Biological Markers in Environmental Epidemiology: Constraints and Opportunities

PETER H. GANN, DEVRA L. DAVIS and FREDERICA PERERA

1 INTRODUCTION

Biological markers (measurements made from biological media) are currently of great interest to epidemiologists in the field of environmental health. After placing markers within an historical and conceptual framework, this paper assesses the practical advantages and limitations of biological markers with respect to epidemiological methods and discusses criteria for the evaluation of markers. It ends by addressing the ethical and legal, as opposed to technical, constraints that such methods imply for future research. Carcinogenesis and mutagenesis have been the subjects of most previous discussions, but there are no conceptual barriers to the extension of biological marker techniques into non-genotoxic processes, and those are emphasised here. The overall purpose is to stimulate dialogue that will lead toward a common language regarding the definitions and boundaries of biological markers for use in environmental epidemiology.

2 HISTORICAL BACKGROUND

Over 50 years ago, Sir Rudolph Peters and co-workers introduced the concept of the "biochemical lesion" into the field of toxicology (Peters, 1969). They built on the advances made by Archibald Garrod and others in understanding the biochemical nature of inherited and induced "errors of metabolism" (Garrod, 1923). Through elucidation of the mechanism of action of classic toxicants, such as the effect of arsenic compounds on the pyruvate oxidase system, experimental toxicology was able to achieve a breakthrough into a deeper level of the disease process, one in which biochemical changes precede measurable changes in physiology or tissue structure. In the case of the arsenicals, this breakthrough provided important leads for therapeutic intervention, such as 2,3-dimercaptopropanol (British Anti-Lewisite).

The convergence of work in various disciplines and development of new

laboratory techniques has continued to clarify the fate of exogenous chemicals in the body, and their molecular and cellular targets or effects. Improvements in analytical techniques for isolating specific cells and molecules alone have been nothing less than spectacular in the last 20 years. In fact, it appears possible to produce measurements in biological samples that are sometimes more sensitive and more precise than we can profitably exploit. In 1982, Perera and Weinstein introduced the term "molecular cancer epidemiology," and began to explore the potential for incorporating advanced laboratory techniques (such as DNA adduct measurements) into epidemiological research (Perera and Weinstein, 1982). Lower and Karenek (1982) and Higginson (1977) had also discussed this approach in general terms.

The use of markers is certainly not new in medical research; it is merely somewhat new to toxicology and particularly to the epidemiological study of environmental toxicity. Markers have been used, for example, in nutritional studies (e.g., serum or urinary vitamin concentrations and fecal bile salts), cardiovascular studies (e.g., serum lipoprotein patterns), and diabetes research (e.g., hemoglobin A_1C measurements). Measures of concentrations of toxicants or their metabolites in body fluids (traditional biological monitoring) have long been used by epidemiologists to study workplace exposures and, in a few instances (as in exposure to heavy metals and persistent organohalogen compounds), to study community exposures as well.

However, the oldest use of biological markers, which might serve as a useful paradigm, was in the study of infectious disease. The infectious agent is analogous to the chemical exposure. Before the advent of methods to isolate (and, in effect, measure) the infectious agent, the presence of an infectious disease could be determined from the observation of gross effects and of the role of non-biological factors, such as living near to a case or drinking from the same well. When microbiological techniques became available, researchers had a useful marker of exposure if they could grow the organism from body fluids or tissues. The discovery of Australia antigen by Blumberg in the early 1960s was a notable advance, in that it provided a marker that could indicate exposure and infection with a simple blood test (Blumberg *et al.,* 1962). Eventually, the ability to identify antibody formation, an aspect of host response to a toxicant, provided easily obtained markers that yielded information on dose, susceptibility, or early response (three types of markers to be described below). Antibodies were especially valuable to the epidemiologist, because they could be detected in the blood long after the infectious agent had vanished. In time, the association between exposure and response became so well-validated that the occurrence of infection could be inferred from knowledge of antibody status alone.

Today, seroepidemiological techniques are powerful tools for infectious disease epidemiology, as shown vividly by the capabilities of recent hepatitis research and the rapidly evolving efforts to understand and control Acquired

Immune Deficiency Syndrome (AIDS) (London, 1984; Schupbach *et al.*, 1985). Environmental health research is highly unlikely to uncover a toxicant that replicates itself *in vitro*, but it can expect to find predictable responses to toxicant exposure that might accurately reflect dose or early disease and that might be persistent in some cases. In fact, antibodies to carcinogen/DNA adducts are now beginning to be used in epidemiological studies of populations exposed to cigarette smoke, chemotherapeutic agents, and workplace chemicals (Maugh, 1984).

3 DEFINITIONS

Broadly defined for environmental health purposes, biological markers are assays performed on body fluids, cells, or tissues that indicate, in biochemical or cellular terms, the presence and magnitude of toxicants or of toxicant/host interactions (including adverse responses). Some markers also indicate the susceptibility of an individual to adverse responses. In a discussion of general concepts, the term "biological marker" is preferable to "molecular marker" or "biochemical marker," because it subsumes measurements made at the subcellular, cellular, and tissue level. Biological marker data should be distinguished, on the one hand, from non-biological data, such as historical information, and, on the other hand, from diagnoses of overt clinical disease. Furthermore, purely physiological data—such as those obtained by cardiovascular, nerve, or pulmonary function tests—should be excluded, because they are not derived from biological samples.

