Methods For Assessing the Effects of Mixtures of Chemicals
Edited by V. B. Vouk, G. C. Butler, A. C. Upton, D. V. Parke and S. C. Asher

Methodological Issues in Epidemiological Studies of Disease Clusters and Environmental Contamination

Peter Greenwald and Gladys Block

ABSTRACT

Contamination episodes or disease clusters provide unique opportunities to evaluate health effects and investigate aetiology. Epidemiological approaches to research in this area require first establishing the existence of a true cluster. Identification of both the cases and the denominator population, determining individual exposure, collecting data on potential confounders, and linking epidemiological and laboratory studies all need careful methodological attention. The Love Canal chemical dump and other examples are used to illustrate these and related issues.

1 INTRODUCTION

The observation of an apparent disease cluster, or the occurrence of an episode of environmental contamination, presents both valuable opportunities and methodological problems for studies of epidemiology. Studies may uncover leads to the causes of the disease and can evaluate the effect of the environmental contamination, both of which may be important to the assessment of human responses to multiple chemical exposures. This knowledge may in turn help to prevent future cases of the disease and can guide policy makers in their decision making. Unfortunately, epidemics and episodes of environmental contamination tend to be accompanied by intense public anxiety which often pressures policy makers for decisions before scientific data can be collected and analysed. For example, contamination of our environment with synthetic chemicals is currently a prominent public health concern. Huge amounts of chemical wastes, expelled into our environment since World War II, have raised concerns about the development of specific diseases and about the potential long-term impact on the quality of life and the ecology. Yet, despite our suspicions, we have few unequivocal data linking human illness to low or moderate environmental exposures of many potentially toxic substances.

In such situations where the pressures and stakes are high, there is a clear need for a high degree of objectivity and careful use of the scientific method in the collection, analysis, and interpretation of the epidemiological data. Attention to the methodological problems involved in cluster and incident investigation by policy makers and potential investigators will enhance the value of the studies and the public policy derived from them. In this article, we discuss (a) the confounding elements that can invalidate study data and (b) elements of study design that are prerequisites to establishing the objectivity of epidemiological data. We illustrate these methodological issues with historical studies of disease clusters and environmental contamination, and comment on their implications for policy making.

2 METHODOLOGICAL ISSUES

2.1 Issue One: The Existence of a True Disease Cluster Must Be Established

Whether it is the disease or the contamination episode which is first identified it must be established that the occurrence of the disease is a true cluster, i.e. that the incidence is greater than what would otherwise be expected in that population. This involves determining the cases and the population from which they arose, and errors can be introduced in both phases. Sophisticated statistical techniques have been used to identify clusters (Greenwald *et al.*, 1979; Knox, 1964; Mantel, 1967), but may not be feasible for studying a given episode. Often in the field study situation, the first consideration should go to identifying the cases, confirming the diagnosis of disease, and describing the demographic and exposure-related characteristics of the affected individuals. In an ideal situation, however, defining the denominator population, a crucially important and often neglected aspect of episode studies, is best done in advance of case ascertainment. One can then determine the disease rate in this defined population and compare it to the expected rate derived from general population incidence figures or from a comparison population (control group). Defining denominator populations after the cases have been reported carries with it the risks of circular reasoning; thus, defining a denominator population around cases that have already occurred may be a type of 'gerrymandering' which can produce misleadingly high (or low) apparent disease rates. A possible solution to this dilemma is to base the denominator population on the initial cases that brought the situation to the attention of the investigator, to search for additional cases, then to exclude the initial cases from further analysis of whether a cluster exists. If field investigation uncovers enough additional cases to yield a high rate with these alone, there is very likely a true cluster.

Identifying cases can be either aided or confounded by an aroused population. Concerned lay groups often have no knowledge of the usual frequency of the disease and may believe that nearly all cases are causally linked to a special exposure (e.g. an episode of environmental contamination). The intensity of their concern may contribute to complete identification of the affected group, or to overdiagnosis. At the same time, case ascertainment is likely to be less complete in a comparison group, thus introducing a potential bias. Objective sources of identification include disease registries, such as population-based cancer registries and health agency records of birth, death, and disease. Unfortunately, such records rarely have information about exposure from, for example, a person's occupation, industrial environment, or life-style risk factors. Registries of exposure with linkage to disease registries would be a useful innovation for making episode studies more systematic and analytical.

