

# POSSIBLE MANIFESTATIONS OF WORSENING ENVIRONMENTAL POLLUTION

WILLIAM R. GAFFEY  
CALIFORNIA STATE DEPARTMENT  
OF PUBLIC HEALTH, BERKELEY

## 1. Introduction

On the assumption that many mathematical statisticians are not well acquainted with the kinds of data available to describe health, I propose to review those aspects of health which may be affected by the environment, and to make some comments on the measurement of the environment, as well as the statistical, or rather metastatistical, problems involved in establishing an association between environment and health.

## 2. Definition and measurement of environment

It is possible (and useful for some purposes) to define the "environment" of a human being as everything outside of his epidermis. Such a view would define the smoking habits of parents, for example, as part of their children's environment. It would also define the presence of an efficient ambulance service as part of the environment of the population of a city. In both these cases, the life expectancy of an individual may be effected by the factors mentioned, thus both are examples of environmental characteristics which affect health.

For our purposes, however, it is clearly expedient to take a narrower view, and to define "environment" as the sum total of the physical phenomena which an individual encounters: food, water, air, and other substances with which he comes in contact.

Those aspects of this environment which are thought to be harmful, or potentially harmful, are now often monitored more or less systematically. However, several qualifying comments must be made about the nature of the monitoring process.

First, the purpose of monitoring is usually to implement public policy regarding maximum levels of pollution. As a result, monitoring is often designed not to estimate the average level of pollution, but to detect violations of some maximum permissible level. It therefore tends to take the form of what might be called "suspicion" sampling. For example, a large proportion of our data on

DDT contamination of the environment comes from sampling of critical incidents, such as an overturned DDT tanktruck, or a suspiciously large die-off of birds in an area previously treated with the pesticide. The necessity for such sampling to implement safety regulations is clear, but the results are not terribly useful in estimating population exposures.

Second, the environmental characteristics which are monitored are those which are already known or suspected to be harmful, or those which public opinion has identified as an aesthetic hazard. Therefore, if one uses existing data to assess the effect of environment on health, one is limited to verifying the hypotheses which justified the establishment of the particular monitoring systems, or hypotheses about further health effects of those same pollutants. For example, one can investigate the effects of photochemical air pollution on cardiovascular disease because such pollution already causes subjective discomfort, and is regularly monitored in several areas for that reason.

Third, and somewhat paradoxically, the measurement of environmental effects may be less of a problem than the measurement of health effects in many situations. The reason is that suspected environmental pollution usually comes from an identifiable source, such as a nuclear power plant or a sewage outfall pipe. Therefore, if untoward health effects are found to occur near such sources, we are entitled to suspect environmental contamination even if we cannot immediately define and measure the precise pollutant involved.

The point here is that the possible spectrum of environmental pollutants is so large that most studies of the effect of particular environmental pollution will probably have to measure that pollution on an *ad hoc* basis. This will have to be done either by monitoring for specific pollutants, or by verifying the existence of probable pollution sources and accepting this as *prima facie* evidence of pollution.

### 3. Definition of health

The dependent variable in a study of the effects of environmental pollution is called for convenience "health," but when we get down to cases, we turn out in fact to be measuring lack of health, since this seems to be easier to talk about in specific terms.

In times past, when health hazards manifested themselves as severe communicable diseases or natural catastrophes, it was natural and easy to measure ill health simply by mortality, or possibly by the occurrence of illness, where by illness is meant a communicable disease diagnosed by a physician. It is apparent that these measures have become increasingly inadequate indices of ill health [2]. Recently we have attempted to measure disability and symptoms, and even more recently the presence or absence of factors such as immunization and medical services, which presumably affect the risk of future illness or death.