As discussed by Dr. Lauwerys elsewhere in this symposium, biological markers reflect the discrete biological processes that accompany a toxic agent from the point of exposure in the external environment through interaction with the host to excretion or depot storage and possible recycling. Biological markers provide windows on one or more of these processes in the "black box" of the organism, including absorption, serum protein binding, distribution, metabolism, repair, immune response, molecular alteration, and cell injury. The term "marker" itself connotes an indirect, indicator function—a measured quantity that marks or indicates something else. The "something else" of ultimate interest is often either a biologically effective dose (concentration of the toxicant at its target sites) or an early, perhaps irreversible, event in the pathogenic chain. Hence, most conceivable biological markers, and nearly all other types of data used in environmental health research, can be viewed as surrogate data.

Markers can be generally classified as indicating susceptibility, dose, and response. As will be discussed, there is overlap between these categories, since they actually correspond to a continuum of biological processes. Therefore, categorisation of a particular marker depends on use and context. To focus on the usefulness of markers as research tools for epidemiology,

terms are adopted that are operational and as flexible as possible. The following definitions are modified from Perera and Weinstein and the recent National Institute of Environmental Health Sciences Task Force III report (Perera and Weinstein, 1982; USNIEHS, 1984):

- "Susceptibility marker"—a biologically measurable indicator of a state, unrelated to exposure or response alone, that influences the probability that disease will result from a given exposure.
- "Dose marker"—a biologically measurable indicator of internal (absorbed) dose or of biologically effective dose that can be used as an independent variable in a study of exposure/disease associations.
- "Early response (or effect) marker"—a biologically measurable indicator of the occurrence of an event that is a stage in a disease process or a disease itself; this indicator can be used as a dependent variable in a study of exposure/disease associations.

Susceptibility markers are a subset of risk or host factors. While a risk factor might be defined (according to the Dictionary of Epidemiology, Last, 1983) as an attribute or exposure that is associated with an increased probability of developing a disease and is not necessarily a causal factor, susceptibility markers are biological measures that predate or are otherwise independent of exposure and effect. This avoids confusion with such risk factors as age and sex or with dose and response markers, which might in themselves be associated with increased risk of disease. Susceptibility markers can be genetic and/or acquired and can include defects in defence, repair, or adaptive mechanisms, or decreases in functional reserve capacity. Operationally, the definition implies that susceptibility markers will often be viewed as confounders of the exposure/effect association and will be handled by stratification of populations or multivariate adjustment. The growing field of pharmacogenetics (or ecogenetics), which studies the genetic basis of individual differences in response to drugs, has provided a rich and largely unexplored territory for environmental health scientists (Vesell, 1984).

Dose markers are often very indirect indicators of the true dose at the site of toxic action, because the specific site of action is often not known or that site is not accessible. Even the simplest-appearing dose markers, such as blood lead (Pb) concentrations or other measures of a toxicant itself in body fluids, reflect extensive interaction with the host and are basically static, summary indicators of various dynamic pharmacokinetic processes.

Effect or early response markers have three functions: (1) to indicate that a reaction has occurred at a biologically responsive site (not necessarily the site of highest toxic significance or greatest sensitivity); (2) to indicate the inception of a pathological process; and (3) to indicate that parallel toxicologically significant effects have probably occurred at inaccessible sites. Effect markers are often surrogates for the true initial "biochemical lesion." For studies of relatively inaccessible target areas such as the nervous and

reproductive systems, biologically analogous targets can be helpful. For example, changes in biogenic amines in platelets might be used as analogous markers for similar changes in neurons, and changes in chromosomal structure in circulating lymphocytes might be used to parallel changes in germ cell chromosomes (Lambert *et al.*, 1982). Obviously, the use of such analogues involves some uncertainty in extrapolating to the target cell of interest, in addition to the inherent uncertainty involved in using a marker to indicate a critical dose or response. A special type of effect marker results from the use of "chemical probes," such as reference drugs or provocative agents, which can be used to indicate function of some inaccessible systems *in vivo*. For example, the serum half-life of antipyrine is a well-known indicator of hepatic metabolic functions (Dossing *et al.*, 1983).

We cannot always tell whether a given marker is of the dose or response type. For example, a marker of dose for one end point (e.g., cancer) might be a marker of response for another (e.g., mutation). Erythrocyte protoporphyrin (EP) has been validated as a reflection of dose (at EP levels above 35 μg/dl), but also can be interpreted as an early response to lead in target cells (CDC, 1985). Consequently, protoporphyrin concentration can be used as either an independent (dose) or dependent (response) variable in epidemiological studies, depending on the research question involved. Some markers indicate either response or susceptibility, as exemplified by delta-aminolevulinic acid dehydratase (delta-ALAD). A decrease in the activity of this enzyme has traditionally served as an indicator of early biochemical response to lead. Recently, workers in Germany demonstrated that an inherited deficiency in delta-ALAD places some individuals (including heterozygotes) at increased risk of serious illness following lead exposure (Doss *et al.*, 1984).

4 MARKERS FROM AN EPIDEMIOLOGICAL PERSPECTIVE

The benefits of biological markers vary when viewed from the perspectives of different disciplines. To the toxicologist conducting animal research, markers can yield better information on susceptibility, dose, and early response and thus improve understanding of interspecies extrapolation, dose/response relations at very low exposures, and the critical relation between acute insult and chronic damage. They also afford the opportunity to validate, for example, a potential serum marker for neurological response in humans in terms of pathological abnormalities in the brain itself. To the clinician or preventive medicine specialist, markers promise a means of detecting persons at risk of disease or with early disease and of initiating primary or secondary preventive intervention. To the epidemiologist, markers are tools for improving research methods and constitute a means of

overcoming technical barriers in study design and inference. To the assessor of quantitative risk, or "modeler," whose work integrates data from other disciplines, all of the above information is valuable (Hattis, 1985).