The likelihood of successfully uncovering an aetiological factor in a disease cluster depends on the frequency of the disease and the type of exposure (Greenwald and Lawrence, 1979). An example of an episode in which a disease cluster was clearly established (Herbst *et al.*, 1971) and quickly confirmed (Greenwald *et al.*, 1971) was the causal association between diethylstilboestrol ingestion during pregnancy and adenocarcinoma of the vagina in daughters born of these pregnancies. The extreme rarity of the disease allowed for confidence that a disease cluster really existed. Because the exposure was to a drug, retrospective identification of the exposure was easier than in some environmental situations where, for example, the soil might be contaminated but it is unclear that any of the chemical was absorbed into the body. Another example in which the rarity of the disease aided in the discovery of its cause may be that of liver angiosarcoma after vinyl chloride exposure (Creech and Johnson, 1974).

When exposure leads to a common disease, however, establishing the cluster and aetiology usually requires an intensive research effort and may involve severe problems in interpretation. Two studies illustrate the potential pitfalls of this type of research, both involving the use of records. In the first, an apparent cluster was an artifact of differences in the quality of diagnostic effort for the case group as compared with the rest of the population; in the second, a known cluster could not be uncovered through the use of a registry.

In the first example, death certificates of men with brain tumours were examined to see if an occupation or industry appeared as unusually frequent. One company was found to be more commonly listed on death certificates of brain tumour patients than on death certificates of a comparison group. A collaborative epidemiological study (Greenwald *et al.*, 1981) was then conducted by the company and health department. It was found that the apparent excess among employees was an artifact of ascertainment bias, i.e. workers had a more complete diagnostic evaluation leading to the brain tumour diagnoses than did patients in the state who did not work for the company. This difference in evaluation could account for the initially observed excess. In the general population, many brain tumours apparently are misdiagnosed as strokes or other medical problems, especially in the elderly. Epidemiologists refer to this kind of error as a sensitivity–specificity problem. In most epidemiological studies it is important to verify that the individuals under study actually have the disease, i.e. that the diagnosis is specific. In occupational studies or episode studies, it also may be essential to assure case–control comparability in detecting the disease; that is, equal sensitivity is needed in making the diagnosis.

In the second example, company employment and medical records were used as the initial basis for case ascertainment. Goldwater *et al.* (1965) studied 366 coal tar workers at a chemical plant in Buffalo, New York, and found that 96 (26 %) of these men developed bladder cancer. The finding was consistent with existing data showing that bladder cancer is associated with occupational exposures to the well-established carcinogens β-naphthylamine and benzidine (Cole and Goldman, 1975).

The data from this known epidemic provided an excellent opportunity to examine the usefulness of the New York State Cancer Registry as a monitoring system. Cancer registry data were analysed to determine if retrospective detection of this known occupational hazard would be possible (Greenwald *et al.*, 1979). The result was that the general frequency of bladder tumours in the area made it impossible to detect the small proportion due to this chemical plant. The fact that bladder cancer has a long induction period and had been diagnosed over many years among workers at one plant in the large population in Buffalo obscured its detection since spreading the cases over many years meant few occurred in each year, and these did not significantly raise the level in the community as a whole. In addition, registries often do not contain the occupational, exposure and life-style data which would permit identification of these risk factors. We conclude that better identification of exposure by cancer registries is necessary to enable environmental hazards that may cause a proportion of a common disease to be pinpointed.

