designated a cofactor by infectious disease epidemiologists, but which could be considered an effect modifier as well. The association between HSV 2 and cervical cancer (if indeed the association is a causal one) is modified in some unknown way by geographic locality; the association is considerably less strong in Japan than in the United States, for instance.

MEASUREMENT

Another major concern of epidemiologists is measurement: measurement of exposures, measurement of diseases, measurement of

confounding variables, and measurement of effect modifiers. Although some variables such as ABO blood groups or the occurrence of a hip fracture can be measured with a relatively high degree of accuracy, epidemiologic studies are often limited by inaccurate measurement. A great deal of interest currently exists in the role of diet in the etiology of many chronic diseases, but determining the foods consumed at the time the disease process actually began is often exceedingly difficult, since most people cannot remember what their diet was 20 or 30 years previously. Even the measurement of current dietary intake is difficult because of daily variations in the diet and memory problems. Thus, failure to find an association between diet and disease may indicate either that no relationship exists or that measurement is so poor that an association that does exist cannot be detected.

Failure to measure confounding variables accurately may mean that the confounder is not adequately taken into account. For instance, cigarette smoking and coffee drinking are highly correlated. If one wants to measure the association between coffee drinking and disease, smoking habits must be carefully taken into account; otherwise, an association may be found between coffee drinking and disease that is in reality attributable to the confounding effect of smoking, which has not been measured with sufficient precision.

Another concern is that ascertainment of exposure be comparable in diseased and nondiseased individuals and that ascertainment of disease occurrence be comparable in exposed and unexposed persons. Lack of comparability, frequently referred to as differential misclassification, can result in positive associations between exposure and disease when none exists and in failure to find associations when an association does in fact exist. The possible effects of measurement error thus need to be considered along with various other issues in designing and interpreting studies, and will be discussed in Chapters 11 through 13.

CONCLUSION

Major concerns of epidemiologists trying to learn about disease causation through observational studies include choosing the most reliable sources of data, the correct study designs, representative study subjects, the appropriate methods of measurement; quantifying the magnitude of various risks associated with exposures; controlling for confounding variables; identifying direct and indirect

causes of disease; and detecting effect modification. In practice, it may be impossible to meet all these objectives to the extent desired because study subjects may not choose to participate, optimal measurement may not be feasible, and a variety of other problems may arise. Therefore, it is important to recognize the effects of various inadequacies in different situations, since different inadequacies can affect the study results in different ways.

It is hoped that this introductory chapter makes it apparent that conclusions drawn from epidemiologic studies, as with other types of scientific inquiry, are not always final. Many require additional confirmatory epidemiologic or laboratory studies, the experimental reproduction of the exposure-disease association, or ascertainment of the effect of removal or modification of the suspected risk factor. What was considered a direct cause of a disease at one time may later be found to be an indirect cause when more information has been obtained on biologic mechanisms. What was believed to be a causal association may be found to be attributable to a confounding variable recognized at a later time. The best method of measurement at one point in time may be supplanted by a better method developed subsequently. In any one study, a reported association may have occurred by chance, especially when many possible associations are being considered, and even a true association may not apply to all population groups.

Thus, although it is important to use the most appropriate methods in any one study, it is also essential to keep an open mind as new knowledge accumulates from further epidemiologic, laboratory, and other types of studies and as attempts are made to replicate the results of even the most carefully executed individual studies. It is in this way that knowledge of disease etiology generally evolves.

REFERENCES

1. Acheson, R. M., and A. B. Collart. 1975. "The New Haven survey of joint disease. XVII. Relationship between some systemic characteristics and osteoarthrosis in a general population." *Ann. Rheum. Dis.* 34:379–87.
2. Buhr, A. J., and A. M. Cooke. 1959. "Fracture patterns." *Lancet* 1:531–36.
3. Burkitt, D. P. 1969. "A children's cancer dependent on climatic factors." *Nature* 194:232–34.
4. Centers for Disease Control. 1981. "*Pneumocystis* pneumonia." *MMWR* 30:250.
5. Centers for Disease Control. 1983. "Human T cell leukemia virus infections in patients with acquired immune deficiency syndrome: preliminary observations." *MMWR* 32:233–34.