Some of the major technical barriers that concern the epidemiologist are summarised in Table 1. There are serious problems in conducting observational studies of toxicant-induced chronic disease in populations, particularly when the relative risk is low and when the disease is relatively rare and has either multiple separate causes or multiple necessary-but-not-sufficient ones. Markers are not as helpful in the study of acute health effects, in which onset is close to the time of exposure, the relevant short-term exposure can often be characterised by direct measurement, and the onset of the effect is well-defined.

The greatest current challenge to environmental epidemiology is to detect and measure small increases in the risk of relatively common chronic diseases with vague onsets and long latencies. Relative risks of 1.2 to 2.0 for such diseases as chronic obstructive pulmonary disease and coronary artery disease are important to detect, because they can translate into large numbers of cases and relatively large public health problems. In this context, markers have five major benefits for environmental epidemiology:

(1) Increased relative risk and statistical power. Biological markers of response constitute early events in the pathogenesis chain that are generally more common than the ultimate clinical disease. These events might be necessary but not sufficient for disease causation; for example, not every person with chromosomal aberration will go on to develop cancer. By increasing relative risk (i.e., the incidence in the more-exposed compared to the less-exposed population) through use of a more common outcome variable, markers expand a study's statistical power in a classical sense, allowing detection of relationships with fewer subjects.

(2) Shorter follow-up times. Markers that reflect early events allow epidemiologists to reduce the duration of expensive follow-up, and perhaps

Table 1. Characteristics of difficult diseases for environmental epidemiology

Easy	Difficult
Single (or few) causes	Multiple causes
Monofactorial	Multifactorial
Rare disease	Common disease
High relative risk	Low relative risk
Acute disease	Chronic disease
Clear diagnosis	Arbitrary diagnosis
Examples:	Examples:
Angiosarcoma,	Coronary artery disease,
Vaginal adenocarcinoma	Alzheimer's, Hypertension

to use the savings to increase sample size and statistical power. The ability to observe outcomes sooner also means a shorter wait for exposed cohorts to "mature."

(3) Reduced misclassification error. The likelihood of error in assigning subjects to correct exposure categories, when long-term low-level exposures are involved, is particularly high. This problem is compounded in studies that use group measurements to classify subjects, such as in epidemiological studies of air pollution based on data from monitoring stations atop buildings. Unmeasured or unknown factors that alter exposure/dose relations (e.g., pharmacokinetic variations, exercise, and anatomical characteristics) contribute to these errors. Potential errors in classification by outcome, although usually less severe than exposure classification errors, are sometimes important, as in studies of cerebral deficits and lead exposure in children (Needleman and Bellinger, 1984). By indicating dose and response more reliably, biological markers offer a means of reducing misclassification error. Ordinarily (if misclassification itself is indifferent to magnitude of exposure) these non-random errors produce noise in the data, and bias a study toward the null hypothesis. This has the same effect as reducing a study's statistical power by reducing sample size, although conventional definitions of power consider the impact of only random types of error. With markers, the risk of misclassification is related to the degree of error involved in using a marker as an estimator and to the slope of the dose/response curve. If large doses are required to produce a measurable health decrement, dose categories can be broad, and individuals will be properly assigned despite errors in dose estimation. Several recent papers contain more detailed discussion of the distortion of epidemiological results by misclassification (Gladen and Rogan, 1979; Copeland et al., 1977).

(4) Delineation of mixture components and covarying exposures. Biological markers may point out which components of complex toxic mixtures are reaching and altering target sites. Separating the effects of mixture components by using ambient-exposure data is a largely statistical procedure and is restricted by the degree to which specific agents covary from place to place and time to time.

(5) Linking of epidemiology findings to prevention initiatives. Epidemiological studies are often criticised for failing to provide means for timely intervention in affected populations. The further development of biological markers may allow less "body-counting," and create more direct opportunities for prevention.

5 CRITERIA FOR THE EVALUATION OF MARKERS

Some criteria for evaluation of biological markers are basically standard for any new laboratory test, but application to epidemiology involves some unique concerns.

5.1 COST

Epidemiological research usually requires the study of large populations. The minimal number of subjects involved is sometimes set by a natural experiment and sometimes by a critical tension between study design, statistical power, and available funds. The gains in study efficiency offered by markers (decreasing required sample sizes) can be negated by the high cost per assay. The increasing automation of analytical methods, fueled by developments such as microprocessors and flow injection analysis, is a favorable trend.

Case-control studies, which usually require fewer subjects than cohort studies, are used in some instances. One means of reducing the cost of using markers is to study cohorts that have already been appropriately sampled for other purposes. If a nested case-control design is used, the number of assays is reduced further by analyzing only subjects who develop diseases of interest and appropriate controls. Biological specimen banks and the use of innovative study designs are vital in hastening the introduction of markers in epidemiological research when cost per assay might be a barrier.

5.2 INVASIVENESS

Obviously, any procedures used to detect markers in humans must present very low risk, discomfort, and inconvenience to the participants. The benefit to an individual subject might be minimal, in that most will be asymptomatic or non-exposed. Neutron-activation analysis, X-ray fluorescence, and magnetopneumography—which are non-invasive techniques that allow measurement of metals in storage depots or target tissues (e.g., bone, lung, or kidney)—are attractive developments for epidemiology, notwithstanding cost-per-assay considerations. The need for non-invasiveness, however, can be waived in studies designed to use retrospective exposure assessment (e.g., in most case-control studies) to examine severely affected victims (e.g., in the Bhopal incident or the MGM Grand Hotel fire) or to exploit medical procedures performed for problems that are not likely to interfere with the health injuries of interest (Birky *et al.,* 1983). For example, microscopic changes in lung structure can be studied in relation to preoperative lung function in patients undergoing biopsy or in relation to lifetime exposure to air pollution in autopsies of accident victims (Ishikawa *et al.,* 1969; Cosio *et al.,*1978). Sometimes, commonly collected but unused surgical specimens, such as adipose tissue and ovaries, can be used in appropriate study designs. A search for clinical sources of samples can reveal unusual opportunities. For example, samples from patients undergoing *in vitro* fertilisation have been used to study halogenated compounds in a particularly important target site for reproductive toxicity: the follicular fluid surrounding the maturing oocyte (Trapp *et al.,* 1984). Bronchoalveolar lavage (BAL), although invasive, might eventually be performed commonly enough to permit studies

of the relationship between BAL fluid constituents (and/or macrophages) and specific previous exposures to the lungs (Henderson, 1984). The need for greater collaboration among epidemiologists, clinicians, and pathologists is obvious.