2.2 Issue Two: The Collection of Data on Potentially Confounding Life-style Factors Is Important and Requires Specialists

Life-style risk factors contribute to the development of most human cancers (Doll and Peto, 1981). For a few of these, smoking for example, collecting valid data is relatively easy; for others such as diet, the task is more complex. When the disease under study is known to be related to life-style factors, such factors may well be the focus of the investigation; however, even when this is not the case, investigators of clusters or contamination episodes should make every effort to collect data on these potentially confounding factors. Smoking may dramatically affect the impact of environmental pollution (Surgeon General's Report, 1976); stress may alter biological response to insult, or possibly be causative; in some cases (Janerich *et al.*, 1981a) dietary factors can alter the action of carcinogens in animals and in man (NAS, 1982). Including an examination of such potential modulators of disease strengthens the interpretation of the results.

A major research advance in recent years has been the development of plausible hypotheses about how dietary factors may increase the risk for several human cancers, especially cancers of the gastrointestinal tract, breast, and prostate gland (NAS, 1982). These hypotheses derive from a convergence of evidence from many different research disciplines; epidemiological data are supported by parallel mechanistic studies in animals. Thus, it is important to include an examination of these factors in studies of disease aetiology. Unfortunately, as with data on occupational exposure history, data on life-style habits rarely exist in public records or registries. Existing survey data may offer useful leads but, generally, intensive field studies are necessary if significant progress is to be made. Furthermore, the methodological problems involved in collecting life-style data are complex and may lead to misinterpretation. Dietary assessment, for example, is an area in which several different approaches have been developed, each with its own merits and deficiencies (NAS, 1982).

The diet history method is widely considered a valid method for obtaining information about an individual's long-term eating habits; however, it does require an extensive interview by an interviewer trained in nutrition (Block, 1982). The record method, while probably the most accurate for obtaining quantities consumed, has the deficiency of being limited strictly to current diet. Furthermore, since it requires a high degree of participation by the subject, it does not lend itself well to large-scale survey investigations.

The 24-hour recall method is easily administered and its validity in assessing 24-hour intakes has been demonstrated (Gersovitz *et al.*, 1978; Madden *et al.*, 1976). Its severe limitation, however, is that it is appropriate only for determining mean intake values of groups; it cannot be validly used to categorize an individual's usual intake (Balogh *et al.*, 1971) and, thus, is inappropriate for estimating relative risk or odds ratios (Block, 1982).

The final method of evaluating diet, the frequency questionnaire, can be fairly easy to administer and has shown good preliminary evidence of validity (Block *et al.*, 1986; Hankin *et al.*, 1975; Willett *et al.*, 1985). However, because a standardized form does not yet exist, the design and evaluation of the questionnaire require participation by researchers experienced in this area. In addition, questions about the appropriate translation of the data into nutrient content, or the combining of laboratory and epidemiological approaches, suggest that this is an area in which considerable expertise, as well as further research, is required.

One means of avoiding some of the methodological problems is to study a stable population that has a fixed diet. For example, from Cali, Colombia, registry data, Correa *et al.* (1975) noticed a cluster of stomach cancer in persons living in high altitudes in certain rural areas. A later study reported geographical correlations of stomach cancers with the nitrate content of well-water (Cuello *et al.*, 1976), and another reported an association with a diet of corn (Haenszel *et al.*, 1976). Correa *et al.* (1976) examined 586 volunteers by gastroscopy and found that 75% of the high-risk population had more superficial gastritis or other dysplasias by age 25 than did the low-risk population. Correa and colleagues suggested that superficial gastritis, atrophic gastritis, and intestinal metaplasia are precursor lesions for gastric cancer. They described in detail the histopathology of the gastric lesions, beginning with the normal mucosa and progressing through various degrees of dysplasia to stomach cancer.

