

ASPO Distinguished Achievement Award Lecture

Epidemiological Studies Using Biological Markers: Issues for Epidemiologists¹

Barbara S. Hulka²

Kenan Professor and Chair, Department of Epidemiology, University of North Carolina, Chapel Hill, North Carolina 27599

On April 11, 1991, Barbara Sorenson Hulka, M.D., M.P.H., received the Distinguished Achievement Award of the American Society of Preventive Oncology (ASPO) at the annual meeting of the Society in Seattle, Washington. The award is given by the Executive Committee of ASPO in recognition of contributions to the prevention of cancer over many years, not for a single accomplishment. Any member of ASPO may nominate a candidate for the award. The 15-member Executive Committee votes by secret ballot after considering a summary of the contributions to cancer prevention of the several candidates. Dr. Hulka won in a close election over several other outstanding scientists.

Dr. Hulka has chaired the Department of Epidemiology at the University of North Carolina School of Public Health since 1983. During her tenure she has developed the Epidemiology Department into one of the best in the country. The basis for her award, however, was her 25 years of work in the prevention of cancers of female reproductive organs. Dr. Hulka began her career with a series of critical studies of cervical cancer in underserved populations. These were followed by studies of the role of hormones in endometrial and breast cancers, a field in which she continues to make important contributions. She has brought to these studies a balanced evaluation of the risks and benefits of hormones rarely found in epidemiological research.

Dr. Hulka has been incredibly generous in her service on state and national advisory committees. Since 1973 she has served on 26 national advisory committees, been a member of the editorial boards of seven journals, and been a consultant to numerous public and private organizations.

Dr. Hulka is a graduate of Radcliffe College, the Julliard School of Music, and Columbia University's College of Physicians and Surgeons and received her M.P.H. from Columbia University School of Public Health. It was an honor for ASPO to bestow its Distinguished Achievement Award on Barbara Hulka.

*W. Thomas London
President of ASPO
(1969-1991)*

Biological markers in epidemiological research have a long history. They were used initially in the study of infectious diseases from which came the genesis of epidemiology as an approach to studying health and disease. Over the years, however, the issues in epidemiology have changed; quantitative statistical and epidemiological methods have blossomed and assumed preeminence. During the 1980s when laboratory advancements in molecular and biochemical techniques were rapid, biological elements in epidemiological research again became prominent. There are biological markers applicable to almost all diseases and conditions that we study in epidemiology. How these markers are used and the influence they have on our study designs and methods,

particularly in the area of cancer epidemiology, are the focus of this paper.

I will review several topics, starting with the conceptual framework for characterizing biomarkers and the rationale for their use. Next I will mention several methodological issues pertinent to studies using biomarkers and show how these lead to modifications in traditional epidemiological study designs. A series of studies that fit within the conceptual framework for biomarkers will serve as examples. Finally, I will offer some suggestions about future directions for studies using biomarkers.

Conceptual Framework

Fig. 1 is a modification of a diagram that has appeared in many publications and presentations (1). It represents a sequence of events from a toxic environmental exposure to the outcome of disease. Markers for these events are labeled "internal dose," "biologically effective dose," "biological response," "altered structure/function," "disease," and "prognosis." I have added the concepts of "susceptibility" and "environmental/life style" factors, noting their potential importance as modifiers of the steps along the spectrum from exposure to disease.

Received 6/18/91.

¹This work was funded in part through Cooperative Agreement CR813485 between the Health Effects Research Laboratory of the U.S. Environmental Protection Agency, and the University of North Carolina, Department of Epidemiology. Presented at the 15th Annual Meeting of the American Society of Preventive Oncology in Seattle, Washington on April 11, 1991.

²To whom requests for reprints should be addressed.

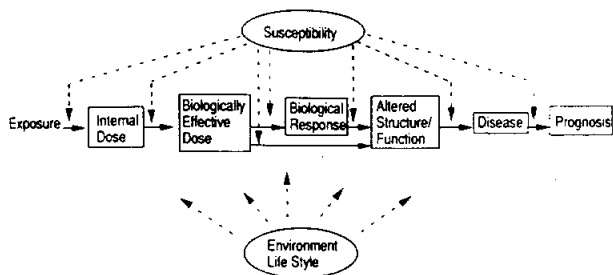


Fig. 1. The relationship of biological markers to exposure and disease.