5.3 PRECISION (REPRODUCIBILITY)

The variation in measurements performed repeatedly on the same biological sample is attributable to variability in a laboratory, assay method, or specimen itself. This criterion could be a problem for epidemiological applications of markers until reagents and methods are standardised.

5.4 CHARACTERISATION OF INTRAPERSONAL VARIATION

For a marker to be used effectively, the variation in a person's measurements over time will have to be understood. A good biological marker will be relatively impervious (or well-characterised) with respect to time-varying factors in the host, assay, or toxic substance itself, such as diet, other concomitant exposures, biological rhythms, storage and transport conditions, or bioavailability of the toxicant through changes in its physical-chemical matrix. Once some understanding of intrapersonal variation is achieved, sampling strategies can be devised to avoid unnecessary variation and select the appropriate frequency, timing, and numbers of samples. For example, characterisation work on aryl hydrocarbon hydroxylase (AHH), a cytochrome P-450 enzyme that metabolically activates polycyclic aromatic hydrocarbons, revealed that enzyme inducibility in an individual is subject to striking seasonal variation (Paigen et al., 1977). Although the basis for this seasonal effect is unknown, it dictates restriction of sampling to summer months, when AHH inducibility is highest.

5.5 CHARACTERISATION OF INTERPERSONAL VARIATION

Ultimately, in order for a marker to be useful, interpersonal variation (which contains the relevant information value) must be large relative to intrapersonal variation. Interpersonal variation is affected by the same kinds of factors as intrapersonal variation, but the variations attributable to unknown or unmeasurable confounding factors, such as diet and concomitant exposures, might be greater. Once again, AHH—one of the few potential susceptibility markers to be extensively evaluated in populations—provides an example: measurements of AHH activity are confounded by consumption of charbro-iled meats and cruciferous vegetables such as cabbage and broccoli (Paigen et al., 1982). Frequency distributions of a marker in populations subjected to various exposures (including normal persons and non-exposed controls) should be carefully examined. The distribution in these control populations

is preferably narrow and Gaussian. Overlap in these distributions often dictates broad assignment of marker levels (with resulting reduction in the discriminating power of studies) and restriction to the use of group, as opposed to individual, comparisons. For example, although smokers of one pack of cigarettes per day or more had sister chromatid exchange frequencies 20 to 30 percent higher than nonsmoking controls, substantial differences were observed between individuals in that group. A dilemma arises: pilot studies designed to determine the feasibility of a particular marker are elaborate in themselves, requiring fairly large-scale surveys to account adequately for all potential confounding variables. Some sources of variation in a population, such as pharmacokinetic differences, might turn out to be of limited interest to the epidemiologist, if they are very small relative to variation in exposure. Finally, population distributions of markers are needed for clarification of sampling variability and the statistics associated with various sample sizes.

5.6 TEMPORAL SIGNIFICANCE

The most successful epidemiological studies are based on explicit biological models of the dose/response relationship, including as much detail as possible concerning temporal aspects. It is important to know the extent to which a marker reflects relatively recent or past exposure and the extent to which it reflects peaks versus integrated exposures or cumulative versus noncumulative effects. The persistence of a dose or early response indicator depends on the pharmacokinetic properties of the toxicant, the rate of injury repair, and the nature of the media in which the indicator is sampled. Figure 1 shows the persistence of biological markers in various media after a hypothetical single exposure to volatile organic compounds. Breath analysis might provide an accurate indication of the preceding 24-hour dose, but a poor indication of the preceding week's dose. Most chronic disease is related to highly repetitive or nearly continual dosage, which may lead to a steady state in the target tissues, and indicators sensitive to short-term variations in exposure do not reflect that steady state as well as do more persistent indicators. Figure 2 shows how use of a nonpersistent marker to indicate a cumulative or integrated biologically effective dose or early response can lead to severe misclassification of study subjects. In Figure 2, the subjects have the same integrated dose, but measurements taken at times A or B yield over- or underestimates of one subject relative to the other. Three strategies might be used to overcome these difficulties in estimating integrated phenomena:

(1) Use of dose-integrating biological media with low turnover. For example, measurements of hemoglobin adducts might integrate dose over an average of 120 days, the normal lifespan of the red cell. Some lymphocytes might persist much longer, possibly up to 20 to 30 years. Cumulative damage
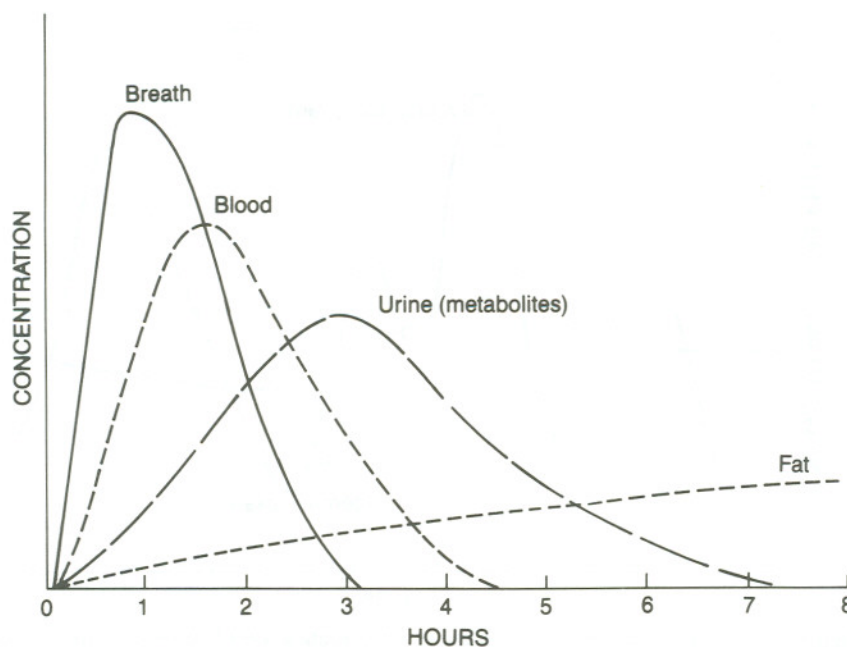
Figure 1. Persistence of volatile organic compounds in various biological media following a single exposure

to genetic material in lymphocytes from long-past exposure to ionizing radiation has been observed (Awa, 1975). The lens, which contains some of the most dense concentrations of protein in the body and which is isolated from outside cells early in gestation, "carries on its metabolism throughout life in a kind of tissue culture" (Weiss, 1983). Samples obtained from cataract surgery or autopsy might provide novel markers of cumulative dose. Hair, nail, and dentine samples have long been used to estimate integrated exposure to heavy metals, but detection of protein interactions with other toxicants in these media have not been explored. In general, progenitor or stem cells, such as the basal epithelial cells in the airway, might provide better opportunities to measure markers than terminal cells with higher turnover.

(2) Repeated sampling or long-term collection. Averaging of multiple samples or long-term collection of media (such as urine or breast milk) is possible, but is often hampered by constraints on study cost and cooperation of subjects.

(3) Retrospective study designs using pathological samples of target tissue. Specially designed studies can exploit biopsy or autopsy samples from the target tissue of interest.
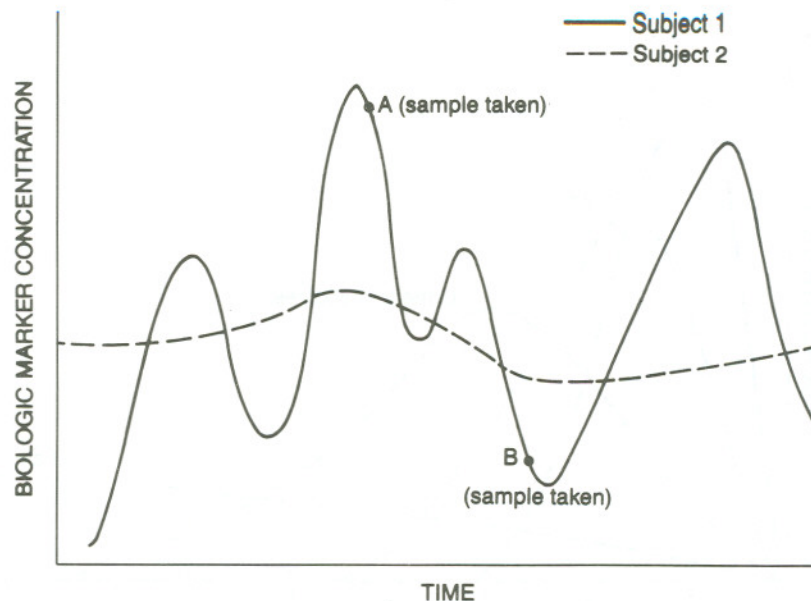
Figure 2. Non-persistent marker leading to misclassification of subjects with equivalent integrated dose

Markers are not likely to be very useful in epidemiological studies until these temporal problems are resolved, since most studies of chronic disease require information on long-past exposure. Epidemiologists should obtain historical information (such as reconstructed workplace histories through job/exposure matrices) in addition to using markers, even when the markers are long-term-dose integrators and the relevant biological model involves cumulative dose and response phenomena. This information would be invaluable in resolving questions about the temporal significance of markers. There is some hope that, as response markers move to points earlier on the pathogenetic chain, the duration of exposure that must be summarised will become shorter and, therefore, less subject to error.