2.3 Issue Three: When the Disease under Study is Related to Environmental Contamination, Actual Exposure in the Cases is Difficult to Establish and Expected Disease Outcome May Be Uncertain

A fundamental problem in conducting a sound epidemiological study is the need to identify who is exposed, to what they were exposed, who develops the disease, and who are in the population from which the cases arose. These identifications may be perplexing in contamination episodes when specifics about levels of individual exposure and absorption into the body are unknown, and when the potential disease outcome is uncertain. Therefore, several issues of statistical methodology are important for both design and analysis. First, as mentioned earlier, failure to identify all cases could result in underestimating the incidence rate; on the other hand, public concern could lead to overdiagnosis and an overestimate of the rate. Second, the anticipated lag-time between exposure and disease onset also must be taken into account, because it can further complicate the problem of retrospectively identifying exposure or of prospectively following exposed individuals to establish the disease relationship. Third, choosing an inappropriate comparison population may lead to a falsely inflated or deflated risk estimate. Is it appropriate, for example, to have a control group or to determine expected rates from some general data base, such as a cancer registry? Finally, cases must be characterized for potential confounders (e.g. cigarette smoking and other life-style factors) if aetiology is to be correctly interpreted.

When possible, laboratory measures should be part of the epidemiological research; these can help establish both exposure and disease. A detailed exposure assessment preferably should include analysis of biological samples such as the blood and urine of exposed individuals. For risk assessment studies, more consideration should be given to the possibility of developing field laboratory teams capable of working together with epidemiologists in real situations. Laboratories prepared to study episodes in the field might very well complement epidemiological studies of episodes. Laboratory simulations also can be helpful for estimating past exposures or effects of exposures on animals. For example, at the Love Canal, a pilot test was attempted on the effects of contaminated basement air on caged pregnant mice, with control mice in cages equipped with air filters; however, the studies were discontinued because no laboratories were tooled up, experienced in this type of field research, and ready to respond to environmental episodes.

Several aspects of these problems of epidemiological research are illustrated in studies of the Love Canal chemical dump. The Love Canal landfill has become symbolic of the problem of chemical wastes. This rectangular, 16-acre tract of land located in Niagara Falls, New York, was a partially constructed, then abandoned, canal used for dumping chemical wastes from 1942 to 1953 (Nailor *et al.*, 1978; Tarlton and Cassidy, 1981). The largest user, a chemical and plastics company, deposited approximately 21 800 tons of chemical waste (some drummed and some not) into the canal. The city of Niagara Falls also used the site for municipal waste disposal until 1953. In 1954 a public elementary school was built right on top of the filled canal, and in the mid-1950s home construction was under way, adjacent to the landfill. Corrosion of the drums and the rising water level from heavy snows and rains apparently combined to bring chemicals to the surface in the mid-1970s. Residents complained of chemical odours. An oily coating was discovered in some basement sump pits. In spring 1978, the New York State Department of Health began intensive laboratory and epidemiological investigation. At this time there were 100 homes in which 235 adults and 143 children lived, with backyards directly abutting the canal; 410 students were enrolled in the elementary school. A total of 2065 people lived in the surrounding neighbourhood (within four blocks of the landfill); most lived in single-family homes. State Health Department studies related to the Love Canal are summarized in Table 1.

The early public health actions were taken largely on the basis of the initial laboratory demonstration of potentially hazardous chemicals in the soil and in basement air. Laboratory analysis of soil and sediment from the Love Canal showed the presence of 200 distinct organic chemical compounds, including many chlorinated hydrocarbons; dioxin at a level of 300 ppb was found in a storm sewer adjoining the canal. A composite sample of sumps from houses next to the canal contained toluene, chlorobenzenes, and other compounds. Basement air in some homes immediately adjacent to the canal showed the presence of similar compounds, including some that might be found in solvents used by home-owners.

Epidemiological studies were only beginning at that time. Epidemiological studies focusing primarily on adverse pregnancy outcomes were conducted by Dr Nicholas J. Vianna and others at the New York State Department of Health (Nailor *et al.*, 1978; Tarlton and Cassidy, 1981); the primary end-points were congenital defects, spontaneous abortions, and low birth weight (Table 2). Pregnancy outcomes for women living immediately adjacent to the canal (97th and 99th Streets) were compared to those for women living farther from the canal, whose homes were built on historically dry land ("Houses not built on contaminated fill", in Table 2) and those whose homes were built on former pathways of creeks that flowed through the area ("Houses possibly built on contaminated fill", in Table 2). Both the 'fill' and 'non-fill' areas are in the general vicinity of the Love Canal but outside of the immediately adjacent rows of houses. Just outside the Love Canal area, another neighbourhood was selected as a second control ('North of Colvin Blvd.', Table 2). In addition, comparisons were made to the expected numbers of spontaneous abortions from a report by Warburton and Fraser (1964). In these Love Canal studies, data were collected by means of field survey and examination of medical and vital records; data were included only when they could be confirmed by these records.