One drawback of the diagram is the impression that every marker moves inexorably on to the next. For many of the early stages in the marker spectrum, progression is rare. Almost all biological alterations, as measured by a marker, are repaired, stabilized, or destroyed. Repair capability is greatest and progression least frequent on the left side of the diagram.

Rationale

The rationale for the use of biological and biochemical markers in epidemiological research include the following.

(a) Improve accuracy and reduce misclassification in exposure variables. Misclassification of exposure is a common problem in studies of occupational and environmental exposures. Internal dose or biologically effective dose markers should be more accurate reflections of the true dose absorbed by the body than ambient monitors of the environment.

(b) Allow for study of preclinical disease and provide opportunity for prevention. A classical example is cervical cytology, used in screening for precancerous lesions of the cervix. The Pap smear, having been a component of clinical practice for several decades, has resulted in the diagnosis and treatment of preneoplastic disease with a subsequent reduction in invasive cervical cancer incidence and mortality.

(c) Provide more homogeneous classification of disease. Subsets of acute nonlymphocytic leukemias, for example, have been identified with distinctive chromosomal abnormalities. These cases are more frequently related to chemical exposures than cases without these alterations. More accurate risk estimates can be obtained for the effect of chemicals in producing acute nonlymphocytic leukemia by including chromosomal abnormalities in the case definition rather than diluting the association with a heterogeneous case group (2).

(d) Identify interindividual differences in susceptibility to disease and utilize this information in classifying risk groups. Susceptibility markers such as genetic polymorphisms or differences in metabolic activity are likely to act as effect modifiers, or possibly confounders, in studies that evaluate relationships between exposure and disease.

(e) Improve the methodology of clinical trials. Biological markers may take the form of intermediate outcomes, proximate to the occurrence of actual disease, which is the ultimate outcome. Their employment may reduce the sample size or length of follow-up required

in trials. Biological markers may also be used to classify trial participants for prognostic purposes or as part of the eligibility criteria. Biological markers of compliance with intervention strategies used in clinical trials are often essential.

(f) Increase our understanding of disease mechanisms, allowing for stronger biologically based research designs and interpretation of epidemiological data.

(g) Enhance the credibility of traditional epidemiological research, especially for studies of weak associations. Very little current epidemiological research identifies risk estimates of 3- to 4-fold or greater. Much of our research is centered on studies providing risk estimates of 1.5 or less. Innovative epidemiological studies employing biomarkers, relevant to exposure or disease, can elucidate mechanisms for disease occurrence that support the epidemiological studies of weak associations. Relevant findings from biomarker studies can increase the likelihood that the weak associations are real and not just artifacts due to inadequacies of study design, methods, or analyses.

(h) Improve methods in risk assessment or estimation. Biological markers can be used to identify hazards, clarify dose-response relationships, describe the extent of exposure in human populations, and estimate risks (especially for low dose exposure).

(i) Increase scientific knowledge about disease processes and adverse health effects through multidisciplinary research. The strategy of addressing health problems through multiple disciplines, rather than a singular focus, is gaining ascendancy in scientific inquiry. The ability to integrate knowledge from a variety of disciplines about a defined health problem may be our most fruitful avenue for developing a better understanding of disease causation and prevention.

Validation of Biological Markers

The ultimate validation of any marker is its ability to predict disease occurrence. However, preliminary studies with more limited goals for validation are usually necessary before a marker is proposed for a study to demonstrate its predictive validity.

In epidemiology, sensitivity and specificity are the usual measures of validity for a screening test or other procedure that requires evaluation. But the definition of these terms varies across disciplines. Laboratory scientists consider sensitivity to be the minimum level of an analyte that an assay can detect. (More technically, sensitivity is the smallest single value that can be distinguished from zero with 95% confidence limits.) For analyte levels above this minimum, the assay is positive; for lower levels, the assay is negative. Specificity in the laboratory indicates the ability to detect a unique analyte from a group of closely related structures. This issue frequently arises with immunological assays, such as the enzyme-linked immunosorbent assay, in which antibodies are created in animal systems to interact with specific chemical structures in humans. Often these antibodies detect several closely related structures and are not uniquely specific.

