

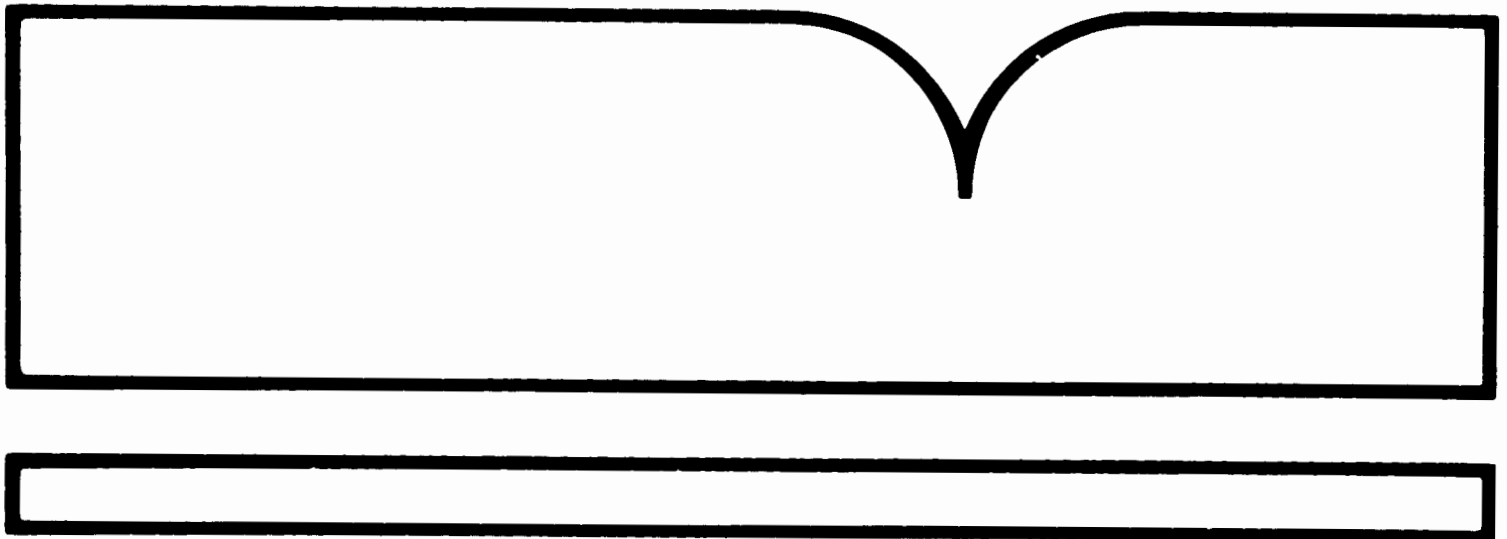
Epidemiology and Health Risk Assessment
Proceedings of a Symposium Held
May 14-16, 1985, Columbia, Maryland

Johns Hopkins Univ., Baltimore, MD

Prepared for

National Inst. for Occupational Safety and
Health, Cincinnati, OH

16 May 85



EPIDEMIOLOGY AND HEALTH RISK ASSESSMENT
PROCEEDINGS OF A SYMPOSIUM
HELD MAY 14 - 16, 1985
COLUMBIA, MARYLAND

Submitted by Leon Gordis, M.D., Professor and Chairman,
Department of Epidemiology, The Johns Hopkins University
School of Hygiene and Public Health, 615 North Wolfe Street,
Baltimore, Maryland 21205

Supported in part by the Centers for Disease Control/
National Institute for Occupational Safety and Health

REPORT DOCUMENTATION PAGE	1. REPORT NO.	2.	3. Recipient's Accession No. PB88 131057/AS	
4. Title and Subtitle Epidemiology and Health Risk Assessment. Proceedings of a Symposium Held May 14 - 16, 1985, Columbia, Maryland			5. Report Date 85/05/16	
7. Author(s) Anonymous			6.	
9. Performing Organization Name and Address Johns Hopkins University School of Hygiene and Public Health, Baltimore, Maryland			8. Performing Organization Rept. No.	
			10. Project/Task/Work Unit No.	
			11. Contract(C) or Grant(G) No. (C) (G)	
12. Sponsoring Organization Name and Address			13. Type of Report & Period Covered	
			14.	
15. Supplementary Notes				
16. Abstract (Limit: 200 words) The 28 contributions to this symposium dealt with various aspects of epidemiology as it relates to the assessment of various risks to the health of workers arising out of hazardous materials or conditions at the workplace. The symposium began with an overview of health risk assessment considering epidemiology and environmental hazards, the scientific basis for policy decisions, epidemiology and risk assessment, and toxicology and epidemiology. Contributions of epidemiology to health risk assessment were considered including estimation of risk and inferring causality in epidemiology, developmental risks, asbestos (1332214) and cancer, genetic susceptibility and the estimation of risk. When considering various refinements to epidemiological approaches to the assessment of exposure both measurements taken in the total environment and measurements in the individual were discussed. The quantitative expression of risk was considered including quantitative expression of dose responses, statistical modeling of dose response relationships, and quantification of risk in defined populations. In the papers evaluating epidemiologic information, methods for combining animal and human data were considered along with papers on saccharin and bladder cancer, estrogens and breast cancer, leukemia risks in relation to benzene (71432) exposure, the influence of theoretical and experimental radiobiology on the epidemiology of radiation carcinogenesis, soft water and hardening of the arteries, and the circumstances of exposure and reproductive consequences.				
17. Document Analysis a. Descriptors b. Identifiers/Open-Ended Terms NIOSH-Publication, Risk-analysis, Risk-factors, Mortality-data, Mortality-rates, Cancer-rates, Carcinogenesis, Worker-health, Body-distribution c. COSATI Field/Group				
18. Availability Statement		19. Security Class (This Report)	21. No. of Pages 591	
		20. Security Class (This Page)	22. Price	

SPONSORING ORGANIZATIONS:

SYMPOSIUM: EPIDEMIOLOGY AND HEALTH RISK ASSESSMENT

American Industrial Health Council

American Petroleum Institute

Centers for Disease Control/Center for Environmental Health

Centers for Disease Control/National Institute for Occupational
Safety and Health

Chemical Manufacturers Association

Electric Power Research Institute

Gas Research Institute

Johns Hopkins University School of Hygiene and Public Health

National Center for Health Statistics

U.S. Department of Energy

U.S. Environmental Protection Agency

PROGRAM COMMITTEE:

SYMPOSIUM: EPIDEMIOLOGY AND HEALTH RISK ASSESSMENT

Leon Gordis, M.D., Dr. P.H., Chairman
Professor and Chairman
Department of Epidemiology
The Johns Hopkins University
School of Hygiene and Public Health

Linda S. Erdreich, Ph.D.
Environmental Criteria and Assessment Office
U.S. Environmental Protection Agency

Carol J.R. Hogue, Ph.D.
Chief, Pregnancy Epidemiology Branch
Division of Reproductive Health
Center for Health Promotion and Education
Centers for Disease Control

Robert K. Miday, M.D.
Procter and Gamble Co.

Abe Silvers, Ph.D.
Electric Power Research Institute

Nat F. Barr, Ph.D.
U.S. Department of Energy

Ian Higgins, M.D., F.R.C.P.
Professor of Epidemiology
School of Public Health
University of Michigan

Genevieve M. Matanoski, M.D., Dr. P.H.
Professor
Department of Epidemiology
The Johns Hopkins University
School of Hygiene and Public Health

Reuel A. Stallones, M.D., M.P.H.
Dean
School of Public Health
The University of Texas Health Science
Center at Houston

PLANNING COMMITTEE:

SYMPOSIUM: EPIDEMIOLOGY AND HEALTH RISK ASSESSMENT

**Nat F. Barr, Ph.D.
U.S. Department of Energy**

**Irwin Billick, Ph.D.
Gas Research Institute**

**Linda S. Erdreich, Ph.D.
Environmental Criteria and Assessment Office
U.S. Environmental Protection Agency**

**Henry Falk, M.D.
Chief of Special Studies Branch
Chronic Diseases Division
Center for Environmental Health
Centers for Disease Control**

**Vernon Houck, M.D.
Director,
Center for Environmental Health
Centers for Disease Control**

**Robert K. Miday, M.D.
Procter and Gamble Co.**

**Martin J. Reape, M.D.
Director, Health Services
FMC Corporation**

**Neil Roth, Ph.D.
Gas Research Institute**

**Leonard Sagan, Ph.D.
Electric Power Research Institute**

**Abe Silvers, Ph.D.
Electric Power Research Institute**

EPIDEMIOLOGY AND HEALTH RISK ASSESSMENT

TABLE OF CONTENTS

I. Overview of Health Risk Assessment

1. Epidemiology and Environmental Hazards	Reuel A. Stallones
2. The Scientific Basis for Policy Decisions	Bernard D. Goldstein
3. Epidemiology and Risk Assessment	Arthur C. Upton
4. Toxicology and Epidemiology -- Strengths and Limitations	Norton Nelson

II. How Has Epidemiology Contributed to Health Risk Assessment?

5. Notes on the Contribution of Epidemiology to Health Risk Assessment	Michel A. Ibrahim
6. Estimation of Risk and Inferring Causality in Epidemiology	Leon Gordis
7. Developmental Risks: Epidemiologic Advances in Health Assessment	Carol J. Rowland Hogue and Marge A. Brewster
8. Asbestos and Cancer	Philip E. Enterline
9. Genetic Susceptibility and the Estimation of Risk	Gilbert S. Omenn

III. Refining Epidemiologic Approaches to Assessment of Exposure

A. Measures in the Total Environment

10. Issues in the Measurement of Exposure	Genevieve M. Matanoski
11. Exposure Measurement for Air Pollution Epidemiology	Benjamin G. Ferris, Jr., J.H. Ware and J.D. Spengler
12. Measurement of Exposure to Arsenic in the Occupational Setting	Ian Higgins, Kathleen Welch and Mary Oh

TABLE OF CONTENTS (cont'd)

B. <u>Measures in the Individual</u>	
13. Relation of Body Burden Measures to Ambient Measures	Philip J. Landrigan
14. Relation of Surrogate Measures to Measures of Exposure	Walter J. Rogan
15. Molecular Cancer Epidemiology: The Use of New Laboratory Methods in Studies on Human Cancer Causation	I. Bernard Weinstein
IV. <u>Quantitative Expression of Risk</u>	
16. Approaches to Quantitative Expression of Dose Responses	Ethel S. Gilbert
17. Statistical Modeling of Dose-Response Relationships	Ron Brookmeyer
18. Quantification of Risk in Defined Populations	William R. Gaffey
V. <u>Evaluation of Epidemiologic Information</u>	
19. Combining Animal and Human Data: Resolving Conflicts, Summarizing the Evidence	Linda S. Erdreich
20. Saccharin and Bladder Cancer	Philip Cole
21. Estrogens and Breast Cancer	Barbara S. Hulka
22. Leukemia Risks in Relation to Benzene Exposures: An Examination of the Data	Darwin R. Labarthe
23. The Influence of Theoretical and Experimental Radiobiology on the Epidemiology of Radiation Carcinogenesis	Charles E. Land
24. Soft Water -- Hard Arteries: An Interpretation of Ecologic Findings	George W. Comstock
25. Circumstances of Exposure and Reproductive Consequences	Jennie Kline and Zena Stein
26. Evaluation of Epidemiologic Information	Moyses Szklo
VI. <u>Conclusion</u>	
27. Risk Assessment and Public Policy	Edward N. Brandt, Jr.
28. Emerging Needs in Epidemiology and Risk Assessment	Leon Gordis

CONTRIBUTORS

SYMPOSIUM: EPIDEMIOLOGY AND HEALTH RISK ASSESSMENT

Edward N. Brandt, Jr., M.D.
Chancellor
University of Maryland at Baltimore

Marge A. Brewster, Ph.D.
Director, Metabolic Laboratory
Arkansas Childrens' Hospital
and Associate Professor
University of Arkansas
School of Medicine

Ron Brookmeyer, Ph.D.
Assistant Professor
Department of Biostatistics
The Johns Hopkins University
School of Hygiene and Public Health

Philip Cole, M.D., Dr. P.H.
Professor of Epidemiology
School of Public Health
University of Alabama at Birmingham

George W. Comstock, M.D., Dr. P.H.
Alumni Centennial Professor of Epidemiology
Department of Epidemiology
The Johns Hopkins University
School of Hygiene and Public Health

Philip E. Enterline, Ph.D.
Director, Center for Environmental
Epidemiology and
Professor of Biostatistics
Graduate School of Public Health
University of Pittsburgh

Linda S. Erdreich, Ph.D.
Environmental Criteria and Assessment Office
U.S. Environmental Protection Agency

Benjamin G. Ferris, Jr., M.D.
Professor of Environmental Health
and Safety
Harvard School of Public Health

William R. Gaffey, Ph.D.
Manager, Epidemiology
Monsanto Company

CONTRIBUTORS (cont'd)

Ethel S. Gilbert, Ph.D.
Staff Scientist
Battelle-Pacific Northwest Laboratories

Bernard D. Goldstein, M.D.
Assistant Administrator for Research
and Development
U.S. Environmental Protection Agency

Leon Gordis, M.D., Dr. P.H.
Professor and Chairman
Department of Epidemiology
The Johns Hopkins University
School of Hygiene and Public Health

Ian Higgins, M.D., F.R.C.P.
Professor of Epidemiology
School of Public Health
The University of Michigan

Carol J. Rowland Hogue, Ph.D.
Chief, Pregnancy Epidemiology Branch
Division of Reproductive Health
Centers for Health Promotion and Education
Centers for Disease Control

Barbara S. Hulka, M.D., M.P.H.
Professor and Chair
Department of Epidemiology
University of North Carolina

Michel A. Ibrahim, M.D.
Dean and Professor of Epidemiology
University of North Carolina
School of Public Health

Jennie Kline, Ph.D.
Research Scientist
New York State Psychiatric Institute and
Columbia University

Darwin R. Labarthe, M.D., Ph.D.
Professor and Convener of Epidemiology
The University of Texas
Health Sciences Center at Houston