One is tempted to look for a single summary index which would adequately express these factors in one descriptive number, and indeed such indices have

been devised for combining such aspects as occurrence and duration of disease, and mortality, [5], [6]. But the construction of such an index, however useful it would be for public information and certain aspects of health administration, might conceal phenomena of interest to the scientist. Consider the problems of such an index for mortality. It is clear that we can measure mortality by the number of deaths per 1000 population per year. If we wish to compare two populations whose age, sex and racial compositions are different, this crude mortality rate can be corrected in one of several ways to compensate for these differences. Using such a measure for California, we would find that the risk of death had decreased over the last several decades. However, the overall rate would conceal the fact that the decrease had been greatest in childhood and young adults, and least in older people. More important, both to the scientist and the planner of health programs, the causes of death have changed.

TABLE I

THE FIVE LEADING CAUSES OF DEATH BY SELECTED AGE GROUPS, CALIFORNIA, 1950 AND 1968

Source: State of California, Department of Public Health, Bureau of Vital Statistics.

Cause	Age group							
	1-4 Rank		15-24 Rank		34-44 Rank		55-64 Rank	
	1950	1968	1950	1968	1950	1968	1950	1968
Accidents, including motor vehicle	1	1	2	2	3	5		
Motor vehicle accidents	2	2	1	1	4			
Cancer	3	3	4	5	2	1	2	2
Pneumonia and influenza	4	5						
Suicides			5	3		3		
Homicides				4				
Tuberculosis			3		5		4	
Diseases of the heart					1	2	1	1
Cirrhosis of the liver						4	5	4
Stroke							3	3
Respiratory diseases commonly designated as obstructive								5

Table I shows the five leading causes of death in 1950 and 1968 in California for four large age groups. No changes have occurred in the group under age 5. However, in the 15-24 age group we see the disappearance of tuberculosis, the decline of cancer, the appearance of homicide, and the increase of suicide. In the older age groups we see the disappearance of tuberculosis, the appearance of cirrhosis of the liver, and the appearance of emphysema and similar respiratory diseases. One could say that the crude mortality rate has concealed the fact that behavioral disorders are becoming a more important cause of death, to

the extent that alcoholism and smoking can be blamed for cirrhosis and emphysema.

If we broaden our consideration to other measures of ill health, the problem becomes more complex. Consider Table II, which lists the principal criteria for physical ill health which have been proposed in recent years [1], excluding certain measures of social function which seem to involve mental health.

TABLE II  
POSSIBLE MEASURES OF REAL OR POTENTIAL ILL HEALTH

---



---

1. Mortality
2. Illness diagnosed by a physician
3. Disability
4. Other symptoms
5. Utilization of health services and facilities
6. Presence of metabolic or physiological abnormalities without illness, disability or symptoms
7. Lack of good health practices (proper immunization, and so on)
8. Absence of adequate health services and facilities
9. Presence of potentially threatening environmental conditions

---

Clearly, the content of any index calculated from these measures will depend on the use to which it is put. Health authorities would probably wish to consider all nine elements in planning programs to improve health. Those of us concerned with the health effects of environmental pollution will want to consider at most the first six, since the rest are clearly measures of education, public policy, economic status, or (in the case of item 9) redundant.

It seems clear that a descriptive vector, rather than a descriptive scalar number, is required to describe ill health.

#### 4. Measurement of health

There are three possible mechanisms for recording or estimating the occurrence of the events in question: (1) compulsory legally sanctioned recording of all events in a given category as soon as they occur, or within some limited time thereafter; (2) sampling those persons (physicians) and/or facilities (hospitals) which provide medical care to a population; (3) sampling the population concerned.

Note that, by definition, some of these techniques are unusable for some measures of ill health. For example, mortality cannot be estimated from a population sample, and symptoms and disability cannot be estimated in their entirety from a physician sample.

More important, perhaps, compulsory registration and sampling of physicians and hospitals cannot in themselves provide measures of risk of ill health, since they do not count those persons to whom the events in question have not occurred. With these techniques, measures of the population at risk must be obtained for the place and time in question, which is frequently a difficult task,

since definite population counts will in general be available once every decade through the U.S. Census. Estimates of small area populations in California are made annually by the State, but usually not for areas smaller than a county.