Epidemiologists actually have two approaches to considering validity and, therefore, sensitivity and specificity. These approaches relate to the goals of the studies and to whether they focus on disease or exposure. Those

of us who have learned the concept of validity from the standpoint of screening for disease evaluate marker sensitivity and specificity relative to diseased and non-diseased persons. Environmental epidemiologists may refer to the sensitivity and specificity of a biological marker relative to exposed and nonexposed persons. They may validate biological markers of exposure against ambient levels of the chemical or pollutant. However, the justification for this approach is not immediately obvious, since the biological marker should be a more accurate indicator of the impact of the chemical on the human organism than the external measurement; otherwise, why bother with the marker? Validating the biological marker against the external measurement appears to lack conceptual coherence.

What are the strategies for developing and validating markers prior to the ultimate validation of their ability to predict disease occurrence? Preliminary work in animal model systems is necessary for assay development, to generate dose-response curves, and to gain an understanding of whether the marker is truly part of the pathogenesis of the disease or merely an adaptive response on the part of the organism. In the conceptual model of biological markers shown in Fig. 1, the focus of interest is an understanding disease causation and markers that fit into the biological pathway between exposure and disease. In those instances where surrogate markers are used (e.g., hemoglobin adducts versus DNA adducts in target tissues), it is important to determine that the surrogate is an accurate, quantitative reflection of the marker of interest that lies in the disease pathogenesis pathway.

Preliminary studies must also be done on human tissues. In comparison to animal studies, humans will likely have very low exposures to the chemical or complex mixture of interest. Therefore, the sensitivity of the assay (laboratory definition of sensitivity) will be stressed. Assay techniques that were successful in animal systems may require modification for human studies. Human tissues need to be obtained from highly selected, heavily exposed persons. With these tissues, the feasibility, reliability, and laboratory sensitivity and specificity of the assay can be evaluated.

Transitional Epidemiological Studies

Transitional epidemiological studies are observational studies in humans, employing biological markers and bridging the gap between laboratory experimentation and population-based epidemiology. These studies may be designed to evaluate exposures, health effects, or susceptibility.

What are their goals? We expect them to address preliminary or developmental questions in anticipation of large-scale epidemiological studies, or they may answer substantive questions. For example, many of the studies required to validate biomarkers would fall into the category of transitional epidemiological studies. Some might be designed to assess intra- and interindividual variability. Sources and extent of variability assume importance in transitional studies, since expansion of sample size to reduce variability is often not a feasible option. Within-person variability is illustrated by varying levels of DNA adducts across organ sites or sister chromatid exchange frequency among WBC. The variability within a person for these markers may be as great or

greater than the variability among persons. Another goal of these studies might be to evaluate the feasibility of using a marker in field conditions. Information must be obtained about the marker's stability and the conditions required for collection, transport, and storage. In addition, there are situations in which transitional studies are designed to identify exposure/effect relationships.

What are the characteristics of these studies? They may use laboratory assays as exposure factors, health effects, confounders, or effect modifiers. These studies will usually not be population based. Subjects will frequently be selected to meet particular disease or exposure criteria. For example, subjects may be selected to be extreme, such as people who have been very highly exposed or totally unexposed.

Commonly applied design criteria will include subject restriction and matching. We may expect large between-group differences in the outcome variable so that small sample sizes (50 to 100 subjects or even less) may be adequate. There will be more use of continuous variables and transformed variables, new scoring methods, and statistical techniques. Thoughtful use of analytic modeling techniques will be necessary to minimize the number of independent variables in the models.

Background levels of markers may cause misclassification. If, for example, one is interested in DNA adducts formed from polycyclic aromatic hydrocarbons and nitrosamines in tobacco smoke, the ubiquitous presence of these chemicals in the general environment and the resulting adduct formation in the tissues of nonsmokers will make the association between adducts and smoking difficult to detect. Recognition of background levels of the exposure should be part of the study design when considering sample size and dose-response relationships for the biomarker.

The issue of confounding assumes some new dimensions in these studies. Often, little is known about risk factors for the marker (other than the exposure of interest), and even less attention may be given to the distribution of risk factors among exposure groups. The potential for confounding is particularly great when nonspecific markers of genotoxicity are used. In this situation, a variety of exposures may produce the same biological response. For example, metabolic products of dietary constituents can produce adducts with the same chemical structures as environmental toxins (3). If diet and environmental exposure are correlated, confounding may occur.

Specific nutrients may act as independent predictors of a marker, confounders, or effect modifiers. Often there is uncertainty about which dietary factors to study and how to measure them, although both biomarkers of dietary constituents and questionnaire information are generally useful. The need for greater attention to dietary constituents and methods for their measurement is evident from the scientific literature both within and outside of epidemiology.