CONTRIBUTORS (cont'd)

Charles E. Land, Ph.D.
Health Statistician
Radiation Epidemiology Branch
National Cancer Institute

Philip J. Landrigan, M.D.
Director, Division of Surveillance
Hazard Evaluations and Field Studies
Robert A. Taft Laboratories, NIOSH

Genevieve M. Matanoski, M.D., Dr. P.H.
Professor
Department of Epidemiology
The Johns Hopkins University
School of Hygiene and Public Health

Norton Nelson, Ph.D.
Professor
New York University Medical Center
Institute of Environmental Medicine

Mary Oh, B.Sci.
Programmer
Department of Epidemiology
University of Michigan
School of Public Health

Gilbert S. Omenn, M.D., Ph.D.
Dean
School of Public Health and Community Medicine
Professor of Medicine (Medical Genetics)
and of Environmental Health
University of Washington

Walter J. Rogan, M.D.
Medical Officer, Epidemiology Branch
National Institute of Environmental
Health Science

John D. Spengler, Ph.D., M.S.
Professor of Environmental Health Science
Harvard School of Public Health

Reuel A. Stallones, M.D., M.P.H.
Dean
School of Public Health
The University of Texas Health
Science Center at Houston

CONTRIBUTORS (cont'd)

Zena A. Stein, M.A., M.B., B.C.H.
Director, Epidemiology of Brain Disorders
Research Department
New York State Psychiatric Institute
Professor of Public Health (Epidemiology)
Columbia University

Moyses Szklo, M.D., Dr.P.H.
Professor
Department of Epidemiology
The Johns Hopkins University
School of Hygiene and Public Health

Arthur C. Upton, M.D.
Professor and Chairman
Department of Environmental Medicine
New York University

James H. Ware, Ph.D., M.S.
Professor of Biostatistics
Harvard School of Public Health

I. Bernard Weinstein, M.D.
Professor of Medicine and Public Health
and Director, Division of Environmental Sciences
Columbia University

Kathleen Welch, M.P.H.
Research Assistant
Department of Epidemiology
School of Public Health
University of Michigan

Part I

Overview of Health Risk Assessment

Chapter 1
Epidemiology and Environmental Hazards
Reuel A. Stallones

EPIDEMIOLOGY AND ENVIRONMENTAL HAZARDS

Reuel A. Stallones*

Risk assessment has been defined as one of four steps involved in public responses to environmental hazards, namely, hazard identification, risk assessment, exposure assessment, and risk management (Table 1-1). Thus, in a biomedical context, risk assessment is a set of somewhat formal procedures that lie between the accumulation of scientific information concerning hazards to human health and the adoption and implementation of measures to control those hazards. It entails assembling and evaluating a body of data, a more general approach than the review of a specific project. A public response to the control of an environmental hazard is an implementation of public policy and therefore risk assessment becomes entangled with policy, contributing to its formulation, and inevitably coming under its influence (Lave and Seskin 1979; Stallones 1982).

Chemical and clinical observations can be important in the first step, the detection of a potential problem, but they do not provide the information needed to develop a formal numerical estimate of risk. For this, epidemiological information or data from experimental exposure of animals is necessary. (Table 1-2)

The results of the latter, the data developed in toxicology laboratories are used in a process that has been labelled quantitative risk assessment, an effort to estimate the risk to humans. The applicability of this procedure is limited almost entirely to carcinogenic agents, for this is the only situation

*Deceased June 22, 1986

in which results obtained from animal experimentation are sufficiently analogous to human experience to serve as the basis for estimating the risk to humans. Quantitative risk assessment has been adopted widely by federal regulatory agencies. Unfortunately, the estimation of hazards to human health from indirect measures involves two large and uncertain inferential steps. The difficulty has been well expressed by Hoel, Kaplan, and Anderson (1983):

The problem of estimating the risk of cancer...from long-term exposure to low doses of a chemical carcinogen is difficult to resolve since there are rarely sufficient data available from humans. An alternative approach to the problem is to estimate low-dose risk in laboratory animals and then extrapolate these results to humans. The main disadvantage of this approach is that the exposure levels of interest are so low that an impracticably large number of animals must be used to obtain a reliable estimate. The strategy for circumventing this problem has been to experiment at doses that are much higher than those encountered in the environment, so as to produce tumors at a higher frequency. The results are then extrapolated to the dose levels of interest.

The standard way of doing the low-dose extrapolation has been to assume a parametric model relating applied dose to tumor response, and to use the high-dose data to estimate the model parameters. The choice of model is controversial since various models may adequately fit the data in the experimental range but produce risk estimates at environmental levels that vary by many orders of magnitude.

Having estimated the risk for laboratory animals at low doses by extrapolating from the scanty high dose observations, the next step is to convert that risk statement into human terms, and this calls for a second order extrapolation, based on some notion of a proper characteristic for conversion,

such as body weight, surface area, or metabolic rate. This choice may introduce a difference in the estimated risk of forty fold, less than the variation among the low dose extrapolation models, but still appreciable. (Office of Technology Assessment 1981).

The final result of this approach to risk assessment is a prediction in the general form that the continued exposure of people to chemical X will, in the population of the United States, be responsible for inducing Y number of cases of cancer Z in a specified calendar year. That number cannot be correct except by accident, and if all of the assumptions involved in reaching that number are taken into account, the confidence limits must extend from zero to infinity. However, the presentation of a number that looks real and that is supported by some arcane mathematics gives an illusion of reality which seems a firm enough basis for governmental regulatory action. The uncertainties attending these extrapolations, together with the effects of the many inferences required to bridge gaps in our knowledge of the relations between exposure to a substance and the development of disease, provide ample ground for disagreement over the results of risk assessment. When the decisions based on these analyses affect the transfer of millions of dollars, the disagreements can often be bitter.

The shortcomings of bioassays in assessing the risk to people of exposure to toxic chemicals include the following:

The methods employed entail extrapolation of effects in animal systems from the large doses used experimentally to the much smaller doses likely to be encountered in environmental exposures. Unfortunately, the mathematical models used for this extrapolation have no biological relevance (Office of Technology Assessment 1981; Park and Snee 1983).

The evidence is indirect; to convert the results of bioassays to effects expected in human beings depends upon a belief that the non-human systems respond to chemical exposures in the same way that people do, and on an extrapolation based on differential size or metabolic rate. If the first is incorrect, the approach fails completely, and if the second is incorrect the quantitative risk estimate will be seriously in error.

Bioassay as an approach to risk assessment is virtually limited to the cancers induced by exposure to chemicals and to ionizing radiation. Most of the illnesses incited by environmental hazards are not even approximately replicated in experimental animals, and for these conditions there is no choice but to rely upon clinical and epidemiological observations of people.

Of primary interest is the epidemiological contribution to the risk assessment process. The basis of epidemiology is the fact that disease and disability distribute predictably in human populations. The methods of epidemiology are designed to measure variations in risk, and identify the characteristics of people and their environments that are predictive of those variations. Some of those characteristics are explanatory and some are capable of modification in such a manner as to reduce the risk; this defines the purpose of epidemiology, to obtain the information needed to prevent disease. Thus, the results of an epidemiological study are a risk assessment, and the best predictions of risk are those based upon the aggregated epidemiological observations, epidemiological risk assessment (Squire 1981; Doll 1981). Indeed, other sources of information become important only when direct information is not available.

The usual methods of epidemiological risk assessment bear little resemblance to the process of quantitative risk assessment; rather they depend more directly on the traditional practices of science. Scientists concerned with the status of a field of research may assemble and assess the body of evidence, and publish the conclusions as a review article or commentary. Some authors bring great insight to bear on a problem, and these publications are immensely valuable in the development of a field of inquiry. These offerings customarily appear in the open literature and are available for criticism. Another useful device is a review conference or a designated committee; a meeting of persons especially competent to evaluate the assembled information on a specified issue. This may be as sharply focused and specific as the assessment of the relation between toxic shock and the use of tampons conducted by an outstanding group of scientists assembled under the aegis of the Institute of Medicine, or as global, geographically and conceptually, as the report of a World Health Organization committee on Health Hazards of the Human Environment (Committee on Toxic Shock Syndrome 1982b, 1982c). The group may consider the various environmental correlates of a specific disease, or the various diseases associated with a selected component of the environment (Committee on Toxic Shock Syndrome 1982a; Breslow and Whittemore 1979).

As exemplified by these mechanisms, epidemiological risk assessment entails a comprehensive and evaluative review of the entire body of research data concerning the health hazard in question, and the development of point estimates of risk for different intensities and circumstances of exposure. If the body of research data is extensive and of good quality, then the usual

measures of risk produced in epidemiological studies provide a clear and reliable basis for projecting risks expected to occur from present and future exposures to a toxic chemical. Commonly, however, the data are sparse, and some aspects of most studies are subject to criticism. Most epidemiological risk assessments, therefore, rely heavily on interpretation and judgment. Some of the most difficult situations arise when the data from different studies indicate widely differing statements of risk. An equally difficult problem exists in the interpretation of studies in which no health effect was found. When the absolute risks are small, as they commonly are, when the number of persons included in a study is small, and when the health effects may occur many years after the exposure, the failure to observe an effect is not convincing evidence that the exposure is innocuous. Even in these unfavorable circumstances negative studies can help to determine the probable upper limits of risk. The projection of risk based on an upper confidence limit does not appear justifiable, although it is defended as a public policy that emphasizes safety. That may be a desirable social goal, but should not be promoted at the expense of making an estimate of risk a function of the number of people who were included in a study.

The nature of epidemiological evidence is not widely understood, and this lack of understanding leads to misdirected criticism. Difficulties that are inherent in research on free-living people are improperly interpreted as deficiencies of the science, of a study method, or of a particular project. The insensitivity of examination procedures to minimal signs of illness, and the inability to detect a small excess of events against the background of occurrence are taken as evidence that negative epidemiological studies are meaningless. Probability statements and confidence limits are treated as though they define an island of truth in a sea of ambiguity. These specious issues coexist with a set of legitimate limitations. A number of important problems attend the use of epidemiological data to characterize the risks of exposure to environmental hazards. (MacMahon 1979)

Stallones

Epidemiological research is commonly protracted and expensive. Studying people in their natural habitat entails costs of tracing and follow-up, and monitoring the experience of a cohort of exposed persons long enough to observe the occurrence of a disease with a latent period as long as a generation is never a simple task. Reconstructing the experience of population groups from records assembled for other purposes may greatly curtail the time required for a study, but, of course, at the risk of losing precision and relevance.

Epidemiological data are observational; only rarely can we perform controlled experiments manipulating environmental hazards. To some, observational data are not convincing; they want to see a smoking gun. Lave and Seskin (1979) wrote, "...when the stakes are large, some scientists will not accept epidemiological evidence as proof of causation, no matter how elaborate it is. Second, most scientists accept a relationship as causal only when there is either laboratory evidence or a theory that specifies detailed functional mechanisms." The counter argument is that experimental research is heavily at risk of introduction of bias of a particularly subtle kind, bias resulting from the presence of the investigator. Observational research is not completely free of this problem, but at least some of the events occur outside the immediate presence of the researcher. Not everyone is happy equating experimental control with inherent bias.

Evaluation of the extent and nature of exposures of people to environmental hazards is commonly uncertain and inferential. This difficulty is compounded by the fact that exposures of concern may have occurred twenty or more years ago, or the effects of contemporary exposures may not appear until twenty or more years in the future. Thus, the current disease experience of employees of some industry may be the result of exposures in the past which were not then measured, and may be inferred only from surrogate information, such as job title. During the years that the diseases are developing, employees usually are exposed to many different hazardous substances, in different workplaces and in the general community environment.

Much of the most valuable epidemiological information was obtained from high dose occupational or catastrophic exposures. If this is to be used to predict low dose effects in a general population, an extrapolation is required, with many of the same problems that accompany extrapolation of data from experimental toxicology (Land 1980).

Individual susceptibility to toxic materials may vary, and may be related to age, sex and, race, and therefore may be confounded with employment, place of residence, and other factors related to exposure.

The health-related effects of exposure to environmental contaminants may be subtle and difficult to detect. This is especially true if the effects are manifested as behavioral change, and if the exposures are low, intermittent, and chronic.

The detectability of disease in the general population is a function of the character of the disease process, the relative risk, the absolute risk and the background expectation. The most difficult conditions to study epidemiologically are those that are very rare and those that are very common. Modestly increased risks of common diseases may be lost against the background occurrence. In the case of rare diseases, the number of persons heavily exposed to an inciting cause may be very small, and, when large numbers of persons are exposed, the dose may be very low. If a disease is distinctive or severe, it may be readily detected even though it is not common. For example, during the swine flu immunization program, Guillain-Barre syndrome was ascertained at a frequency of about one case per 100,000 exposed persons. We are now aware that the risk of paralytic poliomyelitis associated with the oral polio vaccine is well below one case for each million persons immunized. In the period before the disease became rare, to detect an excess risk of less than one per million would have been unimaginable.

Epidemiological risk measurement is a damage assessment; the risk cannot be measured until after the damage has occurred. If there is reason to suspect that current exposures may lead to the development of serious disease problems in the future, no reasonable person would be willing to wait for the aggregation of that disease experience before taking action to control the hazards.

The basic problem with risk assessment, whatever the approach, is the inadequacy of the information base. If there were a full file on each suspected hazard, the statement of risk would be relatively straightforward and strongly defensible. However, for most of the substances and situations

for which there is concern, direct observations of the risk to exposed persons is lacking or incomplete, and the human suffering and death that would be imposed by waiting for secure epidemiological evidence of damage cannot be justified. That inadequacy encourages the use of non-human test systems, but until they become much more reliable predictors of human illness, they are not a satisfactory surrogate for observation of people. Thus, in these terms, neither of the two risk assessment options is acceptable. That may be too harsh; risk assessments have been used to guide risk management, sometimes to almost everyone's satisfaction. However, posing the problem in those terms indicates clearly the need for improvement.