To the best of my knowledge, there is currently no regular sampling of U.S. physicians at this time, although this is being discussed both in California and at the Federal level. There are several systems in existence for obtaining data on hospital patients, but they do not represent either a probability sample or a complete census, so that the problem of determining the population at risk is not easily solvable. For all practical purposes, therefore, usable existing data are limited to those obtained by compulsory registration and population sampling.

Let us now consider the first six variables in Table II, and what is now being done to measure them.

4.1. *Mortality.* In all states, all deaths are recorded and tabulated by age, race, sex, cause and residence. The latter could theoretically be done for any area, no matter how small, but in fact deaths are tabulated by county and sometimes for large cities. Obtaining risk of death for smaller areas is difficult and unreliable because of the problem of estimating population. The exception is infant mortality, where the population at risk is the number of infants born, which is determined by a registration process similar to that for deaths. Tabulations of infant mortality also tend to be made on a county and large city basis, but most health jurisdictions are capable of producing rates for smaller areas if funds can be obtained to pay for the tabulation.

4.2. *Diagnosed illness.* Certain diseases must be reported to the authorities when diagnosed by a physician. The precise list of diseases varies from state to state, but they are generally the classical contagious diseases. Two things can be said of these reporting systems. They are incomplete, often grossly so, and the diseases are not usually the ones we think of in connection with environmental pollution.

A second source of data on diagnosed disease is the National Health Survey [3], which gives patients' reports on their own diagnosed illnesses. This is supplemented by the Health Examination Survey [4], which actually examines a small subsample of the NHS sample. Unfortunately, the NHS data come from a nationwide sample and give little detail for local geographic areas.

A third class of data for a rather specific class of diagnoses comes from birth certificates, which must record congenital malformations. As with all birth data, this information could be provided, given the resources, for very small areas. However, the completeness of reporting of congenital abnormalities is somewhat suspect, since minor abnormalities may escape notice, and others may not develop until some time after birth.

4.3. *Disability, symptoms, utilization of health resources.* Here again the only regular source of data is the National Health Survey, with the limitations noted above for small area calculations. Also, the particular variable of utilization, as measured by days in hospital, number of doctor visits, and so on, is not useful

in most cases, since it is not clear that a low rate does not simply reflect the unavailability of resources, rather than a state of good health.

4.4. *Presence of abnormalities with no other symptoms.* The only data available in this area come from specific research studies of particular environmental pollutants. For example, groups exposed to certain pesticides may experience abnormal cholinesterase levels with no other symptoms, so that the abnormality itself is taken as evidence of exposure. However, there is no systematic surveillance of the vast range of possible abnormalities which might occur.

In summary, it seems clear that no regular health data system is usable in environmental studies. Either existing mortality data must be recalculated on a small area basis corresponding to the area of study, or an ad hoc data system set up, possibly involving a sample survey of the local population.

## 5. Relating health to environment

There are two ways, generally speaking, to establish that health is affected by environmental deterioration. The first is to observe health systematically under different environmental conditions; the second is to make health inferences on the basis of laboratory or animal experiments.

In the first case comparisons can be made either over time in one area, or between areas which differ in environmental pollutants. A practical problem will be the possibility that there is no immediate acute effect, but that cumulative exposure causes some chronic condition. In such a case, any examination of residents of a given area will have to involve some estimation of the length of exposure, and will be complicated by such situations as the individual who lives in a polluted area, but travels some distance to work in a nonpolluted area.

There is no need to discuss the specific statistical techniques which may be used, but I do wish to point out that we are here applying classical inference to a nonexperimental situation, since "treatment" is not randomized. Therefore our statistical decision making, which in the classic situation decides between chance and the experimental variable as an explanation of observed differences, here can only tell us that chance is or is not an acceptable explanation. If it is not, then the difference between say, health in two areas is either due to the environmental difference or to some other difference, associated with environment, which we were not clever enough to detect.