Other factors may act as confounders or possible effect modifiers. Endogenous exposure to normal body constituents, such as estrogens in women, can influence the body's response to xenobiotics. In the female rat model, estrogens act to induce the P450 isoenzyme system in the liver, increasing metabolic production of electrophilic metabolites of genotoxic xenobiotics with a resultant increase in DNA adduct formation (4). In

women, only the induction of liver enzymes by estrogens has been demonstrated; the subsequent biochemical pathways have not been studied.

An example of possible effect modification was reported by Perera *et al.* (5). They showed differences in DNA adduct levels in normal lung tissue from persons with and without lung cancer. Seasonal differences were also observed; higher adduct levels were found in tissues obtained between July and October as compared to other times of the year.

What do we do to minimize confounding? First, maximize design options to reduce confounding, thus minimizing the need to control confounding in the analysis. With small numbers of subjects, analytic control is unlikely to be reliable or stable. Design options for minimizing confounding include restriction, matching, and exclusion. Limit the distribution of potential confounders across subjects through restriction. Consider matching on demographic characteristics, *i.e.*, age, gender, race, and social class, which will also help to control for unknown confounders. The ultimate form of restriction is exclusion. Application of exclusion criteria in epidemiological studies is standard procedure and perhaps is used excessively when the number of people excluded becomes a significant proportion of the total study population. Overzealous application of exclusion criteria requires the recruitment of more study subjects and could bias the study results.

In transitional epidemiology, the first rule of analysis is to keep it simple. Graphic display of the individual data, followed by descriptive statistics and stratified analyses, is informative. Multivariate models should be used sparingly, and the models should be sparse, with few independent variables. Transformation of continuous variables is often appropriate. Log or square transformations are desirable if there is a need to equalize the variance across a distribution of values and to normalize a skewed distribution (6). Transformed data are often more appropriate for analysis.

Metaanalysis may be an important analytical tool when dealing with multiple small studies, none of which have adequate power to demonstrate statistically significant effects. Although we anticipate large differences in the effect measure for exposed and unexposed persons or cases and controls, transitional epidemiological studies are likely to be small, with modest power. A metaanalysis reported by Margolin and Shelby (7) brought together multiple small studies of sister chromatid exchange distributions in different race and gender groups. The metaanalysis demonstrated that women on average have half a sister chromatid exchange more per cell than men, a finding not previously recognized from the individual studies.

Illustration of a Transitional Epidemiological Study

A transitional exposure study, cross-sectional in design, is currently in progress in the Epidemiology and Biostatistics Departments in the School of Public Health at the University of North Carolina in conjunction with scientists in the Health Effects Research Laboratories of the Environmental Protection Agency (8). The study employs cotinine as a marker for tobacco smoke in the semen, serum, and urine of smokers and nonsmokers. The example is relevant to cancer as well as reproductive epi-

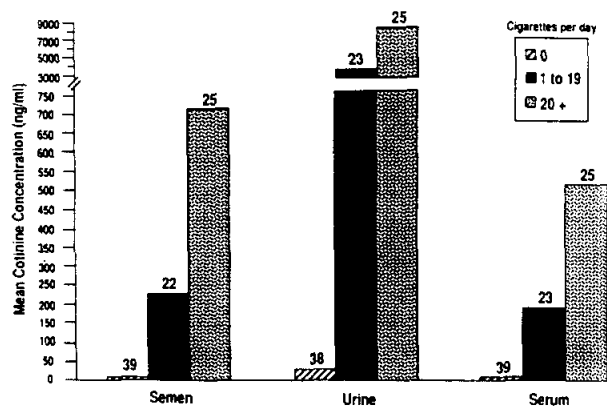


Fig. 2. Cotinine concentrations in semen, urine, and serum by level of active smoking.

demology since the mechanisms of genotoxicity and the agents involved are frequently the same in both fields, although the target cells and organs may differ. Cigarette smoke contains mutagens and carcinogens, which are known to cause adverse reproductive outcomes and cancers of multiple sites. An adverse effect of cigarette smoking on sperm count, motility, and morphology has been suggested in some studies, although statistically significant associations are rare. If cotinine, a metabolite of nicotine, could be detected in the semen of smokers, it would support the concept that mutagenic substances in tobacco smoke could also reach sperm cells and potentially lead to heritable damage. Cigarette smoke also provides a model system for other environmental pollutants that could lead to heritable mutations.