5.7 VALIDITY

Biological markers must be sensitive, i.e., have a high probability of identifying subjects with a given degree of susceptibility, dose, or response. Sensitivity and specificity are important factors in the validity of the relation between a marker and any of the major stages in the exposure/clinical effect continuum, as illustrated in Figure 3. It is important to know the validity of the marker "backwards and forwards;" for example, the relation of urinary hydroxyproline excretion (a possible indicator of pulmonary collagen
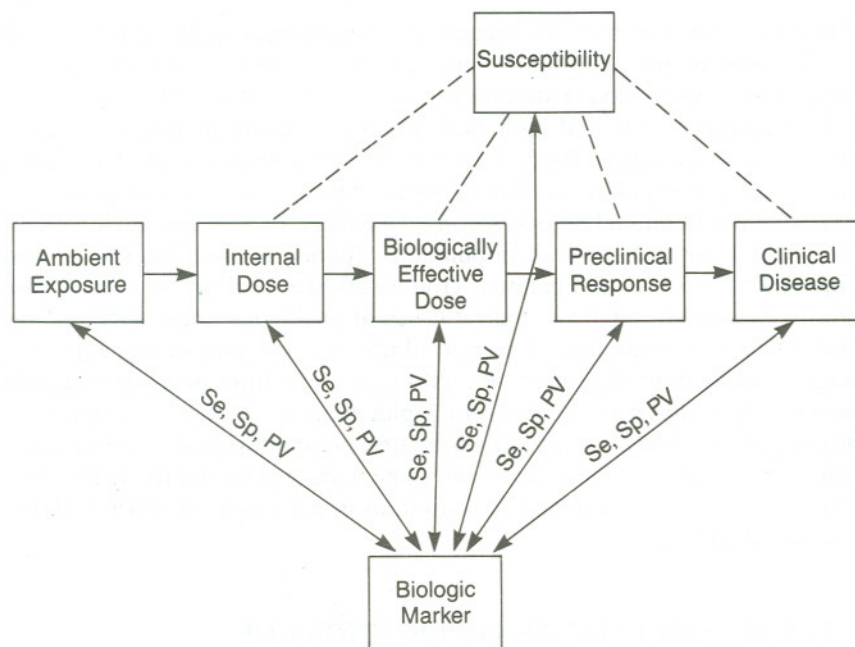
Figure 3. A validation scheme for biological markers

degradation) to both ambient exposure and ultimate signs of emphysema (Matsuki *et al.,* 1981). Such validation, as opposed to calculation of simple correlations, requires comparison of markers to "gold standard" measurements, which might not be available in humans for some of the steps in the process, such as biologically effective dose. Furthermore, these indices of validity involve categorical data; decisions about normal ranges and appropriate categorical levels for continuous data must be made in advance, preferably by examination of extensive frequency distributions for both variables in the validation comparison. In some instances, validation by prospective studies that require waiting for overt disease to develop in large populations will not be feasible. Therefore, validation studies will often be performed in animal models with appropriate extrapolation to humans. Specificity calculations also require large data bases on true-negative results—data bases in which many scientists and funding agencies are reluctant to invest.

Two further points must be raised with regard to validity: predictive value (PV) and low-dose sensitivity. Predictive value—the likelihood that a marker is correct—depends on the prevalence of the event or process estimated by the marker. A given marker that is sensitive and specific might still have a low predictive value for a positive test, PV(+), if the "gold standard" result used for comparison is relatively infrequent in the population studied. Low-

dose sensitivity is important, because epidemiological studies must deal with much lower ambient concentrations than most laboratory animal studies. A dose marker must have a detection limit well below ambient doses.

Evaluation of potential biological markers is likely to follow a more or less ordered sequence. Early activities include assessment of short-term *in vitro* assays (perhaps with cell or tissue cultures or tissue explants) and development of animal models for *in vivo* assays. In humans, early attention is directed toward clinical populations (patients) and heavily exposed populations, such as occupational groups, disaster victims, or, in rare cases, groups exposed to pharmacological doses of analogous drugs. DNA adducts and connective tissue degradation products are examples of markers at this stage of evaluation (Kucich *et al.,* 1984). There is little point in proceeding further with these techniques for epidemiology if they are unable to distinguish chemotherapy patients from untreated individuals or stockbrokers from coke-oven workers. Eventually, some markers become ready for studies of the community or general population, as in the case of biological markers associated with lead.

6 LEGAL AND ETHICAL IMPLICATIONS OF BIOLOGICAL MARKER RESEARCH IN HUMANS

Problems in the policy arena might well outstrip those in the technical arena in complexity and must be resolved so that the basic science can proceed. Shostak reports that workers consider health and safety second only to joblessness as a cause of stress and anxiety in the workplace (Shostak, 1969). In light of these concerns, the development of biological markers can be expected to confront serious practical limitations.

Three major ethical and legal issues require clarification so that research on biological markers can be developed effectively: (1) legal implications of preclinical changes that predict probabilities of future disease, including the development of new protocols or regulations concerning informed consent; (2) the need to provide for understandable reporting of findings; and (3) the need to develop protocols for emergency evaluation of accidents and poisonings, which offer unique opportunities for studying markers.

6.1 NEED FOR CHANGES IN LAWS AND REGULATIONS

The development of biological markers will ultimately revolutionise the tort system. Under traditional law, a person is defined as diseased based on relatively gross accepted evidence, such as tumors or other well-characterised stages of disease. With most chronic diseases, individuals must wait until they are severely incapacitated until they seek compensation for injuries associated with exposure to toxic chemicals. Partly as a consequence of the

long latency between exposure and onset of disease, compensation for occupational disease in the United States remains woefully inadequate (OTA, 1983; Trauberman, 1983). The legal system does not foster equity in these matters, but rather insists on legal principles which cannot be applied, such as the determination that a given exposure was the "proximate" cause of a given chronic disease. Techniques to measure internal dose and dose at the target tissue will provide the theoretical basis for persons to seek compensation for exposure alone, where such exposure can be reasonably argued to place them at risk. Trial lawyers pursued a precedent-setting case involving eight families in Woburn, Massachusetts, who drank water contaminated with trichloroethylene. Damages were sought for emotional anguish and increased risk of serious illness, as well as for several leukemia deaths. In this case, evidence has also been presented that the exposed families suffered cellular and immune system damage which may presage leukemias (Davis, 1985).

6.2 THE RIGHT TO KNOW AND THE DUTY TO INFORM

Twin principles of common law are the right to know about adverse exposure and the duty to inform persons about their exposures. Many laws promulgated during the environmental decade implicitly conferred on the public a right to know about hazards in their workplaces and the general environment. For instance, several states (e.g., California) and cities (e.g., Philadelphia) have vigorous right to know laws, which require that the public be informed about their exposure to industrial chemicals. In addition, the Health Services Research, Health Statistics, and Health Care Technology Act of 1978 indicates a Congressional desire to promote notification and intervention programs for individuals found to be at high risk. Further, sections 8d and 8e of the Toxic Substances Control Act (TSCA) stipulate that any person having knowledge of a study showing that a compound might constitute a risk to public health or the environment must report that finding to the federal government, within 30 days.