Suggestions on how to improve risk assessment have been forthcoming. Risk assessment based on toxicological data would be much more acceptable without the spuriously precise numerical statement, and the method could be employed usefully to make semi-quantitative statements about the relative hazards of different chemicals, thus allowing for a rank order (Squire 1981). The ranking might not be precisely applicable to human health, but could be defended much more confidently than the usual kind of quantitative risk assessment. Additional effort could well be devoted to developing means of relating the limited data on humans more closely to the toxicological research, providing reality checks to the extent possible.

On the epidemiological side, a great amount of human experience has been accumulated over years which has not been fully exploited; plowing back through old records would give useful information about the risks of exposure to some members of the array of chemicals in the environment. Surveillance of exposed populations is still grossly inadequate (Doll 1981), and linkage of existing records of employment, disability, and death would greatly

increase the knowledge base of many industrial hazards. The quality of epidemiological research is not always above reproach and there is a need to continue to train researchers well and to develop more effective research techniques.

Increasing awareness of the health effects of exposure to hazardous substances, both in the workplace and in the general community environment, has contributed to the rising clamor for governmental action to control the hazards, and the Congressional response has been to expand the authority of federal agencies to regulate these hazards. Although many of the regulatory actions have been accepted, some engendered conflict, often carried into the courts, even as far as the Supreme Court, for adjudication. This phenomenon calls for closer consideration and analysis.

Superficially, the issue appears to be one of disagreement over the interpretation of scientific data. In calmer days, a scientist would make results of research known to peers through accustomed channels, and, if the question were important enough, colleagues would attempt to replicate the findings in their laboratories. As information accumulated, a consensus developed, the newly discovered facts would find their way into textbooks, and students would be required to learn them in order to pass their examinations. If serious disagreements arose, the more emotional colleagues might initiate an exchange of polemics in the Letters to the Editor of a scientific journal. This usually terminated if one of the antagonists grew weary or if the aggregation of information made a position untenable. The suggestion that such a question should be resolved by referral to a person trained in the law or that the issue be addressed in legislation while the controversy raged, would have been abhorrent. The reasons for the change

from the situation just described to the present one are found in the nature of public policy and in the nature of the evidence on which risk assessments are based. Public policy requires that hazardous materials be regulated, and regulation cannot wait for the fine points of scientific merit to be settled by traditional mechanisms. When regulatory actions are taken on the basis of insecure evidence, then some will feel themselves aggrieved and will seek redress in the courts. Nothing that scientists say or do seems likely to change any of those imperatives. There is a responsibility to adapt existing procedures to social exigencies to the extent that this is compatible with scientific integrity. Reserving that as a minimum non-negotiable position, scientists must resign themselves to having research translated into legislation before it is ripe, and to an occasional request to explain the fine points of selection bias and confounding to a federal judge.

REFERENCES

- Breslow, N.E. and Whittemore, A.S. (eds.) (1979). Energy and Health. Soc. Indust. Appl. Math. Philadelphia.
- Committee on the Institutional Means for Assessment of Risks to Public Health (1983). Risk Assessment in the Federal Government: Managing the Process. National Academy Press, Washington D.C.
- Committee on Toxic Shock Syndrome (1982a). Environmental Epidemiology: Risk Assessment. Soc. Indust. Appl. Math. Philadelphia.
- Committee on Toxic Shock Syndrome (1982b). Health Hazards of the Human. WHO, Geneva.
- Committee on Toxic Shock Syndrome (1982c). Toxic shock syndrome: Assessment of current information and future research needs. National Academy Press, Washington D.C.
- Doll, R. (1981). Relevance of epidemiology to policies for the prevention of cancer. J. Occup. Med. 23:601-09.
- Hoel, D.G., Kaplan, N.L. and Anderson, M.W. (1983). Implication of non-linear kinetics on risk estimation in carcinogenesis. Science 219:1032-37.
- Land, C.E. (1980). Estimating cancer risks from low doses of ionizing radiation. Science 209:1197-203.
- Lave, L.B. and Seskin, E.P. (1979). Epidemiology, causality, and public policy. Am. Scientist 67:178-86.
- MacMahon, B. (1979). Strengths and limitations of epidemiology. In, The National Research Council in 1979: Current Issues and Studies. National Academy Press, Washington D.C.
- Office of Technology Assessment, (1981). Assessment of Technologies for Determining Cancer Risks from the Environment. USGPO, Washington D.C.

Park, C.N. and Snee, R.D. (1983). Quantitative risk assessment:

State-of-the-art for carcinogenesis. Fundam. Appl. Toxicol. 3:320-33.

Squire, R.A. (1981). Ranking animal carcinogenesis: A proposed regulatory approach. Science 214:877-80.

Stallones, R.A. (1982). Epidemiology and public policy: Pro- and antibiotic. Am. J. Epidemiol. 115:485-91.

TABLE 1-1. Steps leading to the control of environmental hazards to human health

- A. Hazard identification
 - 1. Observation
 - a. Direct information from human experience
 - (1) Clinical
 - (2) Epidemiological
 - 2. Inference from chemical structure
 - B. Risk assessment
 - 1. Direct: epidemiological risk assessment
 - 2. Indirect: quantitative risk assessment (bioassay data)
 - C. Exposure assessment
 - Industrial engineering, environmental science, topography, hydrology, demography, and others
 - D. Risk management
 - Technical feasibility, economic, social and political considerations
-

Source: Committee on the Institutional Means for Assessment of Risks to Public Health (1983).

TABLE 1-2: Comparison of methods of risk assessment.

Methods	Advantages	Disadvantages
Clinical observations	Cases carefully defined	Cannot quantitate risk; limited to moderately severe illness; strong selection bias
Epidemiological studies	Direct evidence; quantitative	Expensive; time consuming; requires expression of the hazard; problems with power to detect excess risk; measurement of exposure difficult
Bioassays, experimental animal exposures	Need not wait for human illness to occur; relatively inexpensive; exposure known	Indirect evidence, requiring extrapolation; limited to diseases that have suitable animal models

Chapter 2

The Scientific Basis for Policy Decisions

Bernard D. Goldstein

THE SCIENTIFIC BASIS FOR POLICY DECISIONS

Bernard D. Goldstein

Introduction

This chapter is based on my experiences as Assistant Administrator for Research and Development of the US Environmental Protection Agency. It will focus on what I believe to be the very major value of Risk Assessment as a formal process to provide an independent scientific basis for policy decisions. Obviously, my experience is not necessarily pertinent to the broad range of regulatory decision-making in federal agencies, nor to another ERA at EPA.

When William Ruckelshaus returned to EPA in 1983 he determined to follow up on the recommendations of an NAS panel which had called for a concerted effort to utilize scientific risk assessment as a basis for regulatory decision making, and to separate the process of risk assessment from that of risk management (Committee on the Institutional Means for Assessment of Risks to Public Health 1983). The time was ripe at EPA for such an approach.

Before describing the advantages and disadvantages of risk assessment as a basis for policy decisions, let me immediately point out that it is not that simple to separate cleanly risk assessment from risk management. This is partly due to the fact that characterization of risk inherently is a process which interfaces between risk assessment and risk management. More important is the fact that science and scientists tend to be used in two different ways in the American regulatory process, and there is much confusion between these two roles.

Two Approaches to Science in a Regulatory Agency

As a simplification, let me describe the two approaches as Scientific Consensus and Legal Confrontation, and use as an example the question of what

is the lowest observed effect level of a well-studied air pollutant for which the Administrator of EPA must set a standard. If one asked every qualified scientist in the world to review the pertinent literature, a graphic plot of the individual judgments of these scientists as to lowest observed effect level would likely show a classic bell-shaped curve with some scientists at either end of the distribution but most clustering in the center. Obviously, the Administrator of EPA cannot poll every scientist. Accordingly, the standard approach has been to hold a meeting of a sample of these scientists, chosen in an unbiased fashion, under the auspices of an advisory group such as the National Academy of Sciences of EPA's Science Advisory Board. At such a meeting the usual outcome is a movement toward the center among those scientists who might be outliers. This is because of the usual fear among scientists of being wrong - we have more to lose by being wrong than we can gain from being right. We tend to move toward the opinion of other scientists whom we respect, thus achieving a consensus.

Now let us consider the Legal Confrontation approach to the use of scientific information and scientists. A well-trained lawyer, acting absolutely appropriately for his profession, will not depend upon a random sample of qualified scientists. Rather, the lawyer will obtain the judgment of every qualified scientist and then choose the outliers on one side of the issue. The lawyer will correctly assume that an opposing lawyer will also be finding those scientists who are two standard deviations or more from the mean on the other side of the issue. Further, the scientists who are at either end of the judgment scale will not have an opportunity to sit down and discuss the issue before being put into a confrontational adversary hearing.

The point is not which of these two processes, confrontation vs. consensus, is the correct one. The legal approach is an inherent part of the American regulatory process. What we must recognize is that both processes occur, often simultaneously, in a regulatory agency such as EPA, and that this leads to much confusion as to what science, and scientists, are saying about a particular issue. It is crucial to any hope of providing a sound scientific basis for policy decisions that we clearly distinguish between these two different uses of scientists in a regulatory process.

I should point out that this possibility of confusion in the use of science and scientists in an agency such as EPA should not be construed as a rationale for removing science from EPA. That experiment, in the form of NIOSH/OSHA, has not proven to be as effective a means of developing health-based regulations, in part because of the lack of direct scientific input in OSHA.

Risk Assessment

The components of risk assessment are hazard identification, dose-exposure estimation, exposure assessment and risk characterization (Risk Assessment in the Federal Government: Managing the Process 1983). Some degree of controversy exists for all of these components. However, we must be careful not to be misled by controversy, an inescapable aspect of almost any regulatory process.

Controversy about risk assessment is inevitable. The stakes are high. There are individuals in all parties to the regulatory process, in industry and in environmental groups or government, who have a vested interest or who honestly believe that their side has a better chance of success if the process is conducted totally in a confrontational manner, and who would prefer to include the scientific assessment of risk as part of the confrontation. These individuals tend to focus on the problems in doing risk assessment, and

Goldstein

problems do exist. This focus on the problems of risk assessment is accentuated by the natural tendency for the press to pay attention to a topic of this nature only when it is controversial. In point of fact, however, most risk assessments are reasonably straightforward. EPA has performed risk assessment of one form or another of literally hundreds of chemicals with relatively few receiving major outside attention.

One way of putting risk assessment in perspective is in terms of data quality objectives. Simply put, what is the quality of data necessary for a particular regulatory decision? Risk managers make use of numerical estimates of all types including assessment of economic costs of a proposed regulation and the percent likelihood of compliance. The scientific desire for exactitude in risk assessment may lead to a narrower extent of uncertainty than absolutely necessary in the context of the economic and policy uncertainties.

Hazard identification, the first step in Risk Assessment, presents an excellent example of the inevitability of controversy when risk assessment is performed for regulatory purposes. One might think that the qualitative identification of a compound as having a certain hazard would be reasonably straightforward. However, the need to regulate suspected hazards can lead to controversy, particularly in the case of carcinogens. Of the 65,000 chemicals in general use, about two dozen are generally agreed to be known human carcinogens. Certainly many of the 65,000 are definitely not human carcinogens. For most, however, the hazard is unsure. We can make statements about each compound based upon structure-activity relationships, and for many we may have some short-term in vitro or occasionally longer term animal data. The point I wish to emphasize is that there is a continuum of evidence ranging from a minor degree of concern, based on structure, to major concern about a compound for which there are ample long term studies showing cancer in multiple

Goldstein

organs of both sexes in multiple species, but for which there is not human data. It is obvious that federal regulatory agencies have been given a mandate to protect the public against the latter compound, despite it not being a proven human carcinogen. The problem is where to draw the line in this continuum of evidence. It is a delusion to believe that one can find a distinct point that will divide this continuum into two discrete boxes containing compounds for which the hazard of carcinogenesis clearly requires or does not require regulation. Wherever the line is drawn there will be compounds just above or just below the line, and reasonable scientists will differ as to where these compounds should be placed.

What makes matters worse is that the regulator has particular difficulty placing a narrow line through this continuum. Regulatory laws generally are not sufficiently clear, often by design, to designate the degree of likelihood that hazard exists upon which regulations should be based. The problem is particularly acute for the Administrator of EPA. In comparison to other federal regulatory agencies concerned with carcinogens, EPA was not organized by a specific act of Congress embodying a regulatory philosophy. Rather, it was put together by Administrative Act from pieces of other federal agencies. There are at least nine different regulatory acts administered by EPA. They come from different ERAs, different Congressional Committees, and use different language in describing the extent to which protection against the potential for hazard is desired. Logically, these laws could be interpreted in a way that a potential human carcinogen should be regulated in one medium or circumstance, but not in another. However, it is most unrealistic to expect the Administrator of EPA, or any other responsible official, to elect to regulate a potential human carcinogen under one environmental law but leave the same

compound unregulated under another. Imagine trying to explain that to the public.

Dose response estimation has received much attention in terms of the many assumptions inherent in extrapolation to regulate ambient pollutant levels. This has been particularly problematic for carcinogens and other genetic effects for which no threshold is assumed. In essence, our society has adopted a policy that all carcinogens are guilty of having a threshold until proven innocent. Of note is that the mechanistic research responsible for providing such scientific proof of innocence has begun to develop, and EPA has started to act on this information.

Exposure assessment is an aspect of risk assessment that is ripe for major improvement. As a generalization from my experience at EPA, the greatest degree of uncertainty concerning the risk of a known hazard has been that surrounding our knowledge of exposure, although it is the extent of hazard which tends to draw the most attention, and the most controversy. Further, for many compounds we have a much greater opportunity at a relatively smaller investment to narrow the extent of uncertainty concerning exposure than we do for hazard.

Exposure can be assessed by several techniques: mathematical modeling of the source receptor relationship which is often based on emission data, knowledge of transport characteristics (e.g. meteorology, hydrogeology), and demography; direct measurement of the pollutant in media such as air, water and food, also including estimation of population distribution of the media; and personal monitoring in which the receptor is directly assayed, in the case of humans usually through sampling of body fluids.