In our study, male volunteers, smokers and nonsmokers, were recruited by newspaper solicitation to provide blood, urine, and semen specimens and to complete a self-administered questionnaire. All specimens were analyzed for cotinine by radioimmunoassay, and a standard clinical semen analysis was performed to determine sperm count, density, morphology, and motility. Exclusion criteria were applied to obtain a group of healthy men, aged 18 to 35, who had not had other exposures that might alter sperm form or function. Eighty-eight men participated in the study.

Fig. 2 shows that cotinine concentrations in semen are similar to those in serum; higher concentrations are found in urine. A dose-response relationship for nonsmokers and light and heavy smokers is also evident. In nonsmokers (Fig. 3) cotinine appears in all three body fluids, presumably due to environmental tobacco smoke exposure. There is a suggestion of a dose-response relationship between h/day of environmental tobacco smoke exposure and cotinine concentration in urine.

We found high correlations between log cotinine concentrations in each of the three body fluids, ranging from 0.87 to 0.94. Fig. 4 shows the linear association between log cotinine concentrations in serum and semen.³ The correlations between log cotinine in each

³M. F. Vine, B. S. Hulka, B. H. Margolin, Y. K. Truong, P. C. Hu, R. B. Everson, M. Schramm, J. Griffith, and M. McCann. Cotinine concentrations in semen, urine and serum of smokers and nonsmokers, submitted for publication.

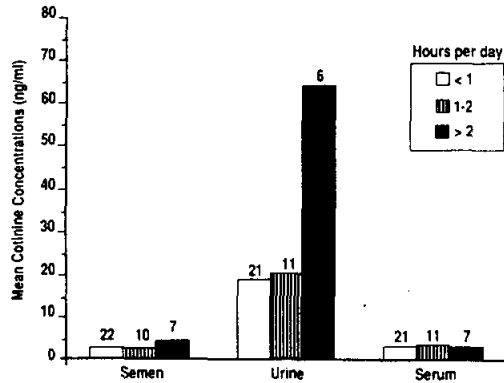


Fig. 3. Cotinine concentrations in semen, urine, and serum by average h/day of environmental tobacco smoke exposure during the previous 3 months (nonsmokers only).

body fluid and number of cigarettes smoked was more modest, ranging from 0.40 to 0.53. There were small negative correlations, -0.20 to -0.22 , between cotinine concentration and sperm density, measured in number of cells/ml.

Cotinine provides an excellent internal dose marker for nicotine in cigarette smoke. But the presence of additional biological markers that appear at later stages in the exposure/disease spectrum would strengthen the argument for biological effects of smoke constituents on sperm cells. In order to provide this evidence, additional studies are in progress. Specifically, we are attempting to detect smoking-related DNA adducts, using ^{32}P postla-

beling assays, in the sperm cells of smokers. If adducts are found, these would provide evidence of mutagenic insult to DNA. Although assay results are not yet available, the design of the adduct study is relevant to an additional aspect of our work.

In the adduct study, subjects were recruited to achieve triplets, composed of a nonsmoker, a light smoker, and a heavy smoker. The heavy smoker and nonsmoker provided the exposure extremes, whereas the light-smoking member provided the opportunity for dose-response analysis. Members of each triplet were matched on age and education; exclusion criteria were applied such that only healthy Caucasian men without exposure to other suspect mutagenic agents were included. Furthermore, only men with a sperm count of at least 40 million were included in order to ensure an adequate amount of DNA for the ^{32}P postlabeling assay.

As the study progressed, it became evident that heavy smokers were being excluded disproportionately from the triplets because of the sperm count criterion. In fact, it was necessary to supplement the advertising to recruit more heavy smokers. This experience caused us to return to the literature on studies of smoking and sperm count or sperm density. The literature yielded several observations. There were 17 interpretable studies in the English language, which fell into two groups, those that were based on healthy volunteers and those that were based on infertility clinic patients. Summary data from these studies are plotted in Fig. 5 (9). Each study is represented by a data point (triangles for infertility clinic patients and squares for healthy volunteers). The data points plot the difference in mean sperm density between nonsmokers and smokers by mean sperm density

