Environmental Health Perspectives
Vol. 52, pp. 3-8, 1983

Practical Limitations of Epidemiologic Methods

by Abraham M. Lilienfeld*

Epidemiologic methods can be categorized into demographic studies of mortality and morbidity and observational studies that are either retrospective or prospective. Some of the limitations of demographic studies are illustrated by a review of one specific mortality study showing possible relationship of nuclear fallout to leukemia. Problems of accuracy of diagnosis or causes of death on death certificates, estimates of population, migration from areas of study, and the issue of "ecological fallacy" are discussed. Retrospective studies have such problems as recall of previous environmental exposure, selection bias and survivor bias. In environmental epidemiology, prospective studies have been used. The problems associated with these studies are illustrated by reviewing some of the details of the study of effects of microwave radiation on embassy employees in Moscow. The study population had to be reconstructed, individuals had to be located and information on exposure status had to be obtained by questionnaire. The relatively small size of the exposed group permitted the detection of only fairly large relative risks. Despite these limitations, epidemiologic studies have been remarkably productive in elucidating etiological factors. They are necessary since "the proper study of man is man."

The contribution of epidemiologic methods in elucidating the determinants and etiological factors in human disease has a long history. This has been particularly true with regard to such features of the environment as occupation and such etiologic factors as ionizing radiation and chemicals. One need only recall Greenhow's work in 1858 where his analysis of mortality data indicated higher mortality rates for those who worked in certain occupations (*1*). Prior to discussing some practical limitations of epidemiologic methods, it should be understood that the limitations do not negate the many positive contributions of epidemiology.

It would be well first to review briefly the different epidemiologic methods that have been used to elucidate the deleterious effects of human exposure to environmental agents. These methods can be broadly classified into two general categories: (1) demographic studies of mortality and morbidity and (2) observational studies that are either retrospective or prospective. This second category of epidemiologic studies consists of studying the exposure or history of exposures of individuals to environmental agents. This contrasts with demographic studies where the investigator generally studies the characteristics of population groups as a whole, i.e., in the aggregate.

In retrospective case-control studies, one usually starts with a group of individuals who have the disease or effect of an agent, determines their prior history of exposure and compares the frequency of such a history with that ascertained in a comparison or control group. In a prospective study, the investigator starts with an exposed group and follows that group to determine the frequency of occurrence of possible effects, which is then compared with an otherwise similar group that has not been exposed. This can be accomplished in two ways: exposed and unexposed individuals are selected and concurrently followed to determine the frequency of occurrence of possible effects; alternatively, for suitable groups, an investigator goes back to a certain point in time and traces the exposed and unexposed groups of individuals to the present, to determine the frequency of occurrence of disease or other possible effect. These different types of observational studies have resulted in ascertaining the effects, either in the form of mortality or morbidity for a variety of environmental agents such as ionizing radiation

*Department of Epidemiology, The Johns Hopkins University, School of Hygiene and Public Health, 615 North Wolfe Street, Baltimore, MD 21205.

or specific occupational exposures, some of which will be presented during the course of this symposium.

Epidemiology also includes the experimental or intervention approach with regard to human population groups. There is no need to point out that the ability of the investigator to intervene and exercise direct control over the population groups he or she is studying markedly increases the strength of the derived inferences. In observational studies, by contrast, the investigator essentially accepts conditions as they are. Epidemiologic studies of environmental factors are predominantly observational in character. This represents a major practical limitation and will become evident when specific illustrations of different types of studies are reviewed. The fact that the investigator has to accept the types of information of interest as they have been or are being recorded results in a variety of problems, thereby limiting the inferences that can be derived from the study. On the other hand, since the "proper study of man is man," and the ability to intervene in the activities of human population groups is limited, observational studies of man also represent one of the strengths of epidemiologic methods. After all, the objective is to make inferences about the effect of environmental agents on humans, the major subject of interest.