EPA has done an excellent job of developing mathematical models to estimate population exposure. Unfortunately, this has sometimes become an end in
Goldstein

itself, an example being the arsenic-producing smelter in Tacoma, Washington. In this case Administrator Ruckelshaus attempted to inform the local public about the risks associated with this smelter operation and to elicit their comments. These risks were calculated based on a mathematical model which included such assumptions as each individual spending 70 years lifetime outside their front door. The more appropriate approach of measuring urine arsenic, a good indicator of total body exposure, revealed that the risk described to the public had been overestimated by a factor of 15, i.e. the model overestimated the true risk.

We can anticipate further improvement in exposure assessment through use of two techniques that will also be particularly useful for epidemiology: advances in our ability to devise personal monitors should allow characterization of individual exposure throughout complex daily activities, particularly to air pollutants; the techniques of molecular biology should allow a rapid increase in the number of compounds for which biological monitoring is a reasonable approach.

Risk characterization is the area in which risk assessment most overlaps with risk management. Some individuals in both industry and environmental groups would like to simply characterize risk in a non-quantitative way - i.e. hazard identification, and then deal with this risk through the political process. Other dislike EPA's choice of a single number, the plausible upper bound, as a means to characterize risk. They rightly point out that this single number gives no hint of the range of uncertainty and is hence misleading to the public. On the other hand, use of a range, which implies to a scientist that some midpoint is much more likely than either extreme, misleads the lay public by implying that all numbers within the range are equally likely. It must also be kept in mind that risk has a non-numerical dimension to the

Goldstein

public. We tend to be biased about certain types of risk and in dread of others, quite apart from the actual risk number.

My view of the entire process is that it really is not two steps - risk assessment/risk management; but rather it is a three step process. The first step can be called science policy. It is a step in which a variety of generic decisions are made, ranging from whether extrapolations from animal to man should be done by weight or by surface area, to the appropriate means to characterize the risk. These first step decisions must be made in advance of the risk assessment. They must combine the best scientific information available, for example the most appropriate model for extrapolating carcinogenic data in animals to man, with openly arrived at policy decisions related to the degree of conservatism appropriate in the extrapolation and characterization of risk. This science policy step must be separate from individual risk assessments so as to not serve as a temptation for the risk manager to put body english on the risk assessment.

Use of Risk Assessment: Benzene as a Case Study

Control of benzene as a toxic chemical in the air provides an example of the use by William Ruckelshaus of risk assessment to help guide regulatory decision making. Benzene is a known human carcinogen which is presumed to have some finite risk of producing leukemia even at one molecule of exposure. Using modern advanced analytical techniques, it is readily detectable in ambient air, although usually in the one part per billion range or less (Occupational standard, 10 ppm TLV). As a component of petroleum products, including about 1% of gasoline, and as a major chemical building block, it is difficult to ban and costly to regulate. On arrival at EPA, Administrator Ruckelshaus was faced with a need to move ahead on a proposal by the Carter Administration to regulate five industrial sources of air benzene. A matrix was laid out for
Goldstein

each of these sources which included information on a number of existing facilities, yearly benzene emissions before and after regulation, percent of total stationary source emissions, maximum lifetime individual risk before and after regulation, expected annual cancer cases due to benzene before and after regulation, capital costs, and annual costs for each source regulated (Goldstein 1985). Based on this total information, the Administrator chose to regulate two of these sources, benzene fugitive emissions and coke by-product emissions which was estimated to result in 0.31 and 2.37 fewer leukemia cases each year respectively. The three other sources, for which the risk estimates range from 0.005 to 0.015 cancer cases per year were left unregulated. Needless to say, this case is now in the courts.

Improvements in Risk Assessment

Obviously much can and should be done to improve risk assessment. One important point is that the type of research that would be most useful to narrow uncertainties about risk is the more mechanistically oriented research which will permit a basic understanding of the interaction of chemical and physical agents with human beings. Unfortunately, many regulators, congressional staff and OMB staff have yet to learn that lesson. They often insist on highly applied studies which produce paltry, albeit predictable, increases in our knowledge of risk. A welcome side benefit of the focus on risk assessment has been an increase in awareness on the part of regulators and other holders of the research purse of the power of mechanistically-oriented research to improve risk assessment. This welcome trend requires reinforcement by scientists, not only through successful research efforts, but also through a willingness to communicate findings in terms pertinent to regulatory decision-making. We also need to continually remind regulators, who tend to be forgetful, of the value added to regulatory decisions by these scientific advances.

Goldstein

A particular area of potential major value in the improvement of risk assessment is that of so-called molecular epidemiology. By building on current rapid advances in molecular biology it should be possible in the next decade to develop new markers of exposure and of hazard which will be powerful tools for the assessment of risk. This topic will be discussed in greater detail by others at this meeting.

Improvements in risk assessment will also come about through advances in risk science itself. Better mathematical models can be anticipated for each aspect of the risk assessment process. Rigorous attempts to validate new and existing models are required.

Perhaps the most important improvement required for risk assessment to be a useful tool is the development of agreed upon guidelines for its performance. We have begun that process in EPA, in part for internal reasons related to getting our own house in order, by providing a consistent, explicit and logical approach to risk assessment to be used throughout the Agency. These Guidelines are under review by the scientific and regulatory communities through formal Federal Register notice and by the Science Advisory Board. Of note is that the Guidelines are not meant to be rigid but in fact incorporate mechanisms by which changes may be made in specific instances based upon the science (e.g. changing the method of extrapolating dose from animals to man based upon pharmacokinetics). When such changes are made, they must be done in a way that is explicit and open to comment and debate by interested parties.

Let me add, parenthetically, that the Risk Assessment Guideline for Cancer is one of only six guidelines being prepared. It is untrue that the bulk of EPA's health-related R&D program is in cancer research: in fact cancer research is less than 20% of the total.

Goldstein

10

The proposed guidelines are intended solely for EPA. However, I strongly believe that it is necessary for all federal and state regulatory agencies to work toward common Risk Assessment Guidelines. It is confusing and misleading to have CDC, FDA and EPA use different dioxin risk numbers for reasons such as employing different methods to straighten the dose response curve for dioxin-induced rat tumors. Given the extent of uncertainty in extrapolation of cancer incidence from rat to man, it represents scientific arrogance for any of these agencies to believe only they know the right way to do it.

Conclusion

We as a society must come to grips with living in a chemical era. Continued improvement in the techniques of analytical chemistry, the vanishing zero, inevitably means that the public will become aware of chemical pollutants in more and more locations. A major challenge facing those of us involved in environmental health sciences is to be certain that we communicate clearly about risk to the public, and to societal decision makers. Success in every other aspect of our science will be negated by a failure to communicate.

Is risk assessment a passing vogue, or will it continue to increase in use as a basis for regulatory decision-making in environmental matters? When appointed to EPA in the Fall of 1983, I developed a talk on Risk Assessment which ended with a comment to the effect that it was a chancy matter as to whether in the 1990's we would wonder why it took so long to establish the value of this very useful tool, or we would vaguely remember that risk assessment had been tried a decade before but hadn't worked. I no longer make this statement. I believe that in the past 18 months it has become more and more clear that although much additional work is needed, risk assessment will be a routine part of our society's approach to the regulation of hazardous chemicals.

Goldstein

REFERENCES

Committee on the Institutional Means for Assessment of Risks to Public Health.

Commission on Life Sciences. National Research Council. (1983) Risk Assessment in the Federal Government: Managing the Process.

Goldstein, B.D. (1985) Risk assessment and risk management of benzene by the Environmental Protection Agency. Banbury Report 19:293-304.

Chapter 3
Epidemiology and Risk Assessment
Arthur C. Upton

Epidemiology and Risk Assessment

Arthur C. Upton

Introduction

Epidemiology has a uniquely important role to play in each element of the process of risk assessment; hazard identification, dose-response assessment, exposure assessment, and risk characterization (Figure 1); however, in assessing the risks of low-level exposure to mutagens, carcinogens, and teratogens, for which the dose-effect relationships are generally assumed to have no threshold, the challenges to epidemiology are formidable. The following report surveys salient applications of epidemiology to quantitative risk assessment and discusses some of the problems and pitfalls in the process.

Hazard identification

Epidemiological studies have traditionally been a principal means of identifying causes of human disease. In spite of the growing importance of ancillary laboratory tests and toxicological surrogates in risk assessment, epidemiological studies continue to be indispensable for detecting, validating, quantifying, and monitoring the influences of risk factors on human health.

Among contemporary challenges to epidemiology is the need to explain the striking temporal, geographic, and ethnic differences in the incidence of cancer and other common diseases which have been observed in recent decades. The decreased incidence of diphtheria, tuberculosis and other formerly prevalent infectious diseases is, of course, attributable to advances in nutrition, sanitation, public health, and medicine during the past century, but concomitant changes in the incidence of cancer and heart disease remain to be fully explained (Figure 2).

Preparation of this report was supported in part by Grants ES00260 and CA13342 from the U.S. Public Health Service and Grant 8-0248-302 from the American Cancer Society.

For most types of cancer, the age-adjusted incidence in the U.S. population appears to have remained relatively constant throughout recent decades (Figure 3); however, the incidence of stomach cancer has fallen precipitously during the same period, while the incidence of lung cancer has risen equally dramatically (Figure 3). The basis for the decline in the incidence of stomach cancer is uncertain, but the increase in the incidence of lung cancer is attributable largely to cigarette smoking, which increased in popularity some 20-30 years before the rise in the incidence of lung cancer in each sex (Figure 4). As might be inferred from the relevant time curves, cancer of the lung has now replaced cancer of the breast as the most common fatal malignancy in U.S. women.

Fully as striking as the temporal changes mentioned above are geographic and cultural variables in the incidence of different cancers (Table 1). Although these variations are yet to be explained, it may be inferred that they are not inherited since they do not persist after migration; migrant populations tend to lose their native patterns of susceptibility and take on those typical for the region to which they migrate (Haenzel, 1975). Large differences in cancer rates are also seen among certain ethnic groups; Mormons (Enstrom, 1980; Lyon, Gardner and West 1980) and Seventh Day Adventists (Phillips, Kuzma and Lotz 1980).

The above differences, although well documented, are yet to be explained in detail. Differences in smoking habits and in alcohol consumption are undoubtedly implicated in some of the observed variations, but the identity and relative importance of specific nutritional, dietary, and other risk factors remain to be determined (Doll and Peto 1981).

Dose-response assessment

Threshold versus nonthreshold responses. Until relatively recently, toxicological risk assessments were based invariably on the assumption of a no-effect level (NOEL). The applicability of this concept to mutagenic effects came to be questioned, however, by the middle of this century (Muller, 1954). Since then, the applicability of the concept to carcinogenic effects also has been challenged (Hoel, Kaplan and Anderson, 1983; Peto, 1977). While, in principle, the existence of a threshold for carcinogenesis can neither be proven nor disproven, and thus the argument for or against it must rest on theoretical as well as empirical grounds (Scherer and Emmelot, 1981), several lines of epidemiological evidence are consistent with the nonthreshold hypothesis.

The elevated risk of lung cancer in underground hard-rock miners, noted centuries ago by Agricola, has since been observed to vary in magnitude with the cumulative dose of radiation to the respiratory tract from inhaled radon and its radioactive decay products (Figure 5). Although the data are not precise enough to define the dose-incidence relationship unambiguously, the curves for the populations of miners studied to date are consistent with a linear nonthreshold dose-incidence relationship in the low-to-intermediate dose range, as are those for laboratory animals similarly exposed to alpha emitters (United Nations Scientific Committee on the Effects of Atomic Radiation, 1977; National Academy of Sciences, 1980).

Among other occupational groups in which the incidence of cancer has been observed to vary with the duration of occupational exposure are chemists who were employed many years ago as distillers of 2-naphthylamine. In one group of such workers, the cumulative incidence of cancer of the urinary bladder approached 100 percent among those with more than five years of occupational

exposure (Figure 6). For those with shorter exposures, the curve rises steeply with the duration of exposure, showing no suggestion of a threshold (Figure 7).

Additional support for the nonthreshold model has come from observations on other populations. For example, the incidence of breast cancer appears to be increased by the same magnitude for a given dose of radiation in women exposed to atomic bomb radiation, women given x-ray therapy to the breast for acute post-partum mastitis, women subjected to repeated fluoroscopic examinations of the chest during treatment of pulmonary tuberculosis with artificial pneumothorax, and women exposed occupationally to external gamma radiation in the painting of radium clock and instrument dials (Rall, Beebe, Hoel et al., 1985). The similarity among these four groups in the magnitude of the excess for a given dose, irrespective of their marked differences in duration of exposure, argues that successive increments of dose are fully additive in carcinogenic effects on the female breast. Such complete additivity is consistent with a linear nonthreshold dose-incidence relationship, as are the shapes of the dose-incidence curves for the individual groups (Figure 8).

Perhaps the most familiar dose-incidence curve, and one which also provides no support for the threshold hypothesis, is the curve relating the incidence of lung cancer in regular smokers to the number of cigarettes smoked per day (Figure 9). This curve implies that even passive inhalation of cigarette smoke by nonsmokers, if sufficiently marked and prolonged, may increase the risk of lung cancer (Sandler, Evernson and Wilcox, 1985).

Multi-causal, multistage models. The dose-incidence relationship depicted in Figure 8 cannot be interpreted without considering the influence of age, sex, duration of smoking, and other variables. Although the incidence rises more

steeply with age in smokers than in nonsmokers, the curves for the two groups become parallel if adjusted appropriately for the duration of smoking (Figure 10). Under such conditions, both curves increase in parallel as a power function of age, and the extent to which the curve for smokers is displaced upward varies in relation to the number of cigarettes smoked per day (Figure 11).

This type of relationship, which is typical for neoplasia in laboratory animals exposed daily throughout life to chemical carcinogens, can be represented by the expressions:

$$t^n = c/d \quad (1)$$

where the median time (age) of tumor appearance t varies inversely with the daily dose d , the exponent n is a constant equal to (or greater than) 1.0 and c is a constant (Albert and Altshuler, 1973). From the slopes of the curves in Figure 11, it may be inferred that development of lung cancer involves four or more discrete carcinogenic steps (Doll, 1978).