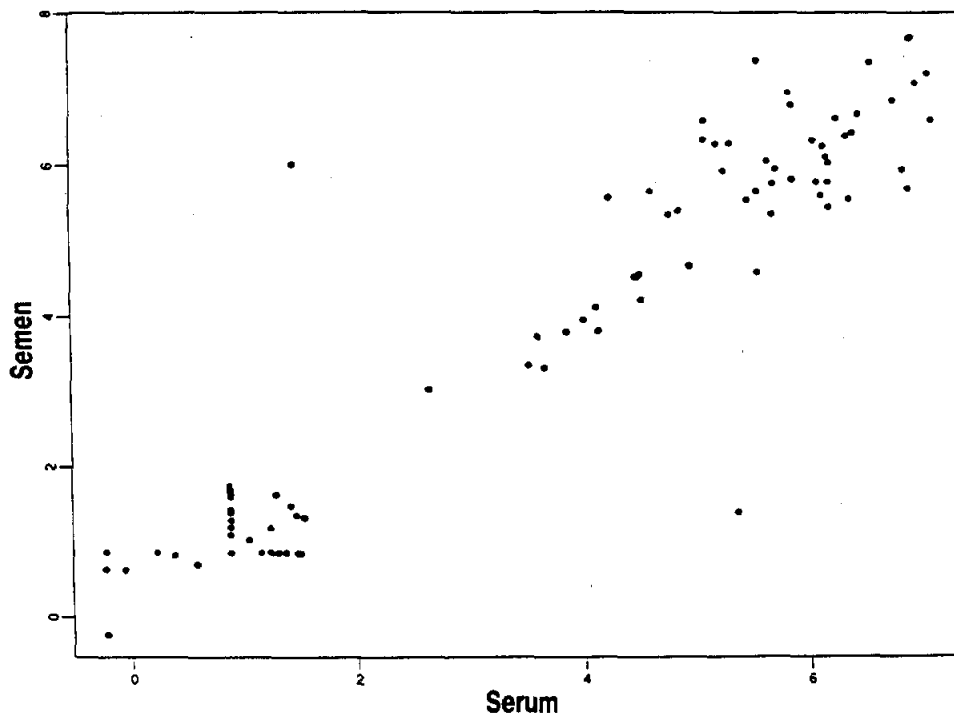


Fig. 4. Log cotinine concentrations (ng/ml) in blood and serum.

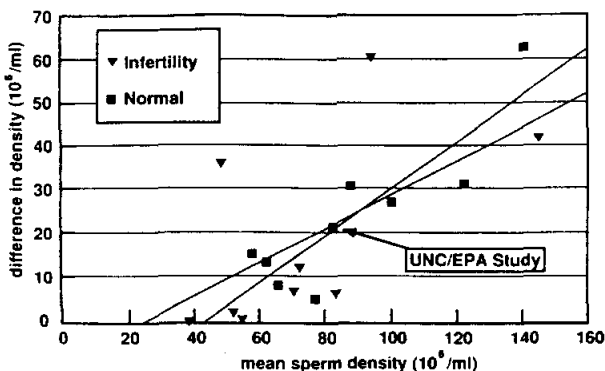


Fig. 5. Studies of sperm density and cigarette smoking in healthy volunteers and infertility clinic patients.

for all subjects in each study. In all of the studies, the difference is positive; nonsmokers have higher average sperm density than smokers, although in the original reports, only one study had indicated that the difference was statistically significant. The regression lines for the two sets of studies are similar in their slopes, and, as it happens, our study lies near the midpoints of these lines. This graphic display is convincing; smokers have a lower average sperm density than nonsmokers. In our study the average sperm density was $93 \times 10^6/\text{ml}$ for nonsmokers and $72 \times 10^6/\text{ml}$ for smokers of 20+ cigarettes/day. A structured metaanalysis is in progress.

If we return at this point to the conceptual framework for biological markers, ranging from exposure through pathogenesis to disease, it becomes clear how the studies I have described relate to the sequence of events in that spectrum. Questionnaire information on active and passive smoking was obtained to measure exogenous exposure. We obtained information on internal dose with cotinine assays in body fluids and on biologically effective dose using ^{32}P postlabeling assays for DNA adducts in sperm cells. Sperm density fits in the conceptual framework at the level of functional and structural alterations. These interrelated studies, using the methods of epidemiology, biochemistry, molecular biology, and biostatistics, have enabled us to identify multiple markers that lie at different points in the biomarker spectrum. This accumulation of information on cigarette smoke, semen, and sperm strengthens the scientific basis for possible adverse reproductive effects of smoking on men.