Demographic Studies

We shall first briefly review demographic studies of mortality or morbidity. An example of such a study is the one recently reported by Lyon et al. (*2*). The investigators found a higher mortality rate from childhood leukemia in areas of Utah that were exposed to radioactive fallout from atmospheric tests of nuclear weapons carried out in Nevada from 1951 through 1958. Lyon et al. reviewed all cancer deaths for those under 15 years of age that had occurred in the entire state during 1944 to 1975. They then defined a low exposure group who were Utah residents during the periods 1944-50 and 1958-75, with a high exposure group during 1951-58. The population censuses of 1940, 1950, 1960 and 1970 and an estimated population for 1975 were obtained. By interpolating between these census years, the number of individuals in the appropriate age groups for each intercensal year was estimated in high and low exposure groups. Death certificates for childhood cancers were obtained and classified into deaths from leukemia and other cancers for the appropriate periods of time, age groups, and high and low exposure groups. Mortality rates were then calculated for the high and low exposure groups and are shown in Table 1. It is sufficient to state that these investigators found a significantly increased mortality from leukemia among children in the high exposure group.

This particular report illustrates some of the limitations of such mortality studies. The age-specific mortality rates consist of a numerator (the deaths) and a denominator (the population), each of which is subject to error. Regarding the numerator, questions can be raised concerning the accuracy of the diagnoses of leukemia on the death certificate, particularly over the time period 1951 to 1975, when improvements in medical care and methods of diagnoses occurred. These investigators were aware of this possibility and indicated that studies of deaths in Pennsylvania (reported in 1958) and the results of a study of the Utah Cancer Registry (which was not in existence during the early period of the study) showed an accuracy of 90%. However, I doubt whether findings of accuracy in Pennsylvania and of a later time period in Utah could be applied to deaths in Utah that had occurred in earlier time periods. Relative to this point, another analysis of these same data indicated a negative relationship between exposure to radioactive fallout and childhood cancer other than leukemia (*3*). Perhaps

Table 1. Leukemia mortality rates per 100,000 population of both sexes, adjusted for age and sex, according to high exposure and low exposure groups, for counties with high and low fallout.[a,b]

	Leukemia mortality rates per 100,000[c]		
Area	Low exposure (1944-50)	High exposure (1951-58)	Low exposure (1959-75)
High fallout counties	2.10	4.39	1.96
	(7)	(32)	(10)
Low fallout counties	3.84	4.21	3.28
	(44)	(152)	(112)

[a]Adapted from Lyon (*2*).
[b]Adjustment was by the direct method to U.S. 1960 white population.
[c]Numbers of deaths in parentheses.

these different results partially reflect differences in diagnostic practices during this time period.

More recently, Percy et al. studied the accuracy of cancer mortality data using cancer deaths during 1970 and 1971 in eight of the nine areas included in the Third National Cancer Survey (*4*). Death certificates with an underlying cause of death of cancer were compared to the hospital diagnoses for 48,826 resident cases. The coded underlying cause of death on the death certificate was found to be accurate for about 65% of cancer deaths. About 80% of cases of lymphocytic, 70% of myeloid and 57% of monocytic leukemia cases were so reported on the death certificate. Of deaths with these diagnoses on the certificate, 86% of lympocytic, 92% of myeloid and 54% of monocytic and 34% of other and unspecified forms of leukemia were confirmed by the hospital diagnosis. In their report, Lyon et al. did not indicate that the diagnoses on death certificates had been validated, at least, by comparison with other medical records. Clearly, it would be desirable to do so.

Secondly, population estimates for each year were estimated from interpolations between census years on a county basis and totaled. There may be a substantial error in this procedure when done for each age and sex group and county, because of migration and population changes. Such migration could be out of state or between counties in the designated high and low exposure areas. The investigators discussed this possibility and concluded that if it had occurred, it would only weaken the association between exposure and leukemia. But, without actually knowing the extent and type of migration, it is difficult to determine its effect on the findings.

Excess mortality from leukemia was concentrated in the 10 to 14 year age group, with a lesser excess in the 5 to 9 year age group and none in the under-4 group. The places of birth and length of residence in Utah of the decedents were not known. The determination of such information is important to help assess the validity of the findings.

Another major difficulty in analyzing data of this type is that relationships of factors occurring in communities such as counties are generally termed "ecological correlations" and may suffer from an "ecological fallacy." The communities may differ in many factors and one or more of those, other than the one being studied, may be the underlying reason for the differences in their observed mortality or morbidity experience. For example, the observed inverse relationship between water hardness and cardiovascular disease has been interpreted by many as representing an ecological fallacy (*5*).