Although the precise number and molecular nature of the successive steps in carcinogenesis remain to be disclosed, the risk of lung cancer in ex-smokers stops rising relatively promptly after cessation of smoking (Figure 12), suggesting that cigarette smoke acts predominantly as a late-stage carcinogen, or promoting agent. In this respect, it differs profoundly from atomic bomb radiation, the carcinogenic effects of which may not be manifest for decades (National Academy of Sciences, 1980; Rall, Beebe, Hoel et al. 1985).

Similarly, the multiplicative interaction between asbestos and cigarette smoke implies that these two agents exert their carcinogenic effects on the lung through different mechanisms, since their combined effects would presumably be no more than additive if both acted in the same way (Table 2).

The time-distribution of carcinogenic effects also varies with the type of neoplasm in question, as well as with age at exposure. For example, the risk of radiation-induced leukemia in atomic bomb survivors was distributed in a wave-like pattern which varied both with the type of leukemia and age at the time of irradiation (Figure 13). With chronic granulocytic leukemia, the risk increased to a peak within ten years after irradiation, irrespective of age at exposure, subsequently returning essentially to normal, whereas the risk of acute leukemia varied markedly in time of onset and duration, depending on age at exposure (Figure 13).

In contrast to the wave-like time-distribution of radiation-induced leukemia, most other radiation-induced cancers in atomic bomb survivors have increased in frequency with age, in parallel with the corresponding baseline cancer rates (Figure 14). These cancers include, among others, carcinomas of the breast in women who were irradiated in childhood, which did not appear until 30-40 years after irradiation when the women attained the age at which breast cancer ordinarily appears in the general population. The findings imply that irradiation before puberty merely initiates carcinogenesis in the female breast, expression of which requires subsequent promoting effects of age-dependent endocrine stimulation. The age-dependent increases in risks of other types of radiation-induced cancers also imply multistage mechanisms, which are still poorly understood. Whatever their explanations, time-age-dependent differences in the dose-incidence relationship must be considered in modeling the effects of carcinogens for purposes of risk assessment (National Academy of Sciences, 1980; Rall, Beebe, Hoel et al., 1985).

Because of the complexity of the dose-incidence relationship, the BEIR Committee could not confidently agree on a single dose-incidence model for

assessing the risks of low-level irradiation, preferring instead to present risk estimates for several different extrapolation models (Figure 15). While none of the models assume a threshold, although the data do not rigorously exclude one, they yield widely differing estimates of risks at low doses.

Without human data, to provide anchor points, risk estimates based on extrapolation from laboratory animal data can vary enormously, depending on the dose-response model (Figure 16) and interspecies scaling factor that are used (Table 3). Hence, the importance of epidemiological data cannot be overemphasized.

Sensitive subpopulations. In modeling risks, hereditary differences in susceptibility should obviously be taken into account, although these differences are as yet poorly known for the most part. One of the most striking examples of inherited susceptibility is the nevoid basal cell carcinoma syndrome which predisposes to radiation-induced skin cancer. In individuals with this trait, radiation therapy may be followed within months by the development of multiple basal cell carcinomas at the edges of the treatment field (Strong, 1977). In familial retinoblastoma, susceptibility to radiation-induced osteosarcoma is increased (Knudson, 1977). Traits involving defective repair of DNA are associated with heightened susceptibility to ultraviolet light and, perhaps, other carcinogens (Knudson, 1983). In general, although the molecular mechanisms of genetically determined susceptibility are not well defined, anti-oncogenes as well as oncogenes have been implicated in certain neoplasms (Table 4). It is likely, therefore, that the determinants affecting susceptibility will ultimately prove to be many and complex.

Exposure assessment

Assessment of the risk posed by an environmental chemical is commonly limited by inadequacy of the relevant information about human exposure. If, for example, the atomic bomb survivors of Hiroshima and Nagasaki had not been stratified on the basis of distance from ground-zero, or radiation dose, the overall excess of cancer other than leukemia in these populations would have been too small to be detected (Table 5).

Clearly, there is need for more systematic collection and recording of information about human exposure to environmental risk factors, along with development of the relevant data bases and record-linkage systems to make the information accessible. Also needed are further development, validation, and exploitation of chemical and biological markers of exposure (Weinstein and Perera, 1982). The growing availability and promise of markers of exposure, markers of susceptibility, and preclinical indices of disease call for the traditional tools of the epidemiologist to be integrated with, and complemented by, those of the laboratory scientist and the clinician.

Risk characterization

In addition to their importance in assessing carcinogenic risks to populations (Tables 2 and 6), epidemiological methods have been utilized for assessing attributable risks to individuals in compensation cases involving exposure to asbestos (Chase, Kotin, Grupp et al., 1985) and radiation (Rall, Beebe, Hoel et al., 1985). Susceptibility varies sufficiently among individuals, however, so that estimates of the attributable risk in any given case of cancer are highly uncertain. Nevertheless, since this application to risk characterization has been mandated by Congress for use with cases involving radiation (Rall, Beebe, Hoel et al., 1985), it is likely to demand wider attention in the foreseeable future.

References

- Albert, R.E. and Altshuler, R. (1973) Considerations relating to the formulation of limits for unavoidable population exposures to environmental carcinogens, In: Radionuclide Carcinogenesis Proceedings of the 12th Annual Hanford Biology Symposium, May 10-12, 1972 (eds. C.L. Sanders, et al. U.S.A.E.C. (A.E.C. Symposium Series, 29 CON 7205), Washington, pp 233-53.
- American Cancer Society (1985) Cancer Facts and Figures. American Cancer Society, New York.
- Boice, J.D., Jr., Land, C.E., Shore, R.E., Norman, J.E., and Tokunaga, M. (1979) Risk of breast cancer following low-dose exposure. Radiology 131:589-97.
- Cairns, J. (1975) The cancer problem. Sci. Sm. 233:64-78.
- Chase, G.R., Kotin, P., Grumt, K., and Mitchell, R.S. (1985) Evaluation for compensation of asbestos-exposed individuals. J. Occup. Med. 27:189-98.
- Doll, R. (1970) Cancer and aging: The epidemiologic evidence. Oncology 5:1-28.
- Doll, R. (1977) Strategy for detection of cancer hazards to man. Nature 265: 589-96.
- Doll, R. (1978) An epidemiological perspective of the biology of cancer. Cancer Res. 38:3573-83.
- Doll, R. and Peto, R. (1981) The causes of cancer: quantitative estimates of avoidable risk of cancer in the United States today. J. Nat. Cancer Inst. 66:1192-308.
- Enstrom, J.E. (1980) Health and dietary practices and cancer mortality among California Mormons. In: Cancer Incidence in Defined Populations. Banbury Report 4, (eds. J. Cairns, J.L. Lyon and M. Skomick), Cold Spring Harbor Laboratory, Cold Spring Harbor, pp. 69-92.

- Haenzel, W. (1975) Migrant studies. In: Persons at High Risk of Cancer, (ed. J. Fraumeni, Jr.) Academic Press, New York, pp. 361-71.
- Hammond, E.C., Selikoff, I.J. and Seidman, H. (1979) Asbestos exposure, cigarette smoking, and death rates. Ann. N.Y. Acad. Sci. 330:473-90.
- Hoel, D.G., Kaplan, N.L., and Anderson, M.W. (1983) Implication of nonlinear kinetics on risk estimation in carcinogenesis. Science 219:1023-37.
- Ichimaru, M. and Ishimaru, T. (1975) A review of thirty years of Hiroshima and Nagasaki atomic bomb survivors. II. Biological Effects. D. Leukemia and related disorders. J. Radiat. Res. 16:89-96.
- Jablon, S. and Bailar, J. (1980) The contribution of ionizing radiation to cancer mortality in the United States. Prev. Med. 8:219-26.
- Jacobi, W. (In press) Carcinogenic effects of radiation on the human respiratory tract. In Radiation Carcinogenesis (eds. A.C. Upton, R.E. Albert, F. Burns and R.E. Shore. Elsevier, New York, 1986.
- Kato, H. and Schull, W.J. (1982) Studies of the mortality of atomic bomb survivors. 7. Mortality. 1958-1978. Part 1. Cancer mortality. Radiat. Res. 90:395-432.
- Knudson, A.G. (1977) Genetic and environmental interaction in the origin of human cancer. In: Genetics of Human Cancer, (eds. J.J. Mulvihill, R.W. Miller and J.F. Fraumeni) Raven Press, New York, pp. 391-99.
- Knudson, A.G. (1983) Model hereditary cancers of man. Prog. Nucleic Acid Res. Mol. Biol. 29:17-25.
- Knudson, A.G. (1985) Hereditary cancer, oncogenes, and antioncogenes. Cancer Res. 45:1437-43.
- Krewski, D. and Van Ryzin, J. (1981) Dose response models for quantal response toxicity data. In: Statistics and Related Topics, (eds. J. Sxorgo, D. Dawson, J.N.K. Rao, and E. Salen) North Holland, New York, pp. 201-31.

- Lyon, J.L., Gardner, J.W., and West, D.W. (1980) Cancer risk and life style: cancer among Mormons from 1967 to 1975. In: Cancer Incidence in Defined Populations, Banbury Report 4, (eds. J. Cairns, J.P. Lyon and M. Skolnick) Cold Spring Harbor Laboratory, Cold Spring Harbor, pp. 3-28.
- Muller, H.J. (1954) The manner of production of mutations by radiation. In: Radiation Biology, Vol. 1: High Energy Radiation, (ed. A. Hollaender) McGraw-Hill, New York, pp. 475-626.
- National Academy of Sciences (1978) Saccharin: Technical Assessment of Risks and Benefits. Part 1 of a 2-Part Study of the Committee for a Study on Saccharin and Food Safety Policy. Panel I: Saccharin and Its Impurities. Assembly of Life Sciences/National Research Council and the Institute of Medicine. National Academy of Sciences, Washington.
- National Academy of Sciences (1980) The Effects of Populations of Exposure to Low Levels of Ionizing Radiation. Advisory Committee on the Biological Effects of Ionizing Radiation (BEIR), National Academy of Science/National Research Council, Washington.
- National Academy of Sciences (1983): Risk Assessment in the Federal Government: Managing the Process. Committee on the Institutional Means for Assessment of Risks to Public Health. National Academy of Sciences Research Council. National Academy Press, Washington, D.C.
- Peto, R. (1977) Epidemiology, multistage models, and short-term mutagenicity tests. In: Origins of Human Cancer, (eds. H.H. Hiatt, J.D. Watson, and J.A. Winsten) Cold Spring Harbor Laboratory, Cold Spring Harbor, pp 1403-28.
- Phillips, R.L., Kuzman, J.W. and Lotz, T.M. (1980) Cancer mortality among comparable members versus nonmembers of the Seventh-day Adventist Church. In: Cancer Incidence in Defined Populations, Banbury Report 4, (eds. J. Cairns, J.P. Lyon and M. Skolnick) Cold Spring Harbor Laboratory, Cold Spring Harbor, pp. 93-102.

Upton

- Rall, J.E., Beebe, G.W., Hoel, D.G., Jablon, S., Land, C.E., Mygaard, O.F., Upton, A.C., Yalow, R.S., and Zeve, V.H. (1985) Report of the National Institutes of Health Working Group to Develop Radioepidemiological Tables, NIH Pub. No. 85-2748, U.S. Government Printing Office, Washington.
- Saffiotti, U. (1973) Metabolic host factors in carcinogenesis. In: Host Environment Interactions in the Etiology of Cancer in Man., (eds. R. Doll, I. Vodopija and W. David) International Agency for Research on Cancer, Lyons, France.
- Sandler, D.P., Evernson, R.B. and Wilcox, A.J. (1985) Passive smoking in adulthood and cancer risk. *Am. J. Epidemiol.* 121:37-48.
- Scherer, R. and Emmelot, P. (1981) Multihit kinetics of tumor cell formation and risk assessment of low doses of carcinogen. In: Carcinogens: Identification and Mechanisms of Action, (eds. A.C. Griffin and C.R. Schaw) Raven Press, New York, pp. 337-64.
- Schneiderman, M.A. (1979) Trends in cancer incidence and mortality in the United States. Testimony before the Subcommittee on Health and Scientific Research, Senate Committee on Human Resources, March 5.
- Strong, L.C. (1977) Theories on pathogenesis: mutation and cancer. In: Genetics of human cancer, (eds. J.J. Mulvihill, R.W. Miller and J.F. Fraumeni) Raven Press, New York, pp. 401-15.
- United Nations Scientific Committee on the Effects of Atomic Radiation (1977) Source and Effects of Ionizing Radiation. Report of the General Assembly, with annexes. United Nations, New York.
- Upton, A.C. (1986) Evolving perspectives on the biology and mechanisms of carcinogenesis. In: Proceedings of the Symposium on Radiation Carcinogenesis, Pergamon Press, Ltd., London, (in press).

- Wakabayashi, T., Kato, H., Ikeda, R. and Schull, W.J. (1983) Incidence of cancer in 1950-1978 based on the Tumor Registry, Nagasaki. *Radiat. Res.* 93:112-46.
- Waterhouse, J., Muir, C., Shammugaratnan, K. and Powell, J. (eds.) (1976) *Cancer Incidence in Five Continents, Vol. IV.* International Agency for Research on Cancer, Lyon, France.
- Weinstein, I.B. and Perera, F.P. (1982) Molecular cancer epidemiology. In: *Indicators of Genotoxic Exposure, Banbury Report 13,* (eds. B.A. Bridges, B.E. Butterworth and I.B. Weinstein) Cold Spring Harbor Laboratory, Cold Spring Harbor, pp. 3-15.

TABLE 1: Range of geographic variation in the incidence of common cancers.