Constraints on Biological Markers in Epidemiological Research

The application of biological markers to epidemiology creates opportunities but also presents difficulties. First is the issue of collaborative research among scientists from different disciplines. Although the application of knowledge and skills from multiple disciplines to a single health problem is a significant reason for doing research with biological markers, it is also a constraint. Scientists from different disciplines have different goals. Laboratory scientists want to develop new assays, to be on the cutting edge of new laboratory technology. Epidemiologists want standardized, reproducible, inexpensive assays

suitable for high-volume work. We work in different laboratories; the epidemiologist's is population based and the laboratory scientist's is bench based, using small numbers of subjects (specimens) under highly controlled conditions. Laboratory scientists tend to have a deterministic orientation toward science, whereas in epidemiology we think in a probabilistic mode. Years of education, training, and experience are reflected in our attitudes toward and values in science. This indoctrination makes it difficult to appreciate fully the importance of scientific disciplines other than one's own. Differences in orientation and assumptions are major reasons for the difficulties that arise in multidisciplinary research.

The greatest constraint in our current environment, however, is research funding. There are few funding sources that cross disciplinary lines. If funding sources are found, the review process may be an obstacle, since both laboratory scientists and epidemiologists must be enthusiastic about a proposal, which may be viewed as hybrid research and may not be attractive to either expert group. To the extent that the individual components of a proposal can be viewed in the context of the whole, the significance of the research will receive higher priority and more proposals will be funded.

The search for environmental exposures that have significant health effects is the task of environmental epidemiology, which is devoted to identifying biological effects in humans from low-level exposures. One problem is that frequently these exposures are neither sufficiently potent nor sufficiently high to produce a measurable effect. In order to counteract these constraints, more sensitive assays are developed, such as the ^{32}P postlabeling assay, for which the sensitivity approaches 1 adduct/ 10^9 nucleotides. But one must ask, at what point does enhancement of sensitivity result in the detection of alterations that are insignificant with respect to human health and disease?

Replication of findings is necessary in order to gain credibility. In epidemiology, we are accustomed to seeking replication in multiple populations, geographic settings, and time periods. This strategy is not satisfying to all disciplines, however, because of the priority that is given to making new discoveries. Replication of assays in multiple laboratories is necessary in order to be assured of assay reliability, and results should be verified in multiple study groups and settings.

More information is needed on a variety of topics relevant to biomarkers. The background level of specific markers in different populations is one of these. Identification of risk factors and potential confounders for different markers is another. Health effects in relation to various markers are virtually unexplored. Sources and extent of variability assume more importance. These include both biological variability (intra- and interindividual), which can have both genetic and environmental sources, and laboratory variability, which has both random and systematic components. Understanding the natural history of biomarkers, analogous to the natural history of disease, will help us to use biomarkers effectively in epidemiological studies.

Future Directions

Biological markers will play a prominent role in epidemiological research during the 1990s. Multiple markers,

including different points on the spectrum of exposure to disease, will be used. Nonspecific markers and chemical-specific markers may be used concurrently, depending on the goals of the study.

The concept of "completing the parallelogram" will receive more empirical validation (10). The parallelogram relates findings from epidemiological studies, both transitional and traditional, to those from animal models and *in vitro* systems. Knowledge about findings from these other systems allows epidemiologists to design studies that can build on existing laboratory data.

Laboratory capability is essential for biomarker studies, and how epidemiologists relate to the laboratory is also important. A number of options for functional relationships between epidemiologists and laboratory scientists exist. Perhaps most desirable for the epidemiologist is to have laboratory capability within one's own domain. This will require the acquisition of laboratory skills by epidemiologists or epidemiological skills by laboratory scientists. Individuals who can cross disciplines and run laboratories themselves will become valued members of epidemiology departments.

Epidemiologists should establish research collaborations with recognized laboratory experts who have the most advanced laboratory technology and capability. This collaboration provides the strongest scientific links and paves the way for eventual in-house laboratory capabilities.

As assays are replicated, and their use for clinical or regulatory purposes becomes evident, technology transfer from research to clinical and commercial laboratories will take place. To the extent that assays can be done in large numbers and hopefully at modest cost, production laboratories, including commercial ones, will be useful for epidemiology.