In general, the epidemiological study (Tarlton and Cassidy, 1981) showed suggestive increases in miscarriages and percentages of low birth weight, but the number of persons available for study was limited, the statistical power was low, and the inability to establish their real exposure made it uncertain whether and to what extent these adverse effects could be attributed to the chemical wastes. It was not possible to demonstrate that ill individuals actually had been exposed in the sense that chemicals had entered their bodies—much less could a dose–response relationship be demonstrated. A report by Janerich *et al.* (1981b) using data from New York State Cancer Registry found no evidence for higher cancer rates associated with residence near the canal as compared to those living elsewhere in upstate New York. A small chromosome damage study of doubtful credibility

Table 1 Love Canal studies by the New York State Department of Health[a]

Study	Results and comments
Laboratory studies	
Gas chromatography/mass spectrometric analysis of a composite sample of eight sumps from houses next to the canal	Sumps contained toluene, chlorobenzene, di-, tri-, tetra-, and pentachlorobenzene and other compounds
Home basement air testing for seven 'marker' chemicals by gas chromatography. Approximately 700 houses tested	Chloroform, benzene, trichloroethylene, toluene, tetrachloroethylene, chlorobenzene and chlorotoluene found in low levels in some basements. Result influenced by air flux through houses and solvents stored in basements by home-owners
Soil samples tested for lindane	Lindane found
Leachate from construction pit tested for dimethoxane (dioxin)	Dimethoxane (dioxin) found
Infrared aerial photography	Low-lying areas (swales) noted which intersected the canal and indicated during earlier years the possibility of preferential drainage paths for leachate
Soil sampling of 700 houses with 5000 soil samples collected; 1000 samples from transect; 50 samples from storm sewers. Soil samples from beneath school playground	Extensive contamination with toxic chemicals of soil near the canal documented. Attempt to define leachate migration inconclusive. Preliminary data suggest swales did not conduct leachate beyond first row of houses. Traces of lindane and dimethoxane (dioxin) found beneath school playground

Table 1 (*Cont'd*)

Study	Results and comments
Radiological survey of ambient (background) levels and soil samples	No serious radiological health hazard
Examination of various forms of plant life	No abnormalities found
Autopsies of 3 dogs, 1 blackbird, and 2 gulls	No histopathological lesions could definitely be ascribed to toxic chemicals
Teratological and pathological effects of Love Canal soil and air above soil on rats	In progress. Earlier Department study showed that large doses of chemical wastes affected mice livers and kidneys
Epidemiological and clinical studies	
Adverse pregnancy outcomes in the Love Canal area	Epidemiological study of residents living near the canal in June 1978. Study showed suggestive increases in miscarriages and percentages low birth weight, but the number of persons available for study and the inability to establish their real exposure make it uncertain to what extent these adverse effects can be attributed to chemical wastes
	No instances of chloracne and no excess of cancer, asthma, or epilepsy were found among these area residents
Blood counts and liver function tests	Twelve clinics were held, and about 4000 blood specimens collected. Blood counts were not unusual. No clinical evidence of liver disease. Some persons were reported as having abnormal liver tests, which diminished after relocation
Medical examinations of 112 construction workers from canal remedial construction project	Eight skin-related conditions noted. Similar number with elevated liver function tests before and after canal work
Retrospective study of cancer in former canal residents	In progress; 975 families who lived in the canal area and moved away prior to June 1978 have been identified
Comparison of cancer deaths in Love Canal census tract to other parts of New York State	Cancer registry data show no evidence for higher cancer rates associated with residence near the Love Canal

[a] Derived from listed references and personal communication with current New York State officials.