Type of cancer	High-incidence area	Cumulative risk ^a (%)	Range of variation ^b	Low incidence area
Skin	Australia, Queensland	>20	>200	India, Bombay
Esophagus	Iran, N.E.	20	300	Nigeria
Bronchus	England	11	35	Nigeria
Stomach	Japan	11	25	Uganda
Cervix uteri	Colombia	10	15	Israel, Jewish
Liver	Mozambique	8	70	Norway
Prostate	USA, black	7	30	Japan
Breast (female)	USA, Connecticut	7	5	Uganda
Colon	USA, Connecticut	3	10	Nigeria
Buccal cavity	India, part	> 2	> 25	Denmark
Rectum	Denmark	2	20	Nigeria
Bladder	USA, Connecticut	2	4	Japan
Ovary	Denmark	2	8	Japan
Corpus uteri	USA, Connecticut	2	10	Japan
Nasopharynx	Singapore, Chinese	2	40	England
Pancreas	New Zealand, Maori	2	5	Uganda
Penis	Uganda, part	1	300	Israel, Jewish

Source: Doll (1977)

^a By 75 years of age in absence of other causes of death.

^b Between high-incidence area and low-incidence area, at ages 35-64 years.

TABLE 2: Relative risks of mortality from lung cancer as influenced by smoking habits and occupational exposure to asbestos.

Smoking habits	Occupational exposure to asbestos	
	Absent	Present
Nonsmoker	1.0	5.17
Regular cigarette smokersmoker	10.85	53.24

Source: Hammond, Selikoff and Seidman (1979)

TABLE 3. Estimated risk of human bladder cancer from ingestion of 0.12 G saccharin per day

Method of high-to low-dose extrapolation	Lifetime cases per million exposed	Cases per 50 million per year
Rat dose adjusted to human dose by surface area rule		
Single-hit model	1,200	840
Multi-stage model	5	3.5
Multi-hit model	0.001	0.0007
Mantel-Bryan probit model (Brown, 1978)	450	315
Rat dose adjusted to human dose by mg/kg/day equivalence		
Single-hit model	210	147
Multi-hit model	0.001	0.0007
Mantel-Bryan probit model	21	14.7
Rat dose adjusted to human dose by mg/kg/lifetime equivalence		
Single-hit model	5,200	3,640
Multi-hit model	0.001	0.0007
Mantel-Bryan probit model	4,200	2,940

Source: National Academy of Sciences (1978)

TABLE 4: Contrasting properties of oncogenes and anti-oncogenes.

Oncogenes	Anti-oncogenes
Gene active	Gene inactive
Specific translocations	Deletions or invisible mutations
Translocations not hereditary	Mutations hereditary and non-hereditary
Dominant	Recessive
Tissue specificity may be broad	Considerable tissue specificity
Especially leukemias and lymphomas	Solid tumors

Source: Knudson (1985)

TABLE 5: Estimated numbers of cancer deaths, 1950-1978, attributable to atomic bomb irradiation in Hiroshima and Nagasaki.

Cause of death	Total number of deaths ^a	Deaths attributable to a-bomb irradiation ^b	
		(No.)	(%)
Leukemia	387	191	49
Other cancers	10,421	336	3
All cancers	10,808	527	5
All causes	67,660	527	0.8

Source: Kato and Schull (1982).

^a Based on 110.8 million person-years for 283,498 atomic bomb survivors identified in 1950 census.

^b Based on assumption of linear increase in cancer incidence with increasing dose.

TABLE 6: Estimates of cancer burden attributable to occupational exposure in various types of work involving ionizing radiation, U.S., 1975.

Type of work	No. of persons exposed annually	Average annual dose per person (mrem)	Annual collective dose (thousands of person-rem)	Lifetime cancer mortality commitment (fatal cancers)
Healing arts	500,000	80-160	40-80	10-20
Manufacturing and industrial	7,000,000	7	50	12.5
Nuclear energy	62,000	840	52	13
Research	100,000	120	12	3
Naval reactors	36,000	220	8	2

Source: Jablon and Bailar (1980).

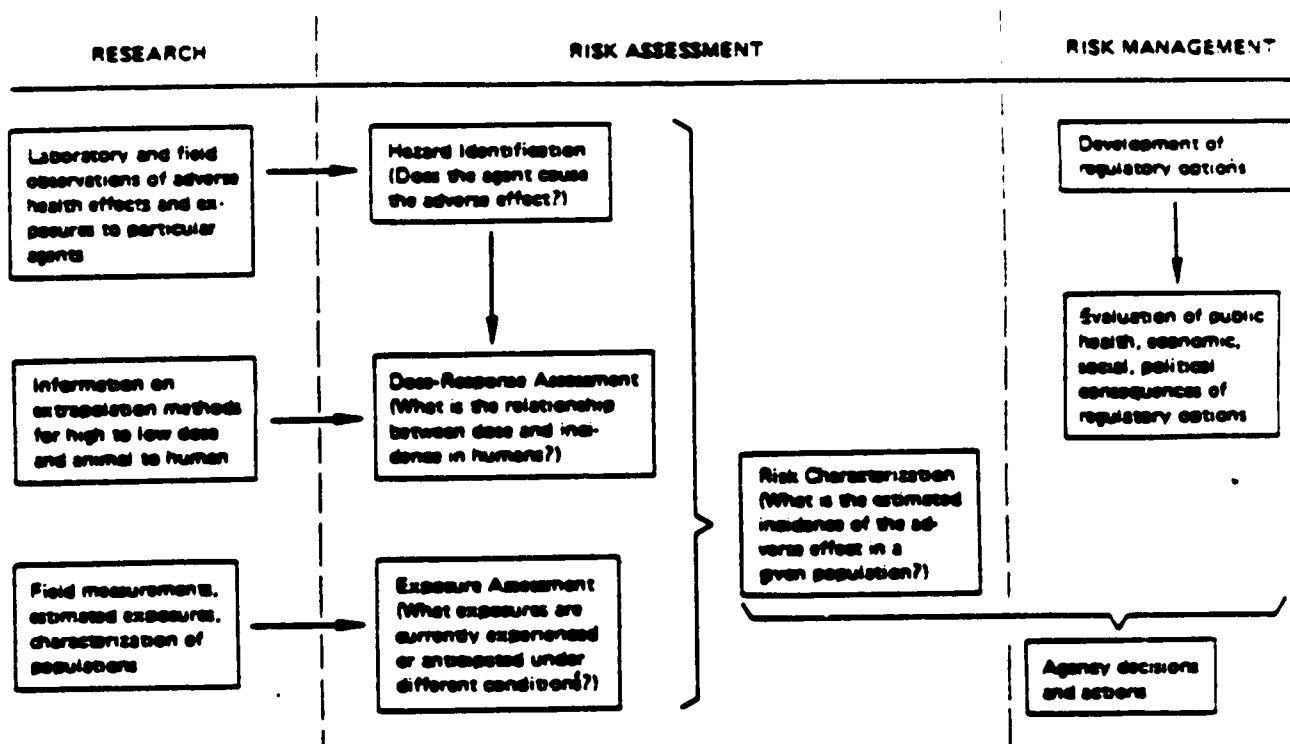


FIG. 1. Major elements of risk assessment and risk management.

Source: National Academy of Sciences (1983).

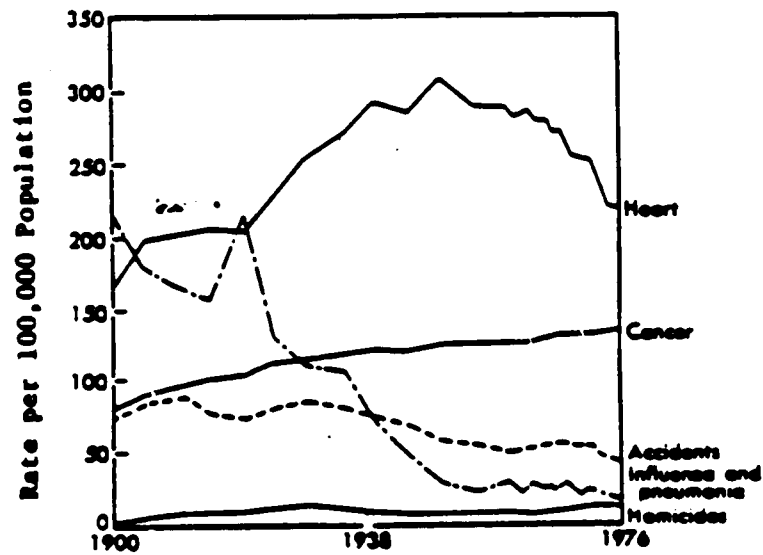


FIG. 2: Trends in age-standardized mortality rates for major causes of deaths in the U.S.

Source: Schneiderman (1979).

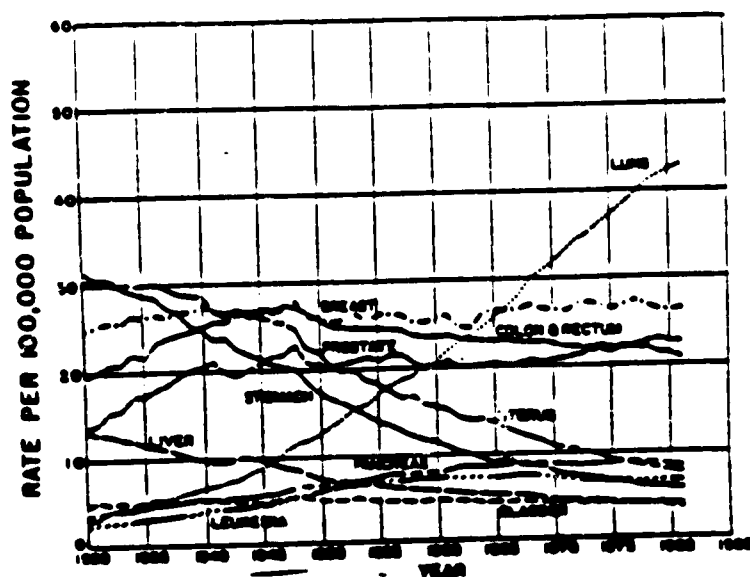


FIG. 3: Trends in death rates* for cancers of the commonest types in the U.S., 1930-1981.

Source: American Cancer Society (1985).

*Rates for the population standardized for age on the 1970 U.S. population.

Sources of Data: National Center for Health Statistics and Bureau of the Census, United States.

Note: Rates are for both sexes combined, except breast and uterus female population only and prostate male population only.

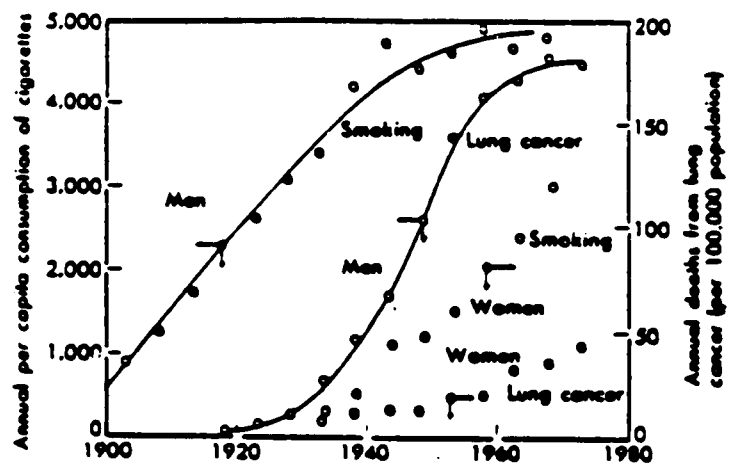


FIG. 4: Time trends in mortality from lung cancer in the U.S. in relation to cigarette consumption.

Source: Cairns (1975).

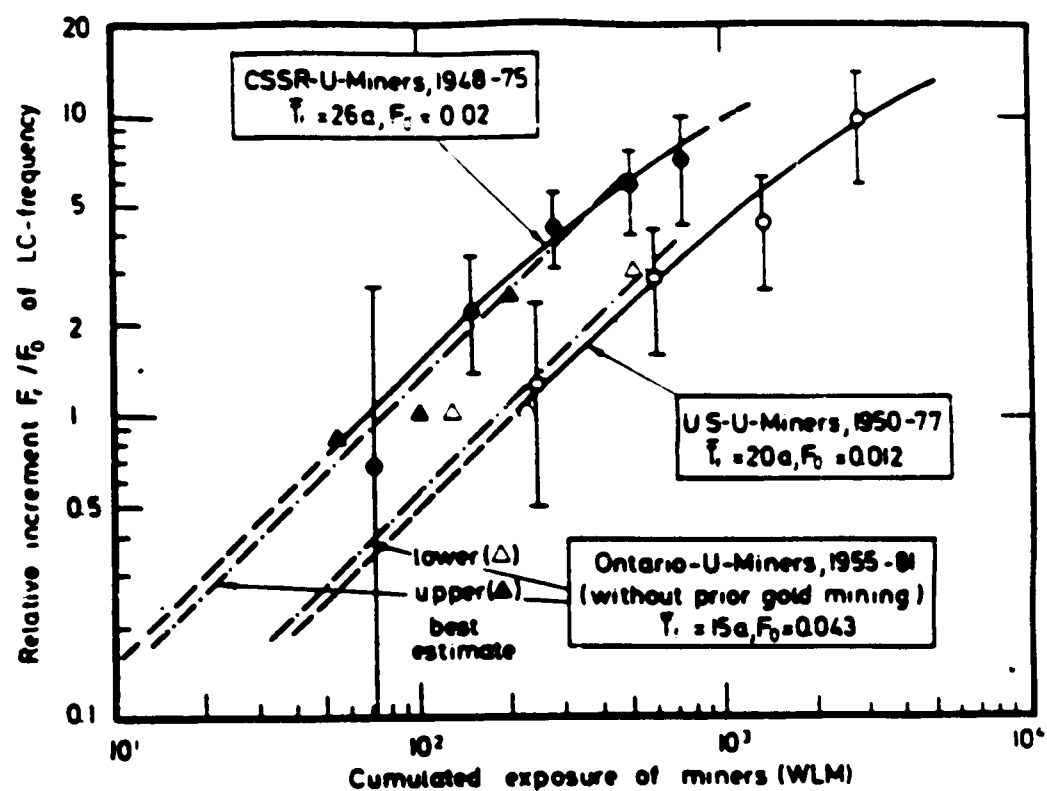


FIG. 5: Increase in relative risk of lung cancer in uranium miners in relation to cumulative exposure to radon daughters (with 95% confidence limits).