It is apparent from considering only a few aspects of this study that there are limitations to the inferences that can be derived from such mortality studies. However, it is important to point out that these demographic studies provide opportunities in suggesting leads to possible effects of environmental agents. Such leads stimulate more intensive and refined investigations, designed to take into account the limitations already mentioned.

Observational Studies

Retrospective studies where cases of disease are compared with some control group, obtain information on past exposure either by interview or by review of past medical records. When interviewing individuals, there may be problems of recall which differ in the groups being compared. Experience with reviews of past medical records regarding exposures of various types indicates that the information that had originally been obtained for some other purpose is usually not adequate for the objective of the particular study. Also, such records usually do not have adequate or sufficient information on the degree of exposure to the environmental agent of interest.

Another problem in retrospective studies results from the possible influence of selection of the groups being compared, i.e., selection bias. This possibility is generic to all types of retrospective designs. It is also necessary to consider the fact that in retrospective studies, the investigator is usually studying the survivors of the exposed group. If the exposures had an effect that resulted in increased mortality during the period between exposure and the time of the study, it may well be that the effect would be underestimated.

Retrospective studies are most productive and less subject to the limitations mentioned if the relationship between exposure and disease (or effect) is of a relatively large magnitude, that is, if the relative risk is at least 3- to 4-fold.

There are other limitations to retrospective studies which we will not discuss here. But, in order to overcome some of the problems mentioned, most studies of the effects of environmental agents have tended to use the prospective or cohort approach.

To illustrate some problems associated with the prospective approach, I would like to briefly review a study in which we were recently engaged, namely, the Foreign Service Health Status Study (*6*). This was a study of the possible effects on mortality and morbidity due to exposure to microwaves among U.S. government employees who had served at the American Embassy at Moscow from 1953 to 1976. Their experience was compared with those in nine Eastern European embassies or consulates during the same period. During 1953 to 1975, microwaves were beamed at the upper half of the Embassy's Chancery Building at a dose of 5 $\mu W/cm^2$ for 9 hr/day. From June 1975 to February 1976, a different part of the building was exposed and the dose increased to 15 μW for 18 hr/day. In the course of the study, over 1800 employees at the Moscow Embassy were finally identified as well as a group of over 2588 employees at comparison posts.

During 1976 to 1978 more than 95% of these employees were located; it was recorded whether they were living or dead (Table 2). Medical records were obtained on 84% of State Department employees and 43% of non-State Department employees; these included nearly 22,000 individual medical examinations, which was similar in the Moscow and comparison groups. The very low percentage for non-State Department employees resulted from the fact that they were mainly military personnel with a unique set of administrative and logistic problems.

An attempt was made to obtain a health history questionnaire from each employee whose current location could be determined by mail and telephone interview. In addition to the health status, the working and living areas were sought while in Moscow from which the exposure to microwave radiation was determined. Questionnaires were completed for only 52% of State Department employees (59% from the Moscow group and 48% from the comparison group) and only 38% of the non-State Department employees (48% from the Moscow group and 34% from the comparison group).

Analysis of the mortality experience showed no evidence that the Moscow group had experienced a higher total mortality or any specific causes of death up to the time of the study than had the comparison groups. Table 3 presents data for total mortality experience.

Every possible effort was made to find if there were any differences in nonfatal morbid conditions between the Moscow and comparison groups. Literally hundreds of comparisons were made based on information obtained in the medical records of the two groups of employees. Only two differences, from the medical review, stood out: (1) the Moscow male employees had a 3-fold higher risk of acquiring protozoal infections between the time of arrival at the post and the time of last observation than did the comparison employees, and (2) both men and women of the Moscow group were found to have slightly higher frequencies of most of the common kinds of health conditions investigated. However, because the group was a very heterogeneous one, it was diffi-

Table 2. Final status of tracing, medical records reviewed, health history questionnaire response for State Department employees by post.