Not much attention has been paid to untangling the latter problem, most likely because it is ill defined and makes us uneasy. The only systematic consideration of the problem of which I am aware is [7], in which Yerushalmy suggests the following rule of thumb: if a suspected "cause" is associated with a very large number of "effects," one should be suspicious of the reality of any of the cause-effect relationships. This is perhaps nothing more than a guide to one's intuition, but alternative suggestions do not appear to be forthcoming.

Laboratory experiments present no such philosophical problem of assessing effect, but they create equally knotty problems of generalization to human

beings. A recent example is the relationship between DDT and liver cancer. Animal experiments indicate that massive doses of DDT will produce cancer-like liver tumors in rats and mice, but in no other laboratory animal. The problem of generalization consists then in deciding whether the human metabolism is closer to that of rats and mice than to, say, dogs.

Obviously there are many situations, especially those involving radiation, where we cannot conduct human experiments. Nevertheless the problem of generalization is there, and the generalizations we make are clearly going to be intuitive, dependent on the consequences of error as we see them, and hopefully, guided somewhat by rational statistical analysis.

In summary, it appears that (1) the routine measures of health as they are presently obtained cannot, generally speaking, be used to assess the impact of environmental pollutants; (2) routinely collected measures of environmental quality will not, as a rule, provide needed data except for a narrow range of investigations; (3) although prudence and necessity lead us to evaluate pollution through animal experiments, the generalization to man is a somewhat intuitive and nonstatistical procedure; and (4) even with all precautions, the results of systematic observations on a population must be further considered in the light of the fact that we do not operate in an experimental situation, and that little attention has been paid to plausible inference in such a situation.

#### REFERENCES

- [1] S. B. GOLDSMITH, "Status of health status indicators," *New Eng. J. Med.*, in press.
- [2] I. M. MORIYAMA, "Problems in the measurement of health status," *Indicators of Social Change*, New York, Russell Sage Foundation, 1968.
- [3] NATIONAL CENTER FOR HEALTH STATISTICS, "Origin, program and operation of the U.S. National Health Survey," *Vital and Health Statistics* (PHS Pub. No. 1000—Series 1—No. 1), Washington, U.S. Gov't. Printing Office, August 1963.
- [4] ———, "Plan and initial program of the Health Examination Survey," *Vital and Health Statistics* (PHS Pub. No. 1000—Series 1—No. 4), Washington, U.S. Gov't. Printing Office, July 1965.
- [5] ———, "An index of health, mathematical models," by C. L. Chiang, *Vital and Health Statistics* (PHS Pub. No. 1000—Series 2—No. 5), Washington, U.S. Gov't. Printing Office, May 1965.
- [6] ———, "Conceptual problems in developing an index for health," by D. F. Sullivan, *Vital and Health Statistics* (PHS Pub. No. 1000—Series 2—No. 17), U.S. Gov't. Printing Office, May 1966.
- [7] J. YERUSHALMY, "On inferring causality from observed associations," *Controversy in Internal Medicine*, W. B. Saunders Co., 1966.

#### Discussion

R. J. Hickey, *Institute for Environmental Studies, University of Pennsylvania, Philadelphia*

I am uneasy about your "rule of thumb" comment to the effect that if a "cause" has "too many effects," the system, or "cause" (hypothesized) is "not real."

If one considers what is known about the biological effects of ionizing radiation, an effect on experimental animals (using relatively high dosages compared with background radiation levels) has been reported to be a shortening of survival, based on population studies. Perhaps the effect might be referred to as an increased rate of biological senescence, possibly based on cumulative somatic genetic degeneration. But if the experimental animals subjected to irradiation "age" more rapidly than the controls, an effect of such more rapid aging can be earlier occurrence of diseases of aging, which are many. Thus, the ionizing radiation "cause" presumably has many effects.