Source: Jacobi (1986).

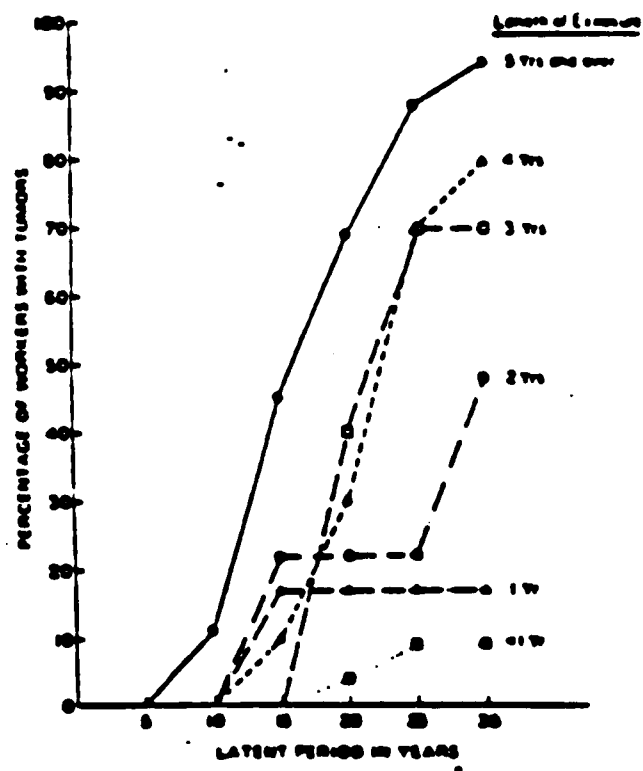


FIG. 6. Latency and cumulative incidence of tumors of the urinary bladder in 78 distillers of 2-naphthylamine, in relation to duration of occupational exposure.

Source: Saffiotti (1973).

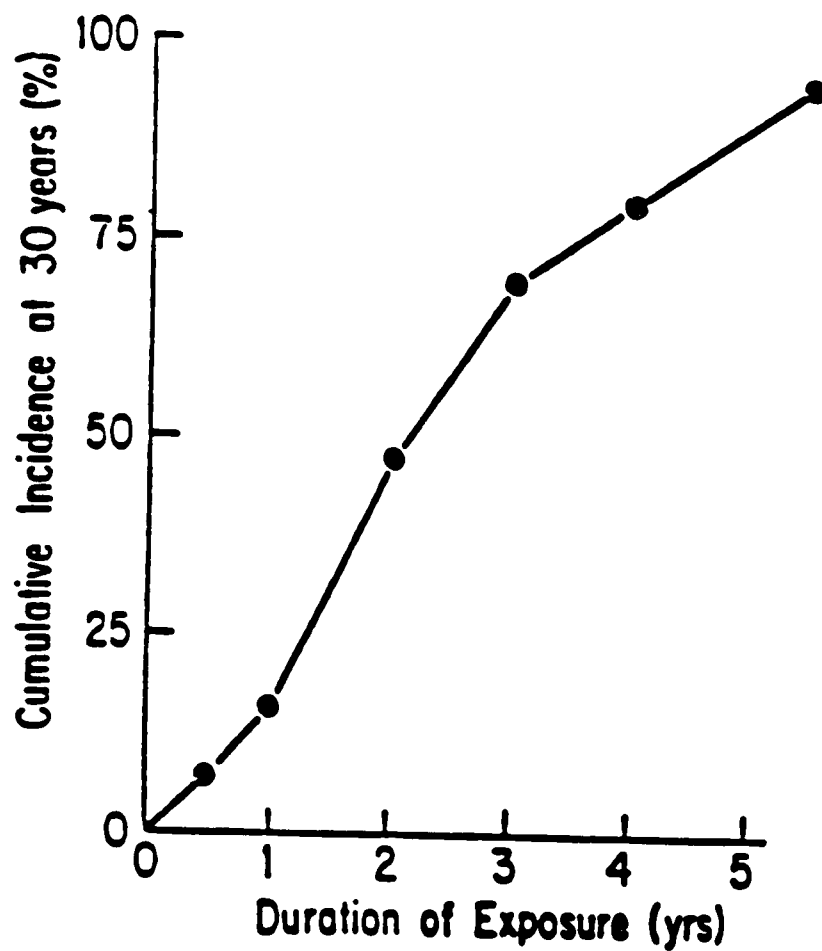


FIG. 7. Cumulative incidence of tumors of the urinary bladder in distillers of 2-naphthylamine, in relation to duration of occupational exposure (data from Figure 5).

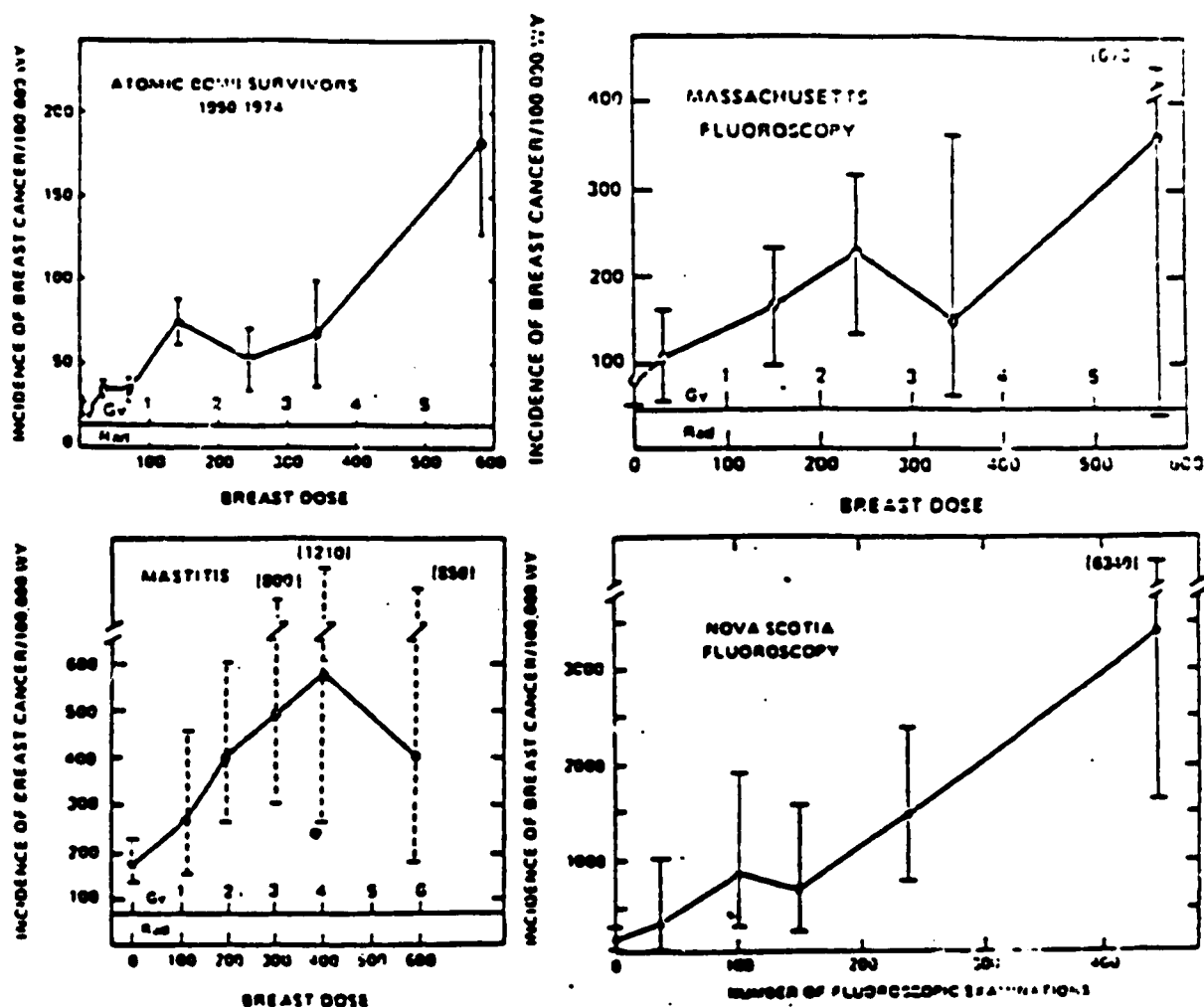


FIG. 8. Incidence of cancer of the female breast as a function of radiation dose in atomic bomb survivors, women treated with x-rays for acute post-partum mastitis, and women subjected to multiple fluoroscopic examinations of the chest during treatment of pulmonary tuberculosis with artificial pneumothorax.

Source: Land, Shore, Boice et al., (1979).

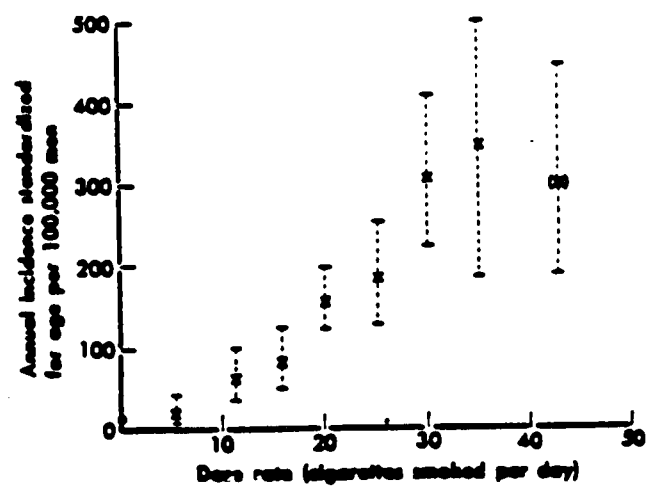


FIG. 9. Incidence of lung cancer in regular cigarette smokers, in relation to number of cigarettes smoked per day.

Source: Doll (1978).

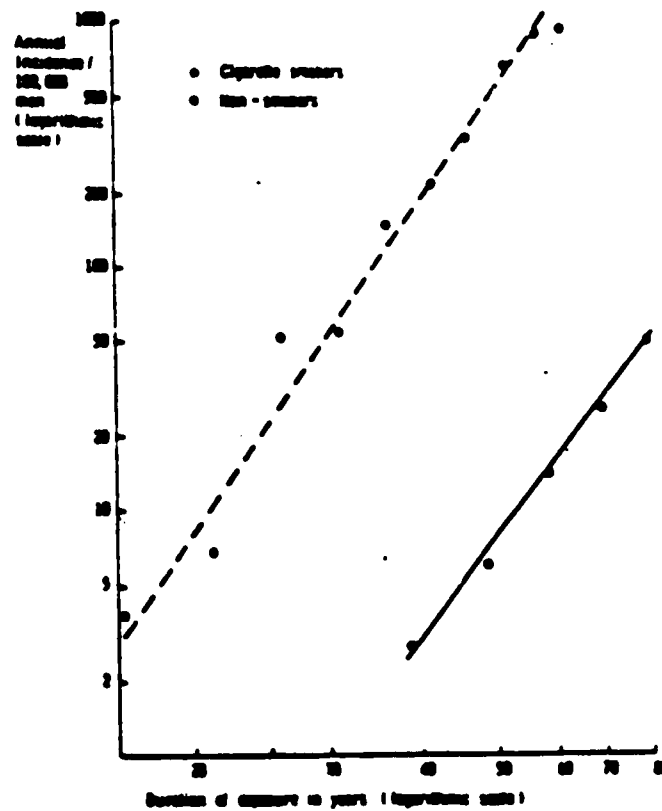


FIG. 10. Incidence of bronchial carcinoma in relation to the duration of exposure to relevant carcinogen in nonsmokers and cigarette smokers.

Source: Doll (1970).

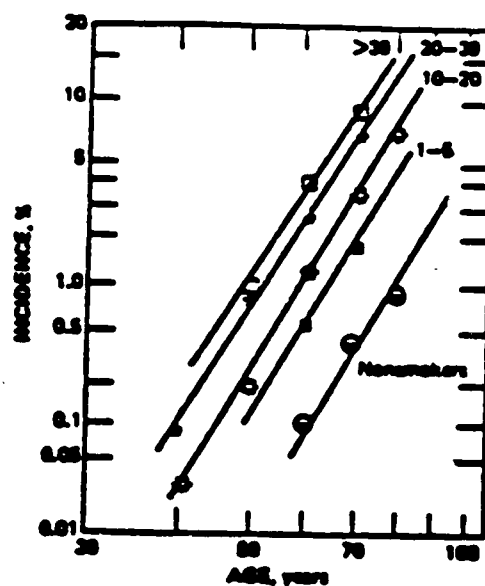


FIG. 11. Cumulative single-risk incidence of lung cancer in cigarette smokers and nonsmokers as a function of age. The number of cigarettes smoked per day is shown at the top of each curve.

Source: Albert and Altshuler (1973).