Table 2 Pregnancy histories of female residents of specified sections of the Love Canal and north of Colvin area

Pregnancy history at present address	Exposed areas				Control areas			
	99th[a]/97th[a] Street		Houses possibly built on contaminated fill		Houses not built on contaminated fill		North of Colvin	
	No.	%	No.	%	No.	%	No.	%
Women ever pregnant	42	100	49	100	98	100	65	100
Number of pregnancies[b]	79	100	108	100	164	100	125	100
(Sets of twins)	(1)		(2)		(2)		(0)	
Women with live births	40	95.2	44	89.8	92	93.9	64	96.9
Number of live births	65	82.2	83	76.9	144	87.8	110	88.0
Women with miscarriages	10	23.8	16	32.7	16	16.3	11	16.7
Number of miscarriages	15	19.0	25	23.1	21	12.8	11	8.8
Women with birth defect child	4	9.5	7	14.3	7	7.1	7	10.6
Number of birth defect children	4	5.1	10	9.3	7	4.3	8	6.4
Women with low birth weight child[c]	1	2.5	10	22.7	6	6.5	3	4.7
Number of low birth weight children[c]	1	1.5	13	12.0	11	7.6	3	2.7
Women with stillbirths	0	0	2	4.1	1	1.0	3	4.5
Number of stillbirths	0	0	2	1.9	1	0.6	4	3.2
Women with no unfavourable[d] event	31	73.8	22	44.9	70	71.4	45	68.2
Women with an unfavourable[d] event	11	26.2	27	55.1	28	28.6	21	31.8

From studies of Vianna, N. J., *et al.* reported in Tarlton and Cassidy (1981). Age distributions of women in exposed and control areas were similar.

[a] Houses on the Canal.

[b] Number of pregnancies = (live births + miscarriages + stillbirths) minus twins.

[c] Percentages based on live births.

[d] Unfavourable event: miscarriage, birth defect child, stillbirth or low birth weight child.

(Kolata, 1980) was commissioned by the federal Environmental Protection Agency. The results of this study could not be confirmed by the US Center for Disease Control (1983).

Kalter and Warkany (1983a, b) recently reviewed research on the aetiology of congenital malformations. They list 16 contaminants and additives (including Love Canal pollutants) and nine occupational exposures as being suspected of causing human malformation, but state that, 'Of these substances only mercury is today accepted to be a proven human teratogen.' Studies of selected industrial situations where more precisely defined exposures have occurred may better allow establishment of epidemiological relationships to birth defects. Ascertainment of bias in identifying spontaneous abortions is also a problem. Spontaneous abortions are more readily ascertained when they occur after the first few weeks of pregnancy; consequently, if environmental chemicals affected early abortions while maternal factors were more associated with those later in pregnancy, those chemically induced would be less likely to be detected.

3 IMPLICATIONS OF THE RESEARCH ISSUES

In June 1980, New York State Governor Hugh Carey convened a panel to review the medical and scientific data related to the Love Canal studies and to comment on these data in relation to the development of public policy (Thomas *et al.*, 1980). The panel's report and the recent US National Academy of Sciences' National Research Council report on 'Risk Assessment in the Federal Government: Managing the Process' (NAS, 1983) provide useful suggestions about policy as related to these types of study. It is obvious from both reports that this type of study requires an extensive effort by well-qualified scientists.

For policy making, it is helpful to have a clear conceptual distinction between risk assessment and risk management (NAS, 1983). Episodes of great importance or concern such as the Love Canal benefit from a clear administrative mechanism to coordinate the research of various agencies. Such a mechanism might include a scientific panel to ensure rigorous peer review of study designs and procedures prior to beginning the major efforts, and ongoing monitoring of the research and its interpretive value. In many situations, especially those involving an aroused public, it is helpful to have some independence of the scientists from the policy makers. This may mean a substantial funding commitment if the episode is to be studied effectively, but in the long run it is likely to reduce both the cost and the time required for a thorough, credible assessment of the situation.