One well-known illustration bears mention at this point, as an example of what to avoid. In 1980, lawyers for the U.S. Environmental Protection Agency sought an analysis of chromosomal damage among a small number of residents of Love Canal, in preparation for its lawsuit for $124.5 million against the Hooker Chemical and Plastics Corporation, charging it with dumping toxic chemical wastes at Love Canal and at three other sites in the Niagara Falls area. Peer review of the data prior to their release was not conducted. Following release of these data and recommendations for relocation in the Love Canal area, a panel of peer reviewers concluded that interpretation of the data was questionable, given the lack of concurrent controls and the failure to use modern staining techniques. Major political fall-out ensued from this chromosome study. For the residents of Love

Canal, release of the results and the decision to relocate 710 families exacerbated their psychological stress. For the scientists, no further work on this topic was allowed, as the affected individuals refused to cooperate.

Prior to their execution, studies of biological markers should be carefully reviewed with all affected parties, to construct procedures for explaining and releasing findings, first, to those affected, and then to relevant federal, state, and local agencies. To date, the duty to inform under TSCA has been underutilised. However, effective implementation could produce difficult questions for the development of biological markers. Are studies of exposure markers to be considered reportable? Can degrees of intake or dose be generally accepted as indicating a risk to public health or the environment? While case by case applications will be required at the outset, it would be useful for regulators, scientists, and policy analysts to air these questions fully, lest they undermine developments of the science, as was the case with Love Canal.

Development of techniques to communicate effectively about markers will be the key. The line will need to be carefully drawn between biological monitoring of routine exposures to ensure compliance and studies of biological activity for research purposes. Markers must be used to complement environmental measurements in standard-setting and compliance, rather than replace them (Ross, 1984). With known or suspected genotoxic agents, no human studies involving deliberate exposure should be conducted. However, monitoring of persons previously and currently exposed to such agents should be conducted. With such exposures, notification of a cohort of workers at risk requires systematic development of criteria as to what constitutes a notifiable risk and of programs to provide understandable notification (Schulte *et al.*, 1985).

6.3 ACCIDENTS AND POISONING

It is important that researchers develop protocols for sampling appropriate tissue specimens. The failure to establish emergency analytical capabilities will only worsen the impact of future accidental contaminations. Public education programs in that context also must be developed. Such protocols must be diplomatically presented, particularly in developing countries where the likelihood of such incidents can be great due to rapid industrialisation. Appropriate international collaboration is essential, lest efforts appear as scientific *noblesse oblige* on the part of developed countries. Opportunities for emergency evaluations can also be expected to continue in the industrialised world, given the prevalence of hazardous wastes problems, transportation accidents, and fires.

7 CONCLUSIONS

Biological markers offer promising opportunities to improve the feasibility and inferential strength of epidemiological observations by mitigating problems with misclassification bias, power, and length of follow-up. At this early stage, the development of markers for environmental epidemiology will be enhanced by attention to historical antecedents in other fields and adoption of a common language regarding central concepts. The transformation of an assay from laboratory technique to epidemiological research tool is an arduous and expensive process, requiring evaluative studies of cost, invasiveness, and determination of all major sources of variation in groups and individuals. Particular attention must be paid to normative data in human populations, development of animal models for validation, and clarification of the temporal characteristics of potential markers. Two types of collaboration will be especially useful: (1) between laboratory workers and epidemiologists in order to match basic science capabilities to field needs; and (2) among clinicians, pathologists, and epidemiologists to fully exploit the innovative and cost-effective use of specimen banks, clinical samples, and stored biological materials from cohort studies.

Finally, several major ethical and legal issues will have to be addressed to ensure that population studies involving markers will proceed without undue impediments.

REFERENCES

Awa, A.A. (1975). Review of thirty years study of Hiroshima and Nagasaki atomic bomb survivors. II. Biological effects. B. Genetic effects. 2. Cytogenetic effects. *J. Radiat. Res.* **16** (Suppl.), 75–81.

Birky, M., Malek, D. and Paabo, M. (1983). Study of biological samples obtained from victims of MGM Grand Hotel fire. *J. Analyt. Toxicol.* **7**, 265–271.

Blumberg, B.S., Dray, S. and Robinson, J.C. (1962). Antigen polymorphism of a low-density bet-lipoprotein. Allotypy in human serum. *Nature* **194**: 66–658.

CDC (Centers for Disease Control) (1985). *Preventing Lead Poisoning in Young Children: A Statement by the Centers for Disease Control.* U.S. Department of Health and Human Services, Centers for Disease Control, Atlanta, Georgia.

Copeland, K.T., Checkoway, H., McMichael, A.J. and Holbrook, R.H. (1977). Bias due to misclassification in the estimation of relative risk. *Am. J. Epidemiol.* **105**, 488–495.

Cosio, M., Ghezzo, H. Hoggs, J.C. *et al.* (1978). The relations between structural changes in small airways and pulmonary-function tests. *New Engl. J. Med.* **298**, 1277–1311.

Davis, E.S. (1985). Trial Lawyers for Public Justice and EI [ecological illness]. *Ecological Illness Law Report* (Evanston, Illinois, USA) **3**(3), 6.

Doss, M., Laubenthal, F., and Stoeppler, M. (1984). Lead poisoning in inherited - aminolevulinic acid dehydratase deficiency. *Int. Arch. Occup. Environ. Health* **54**, 55–63.