6. Dawber, T. R., G. F. Meadors, and F. E. Moore, Jr. 1951. "Epidemiologic approaches to heart disease: the Framingham Study." *Am. J. Public Health* 41:279–86.
7. de The, G., A. Geser, N. E. Day, P. M. Tukei, E. H. Williams, D. P. Beir, P. G. Smith, A. G. Dean, G. W. Bornkamm, P. Feorino, and W. Henle. 1978. "Epidemiological evidence for causal relationship between Epstein–Barr virus and Burkitt's lymphoma from Ugandan prospective study." *Nature* 274:756–61.
8. Doll, R. 1953. "Bronchial carcinoma: incidence and aetiology." *Br. Med. J.* 7:521–27, 585–90.
9. Epstein, M. A., B. G. Achong, and Y. M. Barr. 1964. "Virus particles in cultured lymphoblasts from Burkitt's lymphoma." *Lancet* 1:702–3.
10. Evans, A. S., F. Cox, G. Nakervis, E. Opton, R. Shope, A. X. Wells, and B. West. 1974. "A health and seroepidemiological survey of a community in Barbados." *Int. J. Epidemiol.* 3:167–75.
11. Evans, A. S., C. Jeffrey, and J. C. Niederman. 1971. "The risk of acute respiratory infections in two groups of young adults in Colombia, South America. A prospective seroepidemiologic study." *Am. J. Epidemiol.* 93:463–71.
12. Evans, A. S., J. C. Niederman, and R. W. McCollum. 1968. "Seroepidemiologic studies of infectious mononucleosis with EB virus." *N. Engl. J. Med.* 279:1121–27.
13. Ferris, B. G., and D. O. Anderson. 1962. "The prevalence of chronic respiratory disease in a New Hampshire town." *Am. Rev. Respir. Dis. 86:165–77.*
14. Gallo, R. C., S. Z. Salahuddin, M. Popovic, G. M. Shearer, M. Kaplan, B. F. Haynes, T. J. Palker, R. Redfield, J. Oleske, B. Safai, G. White, P. Foster, and P. D. Markham. 1984. "Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS." *Science* 224:500–2.
15. Gregg, N. M. 1941. "Congenital cataract following German measles in the mother." *Trans. Ophthalmol. Soc. Aust.* 3:35–46.
16. Jordan, W. M. 1961. "Pulmonary embolism." *Lancet* 2:1146–47.
17. Kamoi, M. 1960. "Statistical study on relation between breast cancer and lactation period." *Tohoku J. Exp. Med.* 72:59–65.
18. Kannel, W. B., T. R. Dawber, A. Kagan, N. Revotskie, and J. Stokes, III. 1961. "Factors of risk in the development of coronary heart disease—six-year followup experience." *Ann. Int. Med.* 55:33–50.
19. MacKenzie, I. 1965. "Breast cancer following multiple fluoroscopies." *Br. J. Cancer* 19:1–8.
20. MacMahon, B., and M. Feinleib. 1960. "Breast cancer in relation to nursing and menopausal history." *J. Natl. Cancer Inst.* 24:733–53.
21. MacMahon, B., T. W. Lin, C. R. Lowe, A. P. Mirra, B. Ravnihar, E. J. Salber, D. Trichopoulos, V. G. Valaoras, and S. Yuasa. 1970. "Lactation and cancer of the breast. A summary of an international study." *Bull. WHO* 42:185–94.
22. MacMahon, B., P. Cole, T. M. Lin, C. R. Lowe, A. P. Mirra, B. Ravnihar, E. J. Salber, V. G. Valaoras, and S. Yuasa. 1970. "Age at first birth and breast cancer risk." *Bull. WHO* 43:209–21.
23. Markush, R. E., and D. G. Seigel. 1969. "Oral contraception and mortality trends from thromboembolism in the United States." *Am. J. Public Health* 59:418–34.
24. Martland, H. S. 1931. "The occurrence of malignancy in radioactive persons." *Am. J. Cancer* 15:2435–516.