Studies of clusters of environmental episodes may be complex and expensive. Therefore, consideration should be given to what is the minimum information needed for policy decisions, and what is the optimum needed for a high-quality, scientific study. For the former, *a priori* attention should be given to how the different possible results of limited data collection would be interpreted. For the latter, the best advice may be: 'Do it right, or don't do it at all.'

For some episodes of environmental contamination, especially where the exposure level is low or uncertain and the number of people involved is small, the likelihood of an interpretable research result may be low. Sometimes it may be better to focus the detailed research efforts on more pure occupational situations and use this information, together with minimal field data, in considering policy decisions.

4 REFERENCES

Balogh, M., Kahn, H., and Medalie, J. H. (1971). Random repeat 24-hour dietary recalls. *Am. J. Clin. Nutr.*, **24,** 304–310.

Block, G. (1982). A review of validations of dietary assessment methods. *Am. J. Epidemiol.*, **115,** 492–505.

Block, G., Hartman, A. M., Dresser, C. M., *et al.* (1986). A data-based approach to diet questionnaire design and testing. *Am. J. Epidemiol.*, **124,** 453–469.

Cole, P., and Goldman, M. B. (1975). Persons at high risk of cancer. In Fraumeni, J. F. (Ed.) *Occupation*, Chapter 11. Academic Press, New York.

Correa, P., Bolanos, O., Garcia, F. T., Gordillo, G., Duque, E., and Cuello, C. (1975). The cancer registry of Cali, Colombia—epidemiologic studies of gastric cancer. Recent results. *Cancer Res.*, **50,** 155–169.

Correa, P., Cuello, C., Duque, E., Burbano, L. C., Garcia, F. T., Bolanos, O., Brown, G., and Haenszel, W. (1976). Gastric cancer in Colombia. III. Natural history of precursor lesions. *J. Natl. Cancer Inst.*, **57,** 1027–1035.

Creech, J. L., and Johnson, M. N. (1974). Angiosarcoma of liver in the manufacture of polyvinyl chloride. *J. Occup. Med.*, **16,** 150–151.

Cuello, C., Correa, P., Haenszel, W., Gordillo, G., Brown, C., Archer, M., and Tannenbaum, S. (1976). Gastric cancer in Colombia. I. Cancer risk and suspect environmental agents. *J. Natl. Cancer Inst.*, **57,** 1015–1020.

Doll, R., and Peto, R. (1981). The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J. Natl. Cancer Inst.*, **66,** 1193–1308.

Gersovitz, M., Madden, J. P., and Smiciklas-Wright, H. (1978). Validity of the 24-hour dietary recall and seven-day record for group comparisons. *J. Am. Diet. Assoc.*, **73,** 48–55.

Goldwater, L. J., Rosso, A. J., and Kleinfield, M. (1965). Bladder tumours in a coal tar dye plant. *Arch. Environ. Health*, **11,** 814–817.

Greenwald, P., and Lawrence, C. E. (1979). *The State Use of Records for Environmental Research. The Public Health Conference on Records and Statistics. The People's Health: Facts, Figures, and the Future*. DHEW Publication No. (PHS) 79-1214, pp. 296–299.

Greenwald, P., Barlow, J. J., Nasca, P. C., and Burnett, W. S. (1971). Vaginal cancer after maternal treatment with synthetic estrogens. *N. Engl. J. Med.*, **285,** 390–392.

Greenwald, P., Rose, J. S., and Daitch, P. B. (1979). Acquaintance networks among leukemia and lymphoma patients. *Am. J. Epidemiol.*, **110,** 162–177.

Greenwald, P., Friedlander, B. R., Lawrence, C. E., Hearne, T., and Earle, K. (1981). Diagnostic sensitivity bias—an epidemiologic explanation for an apparent brain tumour excess. *J. Occup. Med.*, **23,** 690–694.

Haenszel, W., Correa, P., Cuello, C., Guzman, N., Burbano, L. C., Lores, H., and Munoz, J. (1976). Gastric cancer in Colombia. II. Case–control epidemiologic study of precursor lesions. *J. Natl. Cancer Inst.*, **57,** 1021–1026.