Final status	Moscow	Comparison	Total
Total number of employees (100%)	1149	1843	2992
Traced, % of total	95	98	97
Medical records reviewed, % of total	81	85	84
Number and percent of total sent health history questionnaire			
Number	1040	1643	2683
%	91	89	90
Returned questionnaire, % of those sent	59	48	52

Table 3. Standardized mortality ratios (SMR) for State Department employees by sex and service post.[a]

	Sex	Person-years	Observed deaths	SMR (95% confidence limits)
Moscow only	M	5,135	14	0.43 (0.2–0.7)
Comparison only	M	14,076	75	0.53 (0-4–0.7)
Both Moscow and comparison	M	3,222	10	0.48 (0.2–0.9)
Total	M	22,433	99	0.51 (0.4–0.6)
Moscow only	F	2,975	9	0.96 (0.4–1.8)
Comparison only	F	8,205	28	0.80 (0.5–1.2)
Both Moscow and comparison	F	1,233	1	0.24 (0.0–1.3)
Total	F	12,413	38	0.78 (0.6–1.1)

[a]SMR computed by using United States mortality experience specific for sex, color, age and calendar time applied to study persons from their time of arrival at first study post to time of follow up to determine expected number of deaths. Confidence limits were derived assuming a Poisson distribution for deaths.

cult to conclude that these conditions could have been related to exposure to microwave radiation, since no consistent pattern could be found.

Based on the analysis of responses to the health history questionnaire, the Moscow group was found to have a variety of symptoms after their study tour more frequently than the comparison group: more depression, more irritability, more difficulty concentrating and more memory loss. However, no relationship was found between the occurrence of these symptoms and exposure to microwaves; in fact, these four symptoms, which showed the strongest differences between the Moscow and comparison groups, were all found to have occurred most frequently in the group with the least exposure to microwaves. In view of the well-publicized possibilities of the increased danger to their health and that of their children, it is not at all surprising that the Moscow group might have had an increase in symptoms such as those reported. The statistically significant disease conditions that occurred after the first tour of duty as ascertained on medical record abstracts, are shown in Table 4 in the form of standardized morbidity ratios by sex and exposure status. Among males, statistical significance for cancer morbidity is mainly due to the low frequency in the uncertain exposed group. In the two other groups of conditions, the frequency is highest in the unexposed group. Among females, no statistically significant conditions were found among these exposure groups.

Many problems were encountered in this study. A major one was the identification of the study population. There were no procedures for maintaining the records of individuals (except for those currently employed by the Department of State) who had served tours of duty at foreign embassies and consulates. Thus, it was necessary to reconstruct the population who had served at any of the study posts during the period 1953 to 1976.

Although it was anticipated that the study group would be most responsive to completing a mailed questionnaire requesting the desired information, the response rate was disappointing (33%). This made it necessary to utilize telephone interviewing, which was very productive, but time and financial constraints did not permit using it to the fullest extent possible. Therefore, the final response rate was 52% for the State Department employees and 38% for non-State Department employees. The study population was very mobile, and it was often necessary to telephone overseas posts, since there was no definitive current list of the location of many active employees.

The limitation due to the relatively low response to the health history questionnaire was somewhat balanced by the large amount of information available in the medical records which contained the findings of the routine, periodic examinations and special examinations performed on this civil and military service population. Some form of health status information, either from a medical record or a completed questionnaire, was available for 92% of the State Department and 64% of the non-State Department groups.

A major problem, mainly due to the incomplete response to the health history questionnaire, was the classification of exposure to the microwave beams for the Moscow embassy employees. It was possible to determine exposure status only if a health history questionnaire was returned and then, only if the individual remembered where he or she had worked and lived within the embassy. Many could not remember enough details of their working and living locations to allow classification of their exposure status.

Another factor must also be considered in interpreting the findings of the study; namely, whether the groups studied were large enough to permit a reasonable chance of detecting statisti-

Table 4. Standardized morbidity ratio (SMBR) for statistically significant conditions reported, after first tour of duty, on medical abstracts by sex and exposure group in Moscow.[a]

Condition (ICDA, 8th)	Sex	Unexposed No.	Unexposed SMBR	Exposed No.	Exposed SMBR	Uncertain No.	Uncertain SMBR	*p*
		(PY = 1912)[b]		(PY = 1787)[b]		(PY = 6827)[b]		
All cancer except skin (140-209)	M	4	1.5	6	2.3	3	0.4	0.02
Other diseases of upper respiratory tract (500-508)	M	22	1.6	17	1.3	41	0.8	0.03
Nervousness and debility (790)	M	20	1.7	9	0.9	30	0.8	0.05
	F	None significantly different						

[a]SMBR adjusted for year of entry and age at entry.
[b]PY = person-years.