A resource for biomarker research is the availability of repositories of biological specimens. Tissue and body fluid repositories have the potential to make studies feasible and efficient from the standpoint of cost versus information obtained. Repositories of biological media that have been collected and stored in the context of prior epidemiological studies are obviously valuable, but archived fixed tissue specimens, which exist in pathology departments, should not be ignored. Although subjects may not be well characterized on epidemiological parameters, innovative means to overcome this deficiency may be possible. Use of archived specimens has important implications for assay technology in that assays using fixed tissues have different technical problems from those using fresh or frozen material.

My last point pertains to those of us in the American Society of Preventive Oncology who are concerned about cancer control and prevention. Biological markers are likely to become a more prominent part of our

research and implementation agendas. Even though past efforts to develop markers for cancer screening have been disappointing, it is hard to envision a significant alternative to the continued search for sensitive and specific biomarkers. The domain of clinical trials is also a fertile ground for marker application. Markers can be used to detect precursor lesions (the response variable in some clinical trials) and intermediate outcomes. They may also serve as part of the inclusion criteria for admission into trials and as indicators of compliance with the intervention.

Acknowledgments

The author gratefully acknowledges the many scientific and technical collaborators on the research reported in this paper. Primary participants from the Health Effects Research Laboratory, U.S. Environmental Protection Agency, include Drs. Richard Everson and Jack Griffith. From the University of North Carolina at Chapel Hill the investigators include Drs. Marilyn Vine, Barry Margolin, Young Truong, and Ping-chuan Hu.

References

1. Subcommittee on Reproductive and Neurodevelopmental Toxicology, Committee on Biological Markers, Board of Environmental Studies and Toxicology, Commission on Life Sciences, National Research Council. *Biological Markers in Reproductive Toxicology*, p. 17. Washington, D.C.: National Academy Press, 1989.
2. Sandler, D. P., and Collman, G. W. Cytogenetic and environmental factors in the etiology of the acute leukemias in adults. *Am. J. Epidemiol.*, 126: 1017-1032, 1987.
3. Poirier, M. C. DNA adduct determination in human tissues and blood cell DNA samples. Presented at the American Cancer Society Mary Lasker Conference, Sarasota, FL, April 3, 1991.
4. Lucier, G. W. Receptor-mediated carcinogenesis. Presented at the workshop on "Approaches to Classifying Carcinogens According to Mechanisms of Action," for the International Agency for Research on Cancer, Lyon, France, June 11-18, 1991.
5. Perera, F., Mayer, J., Jaretski, A., Hearne, S., Brenner, D., Young, T. L., Grimes, M., Grantham, S., Tang, M. X., Tsai, W.-Y., and Santella, R. M. Comparison of DNA adducts and sister chromatid exchange in lung cancer cases and controls. *Cancer Res.*, 49: 4446-4451, 1989.
6. Kleinbaum, D. G., Kupper, L. L., and Muller, K. E. *Applied Regression Analysis and Other Multivariable Methods*. Boston: PWS-Kent Publishing Company, 1988.
7. Margolin, B. H., and Shelby, M. D. Sister chromatid exchanges: a reexamination of the evidence for sex and race differences in humans. *Environ. Mutagen.*, 7 (Suppl. 4): 63-72, 1985.
8. Vine, M. F., Hulka, B. S., Everson, R. B., Margolin, B. H., Hu, E., and McCann, M. Cotinine concentrations in blood (serum), urine, and semen of smokers and nonsmokers. Poster session at the International Society for Environmental Epidemiology, Berkeley, CA, August 1990.
9. Margolin, B. H. Meta-analysis: a quantitative approach to research integration. Presented at the American Cancer Society Meeting, Raleigh, NC, April 1991.
10. Waters, M., Allen, J., Doerr, C., Erexson, G., Hasselblad, V., Huang, S., Kligerman, A., Moore, M., Nesnow, S., Stead, A., Strauss, G., Tabor, J., Westbrook-Collins, B., and Wilmer, J. A preliminary investigation of the parallelogram concept in genetic monitoring and risk estimation. In: M. Sorsa and H. Norppa (eds.), *Progress in Clinical and Biological Research: Monitoring of Occupational Genotoxicants*, Vol. 27, pp. 203-215. New York: Alan R. Liss, Inc., 1986.