Hankin, J. H., Rhoads, G. G., and Glober, G. A. (1975). A dietary method for an epidemiologic study of gastrointestinal cancer. *Am. J. Clin. Nutr.*, **28,** 1055–1060.

Heath, C. W., Jr. (1983). Field epidemiologic studies of populations exposed to waste dumps. *Environ. Health Perspect.*, **48,** 3–7.

Herbst, A. L., Ulfelder, H., and Poskanzer, D. C. (1971). Adenocarcinoma of the vagina: association of maternal stilbestrol therapy with tumour appearance in young women. *N. Engl. J. Med.*, **284,** 878–881.

Janerich, D. T., Stark, A. D., Greenwald, P., Burnett, W. S., Jacobson, H. I., and McCusker, J. (1981a). Increased leukemia, lymphoma, and spontaneous abortion in western New York following a flood disaster. *Public Health Rep.*, **96,** 350–356.

Janerich, D. T., Burnett, W. S., Feck, G., Hoff, M., Nasca, P., Polednak, A. P., Greenwald, P., and Vianna, N. (1981b). Cancer incidence in the Love Canal area. *Science*, **212,** 1404–1407.

Kalter, H., and Warkany, J. (1983a). Congenital malformations: etiologic factors and their role in prevention (first of two parts). *N. Engl. J. Med.*, **308,** 424–429.

Kalter, H., and Warkany, J. (1983b). Congenital malformations: etiologic factors and their role in prevention (second of two parts). *N. Engl. J. Med.*, **308,** 491–497.

Kolata, G. B. (1980). Love Canal: false alarm caused by botched study. *Science*, **208,** 1239–1242.

Knox, G. (1964). Epidemiology of childhood leukemia in North Cumberland and Durham. *Br. J. Prev. Soc. Med.*, **18,** 17–24.

Madden, J. P., Goodman, S. J., and Guthrie, H. A. (1976). Validity of the 24-hour recall. Analysis of data obtained from elderly subjects. *J. Am. Diet. Assoc.*, **68,** 143–147.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res.*, **27,** 209–220.

Nailor, M. G., Tarlton, S., and Cassidy, J. J. (Eds.) (1978). *A Special Report to the Governor and Legislature—Love Canal: Public Health Time Bomb*. New York State Department of Health, Office of Health Communication, Albany, New York.

NAS (1981). *Committee on Nitrite and Alternative Curing Agents in Food. The Health Effects of Nitrate, Nitrite, and Nitroso Compounds*, Part 1 of a 2-part study. National Academy Press, Washington, DC.

NAS (1982). *Committee on Diet, Nutrition and Cancer. Diet, Nutrition, and Cancer*. National Academy Press, Washington, DC.

NAS (1983). *Committee on the Institutional Means for Assessment of Risks to Public Health. Risk Assessment in the Federal Government: Managing the Process*. National Academy Press, Washington, DC.

Surgeon General's Report (1976). *The Health Consequences of Smoking, a Reference Edition—1964–1975*. DHEW Publication No. (CDC) 78-8357.

Tarlton, F., and Cassidy, J. J. (Eds.) (1981). *A Special Report to the Governor and Legislature—Love Canal*. New York State Department of Health, Albany, New York.

Thomas, L., *et al.* (1980). *Report of the Governor's Panel to Review Scientific Studies and the Development of Public Policy on Problems Resulting from Hazardous Wastes*. New York State Department of Health, Albany, New York (unpublished).

US Center for Disease Control (1983). *Cytogenetic Patterns in Persons Living Near Love Canal—New York*. Morbidity and Mortality Weekly Report MMWR 32/261-62.

Warburton, D., and Fraser, F. (1964). Spontaneous abortion risks in man: data from reproductive histories collected in a medical genetics unit. *Hum. Genet.*, **16,** 8.

Willett, W. C., Sampson, L., Stampfer, M. J., Rosner, B., Bain, C., Witschi, J., Hennekens, C. H., and Speizer, F. E. (1985). Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am. J. Epidemiol.*, **122,** 51–65.