Comparably, if one examines the hypothesis that among nonradioactive atmospheric (and other environmental) chemicals are some which are radiomimetic, then one might expect certain effects in populations somewhat comparable to the effects of ionizing radiation, for example, positively correlated relationships between environmental concentration of a suspected chemical with risk in populations to mortality from certain diseases of senescence [1], [2]. Thus, such hypothesized "causes" might, perhaps, be expected to have many effects.

Regarding a "single index" of manifestations of worsening pollution, one might consider life expectancy. This, however, may not be easily estimated in, for example, different metropolitan populations. However, median age is reported by government agencies. But median age is determined in populations by birth, death, and migration. Median age has been "predicted" statistically from atmospheric concentrations of certain chemicals [1].

#### REFERENCES

- [1] R. J. HICKEY, D. E. BOYCE, E. B. HARNER, and R. C. CLELLAND, "Ecological statistical studies concerning environmental pollution and chronic disease," *IEEE Trans. Geosci. Electron. GE-8* (1970), pp. 186-202.
- [2] R. J. HICKEY, "Air pollution," Chapter 9, *Environment: Resources, Pollution, and Society*, Stanford, Sinauer Associates, Inc., 1971.

*John R. Goldsmith, Environmental Epidemiology, California Department of Public Health*

Professor Hickey has both emphasized the multiplicity of radiation reactions, and used the term "radiomimetic" to describe pollution effects. These notions are in conflict, since they obscure the specific effects of radiation on the one hand and of pollution on the other. As a basis for raising questions, the notion of radiomimeticity has merit, but as a basis for describing what is known, I feel it to be confusing.

Professor Gaffey has not mentioned the interrelationship of experimental results and epidemiological ones, but both are a basis for deciding what effect to study and how to interpret results on epidemiological studies. Experimental results are of crucial importance. This meeting offers an opportunity to statisticians, who have been extensively involved in experimental design and analysis,

to help encourage the further interaction of the experiential and experimental branches of science.

*Samuel W. Greenhouse, National Institute of Mental Health*

You have raised very serious questions concerning the populations one studies and the thoroughness of the information elicited from each subject. Some discussants have raised questions about the need for studying interactions of different agents such as drugs, air pollutants and radiation, and so forth—questions which are of considerable importance to epidemiologists and biostatisticians. I wonder if you would react to the following thought which I have proposed at several meetings in the past, namely, borrowing the “panel” concept from the Census Bureau, and establishing in effect population laboratories which might be counties of size 100,000 to 200,000 residents to be observed carefully with respect to health, and to social behavior. In this body, populations could be well defined and information very detailed on each resident. The idea is obviously not new. The Johns Hopkins use of the Baltimore Eastern Health District is certainly in line with this suggestion.

*Reply: W. Gaffey*

I think Dr. Greenhouse's suggestion is an excellent one.

Up until now, our attempts to evaluate the kinds of relationships which Dr. Greenhouse mentioned have usually involved one of two strategies. The first is to make very simple measurements on extremely large groups of people, as typified by the usual calculations of mortality rates by residence, occupation, and so forth. The second is to make extremely careful and detailed measurements on small groups of people. An example is the intensive scrutiny of a wide range of physiological parameters which is often carried out on workers in pesticide plants.

The kinds of relationships in which we are interested fall between the cracks, so to speak. On the one hand, although the health effects in question concern the population at large, they are likely to manifest themselves in forms which are not measured by the routine data collection systems. On the other hand, the untoward events for which we are looking are likely to be rare, so that no observation of a small group, however carefully done, will turn up anything.

I see Dr. Greenhouse's suggestion as a feasible compromise, which recognizes that we must look, admittedly at some cost in money and effort, at a large enough group to find rare events, while avoiding the completely impractical alternative of making intensive measurements on the whole population.