Dossing, M., Baelum, J., and Lundqvist, G.R. (1983). Antipyrine clearance during experimental exposure to toluene. *Br. J. Ind. Med.* **40**, 466–469.

Garrod, A.E. (1923). *Inborn Errors of Metabolism*, 2nd edition. Frowde, Hodder and Stoughton. 5 pp.

Gladen, B. and Rogan, W.J. (1979). Misclassification and the design of environmental studies. *Am. J. Epidemiol.* **109**, 607–616.

Hattis, D.B. (1985). *The Promise of Molecular Epidemiology for Quantitative Risk Assessment*. MIT Center for Policy Alternatives Report No. E40–233. Massachusetts Institute of Technology, Cambridge, MA 02139.

Henderson, R.F. (1984). Use of bronchoalveolar lavage to detect lung damage. *Environ. Health Perspect.* **56**, 115–129.

Higginson, J. (1977). The role of the pathologist in environmental medicine and public health. *Am. J. Pathol.* **86**, 459–484.

Ishikawa, S., Bowden, D.H., Fisher, V. and Wyatt, J.P. (1969). The "emphysema profile" in two midwestern cities in North America. *Arch. Environ. Health* **18**: 660–666.

Kucich, U., Abrams, W.R., Christner, P., Rosenbloom, J., Kimbel, P. and Weinbaum, B. (1984). Molecular weight distribution of elastic peptides in plasmas from human non-smokers, smokers and emphysema patients. *Am. Rev. Respir. Dis.* **129**, 307.

Lambert, B., Hindblad, A., Holmberg, K. and Francesconi, D. (1982). The use of sister chromatid exchange to monitor human populations for exposure to toxicologically harmful agents. In Wolff, S. (Ed.) *Sister Chromatid Exchange,* pp. 149–182. John Wiley and Sons, New York.

Last, John M. (Ed.) (1983). *Dictionary of Epidemiology*. International Epidemiological Association. Oxford University Press, New York. 114 pp.

London, W.T. (1984). Epidemiologic lessons from viral hepatitis B. *Ann. N.Y. Acad. Sci.* **37**, 100–104.

Lower, G.M. and Karenek, M.S. (1982). Risk, susceptibility and the epidemiology of proliferative neoplastic disease: Descriptive vs. mechanistic approaches. *Med. Hypotheses* **9**, 33–49.

Matsuki, J., Osaka, F., Kasuga, H. and Sugita, M. (1981). Epidemiological study on the effects of smoking and air pollution using urinary hydroxyproline: creatinine ratio on the healthy school children and adults. *Jap. J. Public Health* **28**, 505–515.

Maugh, T.H. (1984). Tracking exposure to toxic substances. *Science* **226**, 1183- 1184.

Needleman, H.L. and Bellinger, D. (1984). The developmental consequences of childhood exposure to lead: recent studies and methodological issues. *Adv. Clin. Child Physiol.* **7**, 195–200.

OTA (Office of Technology Assessment) (1983). *The Role of Genetic Testing in the Prevention of Occupational Disease*. U.S. Congress, Office of Technology Assessment. Government Printing Office, Washington, D.C. 243 pages.

Paigen, B., Minowada, J., Gurtoo, J., *et al.* (1977). Distribution of aryl hydrocarbon hydroxylase inducibility in cultured human lymphocytes. *Cancer Res.* **37**, 1829–1837.

Paigen, B., Ward, E., Steenland, K., Bolanowska, W., Gessner, T., *et al.* (1982). Absence of seasonal variation in antipyrine metabolism. *Clin. Pharmacol. Ther.* **31**, 144–150.

Perera, F. and Weinstein, I.B. (1982). Molecular epidemiology and carcinogen-DNA adduct detection: New approaches to studies of human cancer causation. *J. Chron. Dis.* **35**, 581–600.

Peters, R.A. (1969). The biochemical lesion and its historical development. *Br. Med. Bull.* **25**, 223–226.

Ross, R.D. (1984). Biological monitoring: Methodologic progress suggests wider applications. *Am. J. Ind. Med.* **6**, 87–88.

Schulte, P.A., Ringen, K. Altekruse, E.B., *et al.* (1985). Notification of a cohort of workers at risk of bladder cancer. *J. Occup. Med.* **27**, 9–28.

Schupbach, J., Haller, O. Vogt, M., *et al.* (1985). Antibodies to HTLV-III in Swiss patients with AIDS and pre-AIDS and in groups at risk for AIDS. *New Engl. J. Med.* **312**, 265–270.

Shostak, A. (1969). *A Blue Collar Life*. Random House, New York: 299 pages.

Trapp, M., Baukloh, V., Bohnet, H.G., and Heeschen, W. (1984). Pollutants in human follicular fluid. *Fertil. Steril.* **42**, 146–148.

Trauberman, J. (1983). Statutory reform of "toxic torts": Relieving legal, scientific, and economic burdens on the chemical victim. *Harvard Environ. Law Rev.* **7**, 177–296.

USNIEHS (U.S. National Institute for Environmental Health Sciences) (1984). *Task Force III Report*, pp. 197–259. U.S. Government Printing Office, Washington, D.C.

Vesell, E.S. (1984). Symposium: New Directions in Pharmacogenetics. 34th Annual Fall Meeting of the American Society for Pharmacology and Experimental Therapeutics, Philadelphia, PA, American Society for Pharmacology and Experimental Therapeutics. *Fed. Proc.* **43**, 2319–2347.

Weiss, L. (1983). *Cell and Tissue Biology*, 5th edition, pp. 1159–1160. Elsevier Biomedical, New York.