Table 5. Minimum detectable excess relative risks or risk ratios for different comparison groups by sex, source of information and event rate per person-year.[a]

		Moscow vs. comparison posts			Moscow
			Morbidity		Exposed vs. unexposed
Sex	Event rate per person-year	Mortality	Medical records	Health history questionnaire	Morbidity only
M	1/100	1.3-1.4	1.4-1.5	1.5-2	2-3
M	1/1000	2.2-2.5	2.5-3	3.5-4	5-6
M	1/10,000	7-8	8-10	10-15	25-50
M	1/100,000	30-50	50-75	75-100	>100
F	1/100	1.6-1.8	2-2.5	2-3	3-4
F	1/1000	3.5-4	4-5	5-6	10-20
F	1/10,000	15-20	15-20	25-50	50-100
F	1/100,000	>100	>100	>100	>100

[a]Assuming $(1 - \beta) = 0.80$ and $\alpha = 0.05$.

cally significant excess risks that may have resulted from exposure to microwaves. The ability to detect excess risks of any particular disease or condition was determined by the size of the excess risk, the incidence of the condition, and the number of person years of observation of the two groups to be compared. On assuming this probability to be 0.80, at a significance level of $p = 0.05$, Table 5 shows the ranges of excess risks, expressed as relative risks, which the present study could have detected for four hypothetical event rates. Clearly, except for relatively frequent events, only moderate or large differences between the comparison groups could be detected. The size of the study population, and particularly that of the identified exposed population in Moscow, was not sufficient to detect excess risks that were less than 2-fold for many of the medical conditions studies. For all malignant neoplasms, which occurred with a frequency of about 1 per 1000 among males and 5 per 1000 among females, a statistically significant 2-fold increase could have been detected. However, for specific types of neoplasms which occurred with a lower frequency, the study group was not sufficiently large to permit one to find statistically significant increased risks unless they were unusually large, approximately of the order of a 5- to 10-fold excess or higher.

This last problem, that of size of study population, is one particularly relevant to the specific topic of setting standards for environmental agents. When setting such standards, it is important to be able to measure effects at low levels of exposure. To determine such effects with the desired precision, it is usually necessary to have information on fairly large population sizes, and this gets to be a difficult task.

General Comments

In this presentation, I have attempted to discuss rather broadly some of the problems associated with epidemiologic studies. This review has not been exhaustive. Yet, despite these limitations, these types of studies have provided us with much of our knowledge of the effects of environmental agents on human population groups. Some of this is due to the fact that several of the limitations mentioned can be taken care of by improving study designs. Other limitations are overcome by replication of studies in different population groups and taking advantage of information from special types of situations. From such replication, there can be a convergence of observed facts leading to inferences with sufficient credibility to make policy decisions. This general issue will no doubt be discussed in later papers at this symposium.

REFERENCES

1. Greenhow, E. M. Papers Relative to the Sanitary State of the People of England. General Board of Health, London, 1858, Gregg International Publishers, Ltd., London, 1973.
2. Lyon, J. L., Klauber, M. R., Gardner, J. W. and Udall, K. S. Childhood leukemias associated with fallout from nuclear testing. New Engl. J. Med. 300: 397–402 (1979).
3. Land, C. E. The hazards of fallout or of epidemiologic research. New Engl. J. Med. 300: 431–432 (1979).
4. Percy, G., Stanek, E., III, and Gloeckler, L. Accuracy of cancer death certificates and its effect on cancer mortality statistics. Am. J. Publ. Health 71: 242–250 (1981).
5. Comstock, G. W. The epidemiologic perspective: water hardness and cardiovascular disease. J. Environ. Pathol. Toxicol. 4: 9–25 (1980).
6. Lilienfeld, A. M., Tonascia, J., Tonascia, S., Libauer, C. H., Cauthen, G. M., Markowitz, J. A. and Weida, S. Evaluation of Health Studies of Foreign Service and Other Employees from Selected Eastern European Posts. National Technical Information Service, Springfield, VA, 1979.