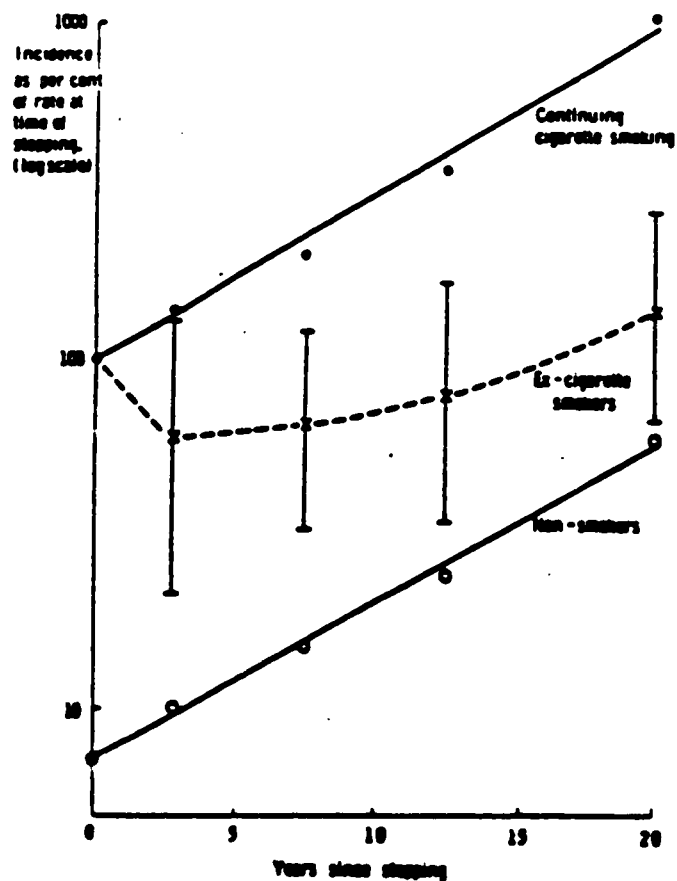


FIG. 12. Incidence of bronchial carcinomas in relation to time after cessation of cigarette smoking, as compared with that in continuing smokers and nonsmokers.

Source: Doll (1970).

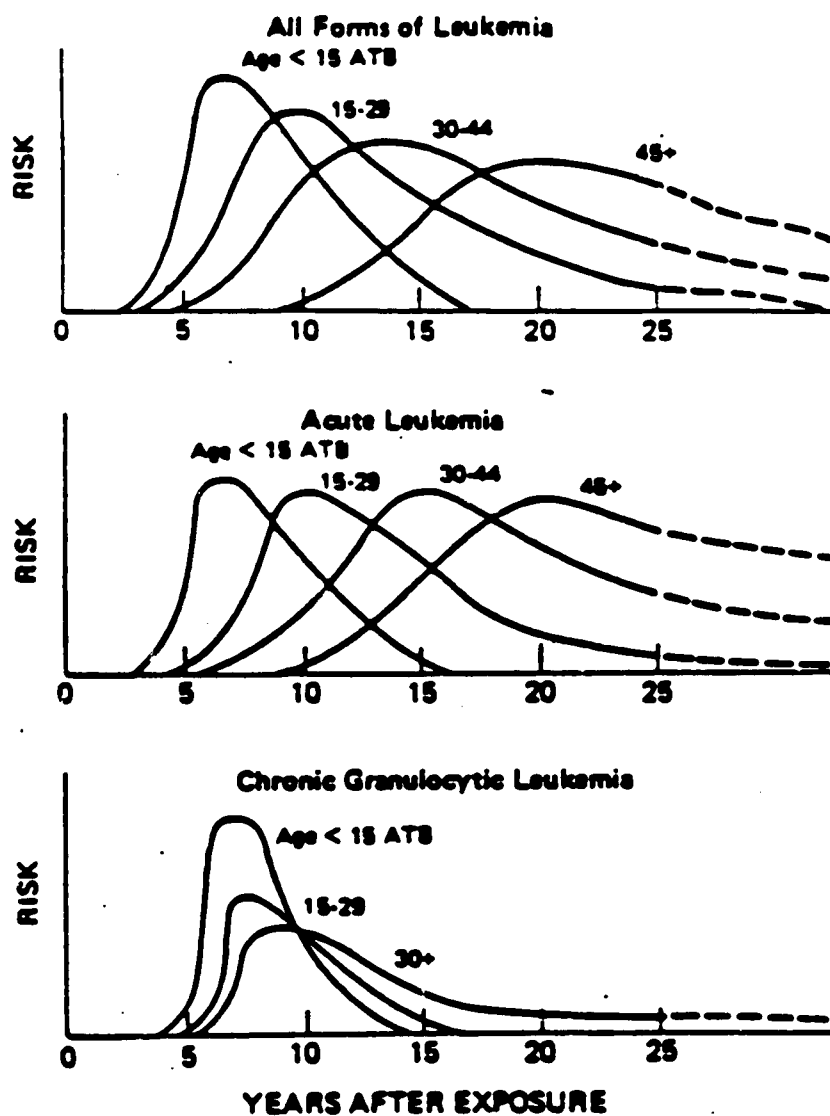


FIG. 13. Temporal distribution of radiation-induced leukemias in atomic bomb survivors, in relation to type of leukemia and age at the time of irradiation.

Source: Ichimaru and Ishimaru (1975).

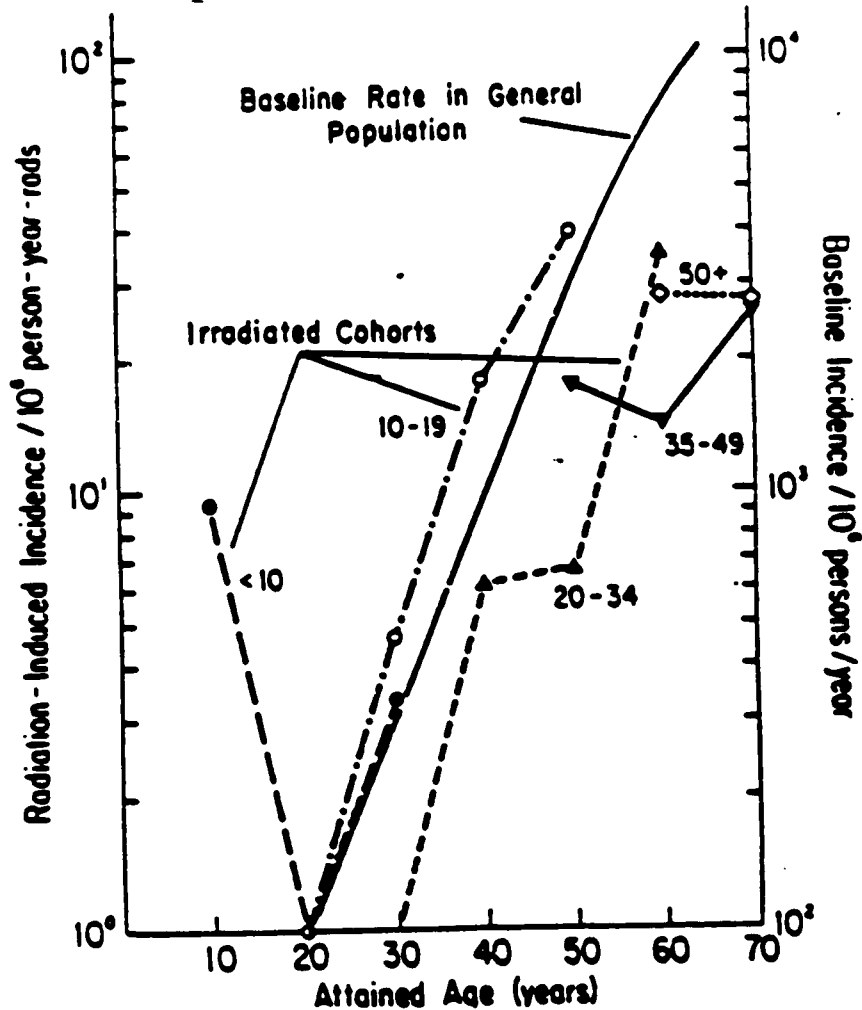


FIG. 14. Incidence of radiation-induced cancer (excluding leukemia) in relation to attained age in different birth cohorts of heavily-irradiated Nagasaki atomic bomb survivors, as compared with the baseline incidence in the general population of Japan. Data for atomic bomb survivors (Wakabayashi, Kato, Ikeda et al., 1983) denoted by dashed curves, with age of each cohort at time of irradiation as indicated. Data for the general population of Japan (Waterhouse, Muir, Shammugaratnan et al. 1976) denoted by solid curve.

Source: Upton (1986).

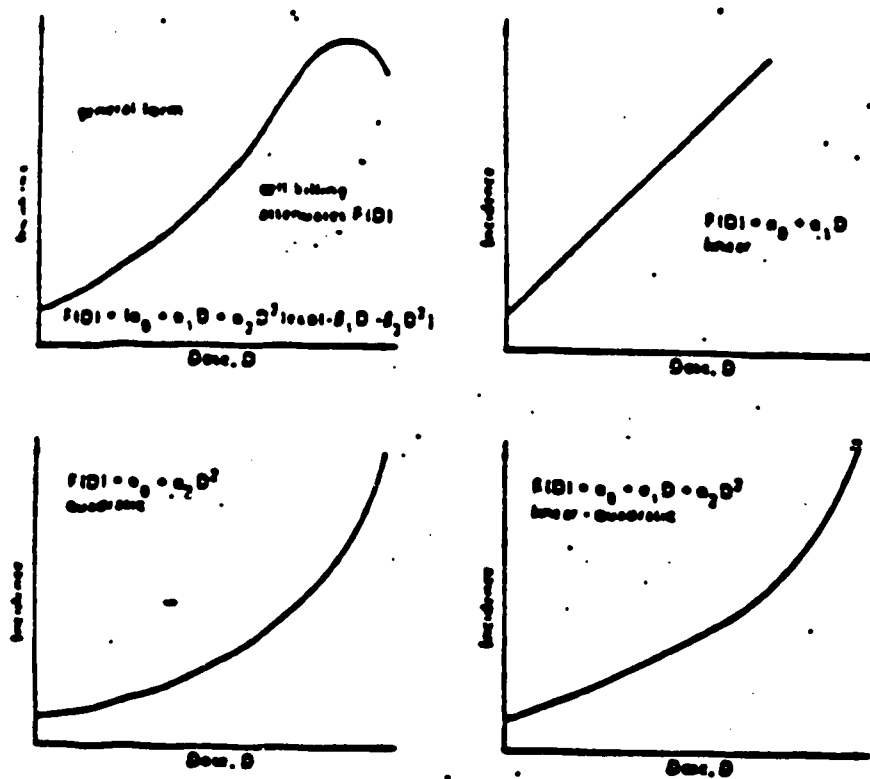


FIG. 15. Dose-response curves for four different mathematical models relating cancer incidence to radiation dose.

Source: National Academy of Sciences (1980).

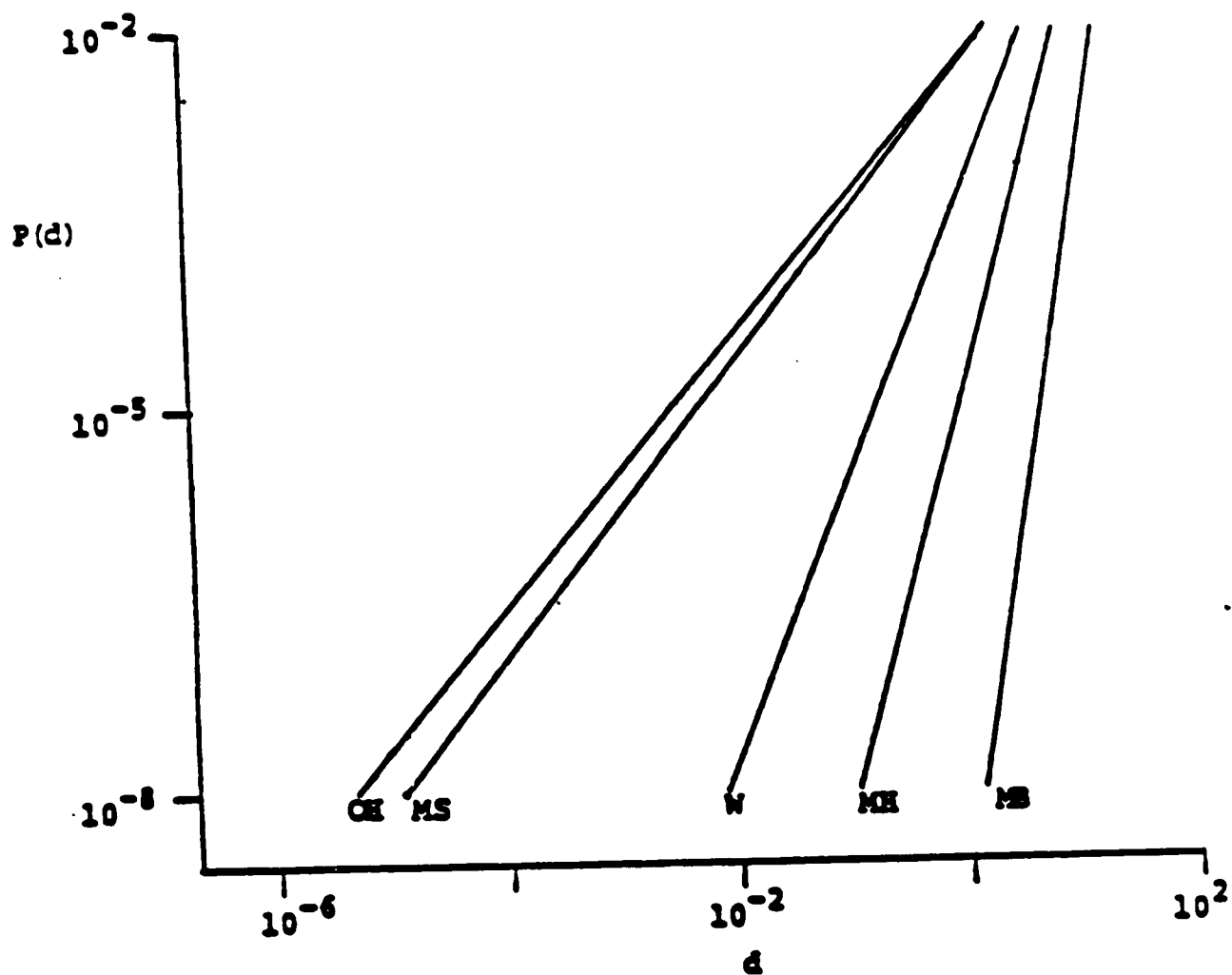


FIG. 16. Estimated risk of liver cancer, $P(d)$, in relation to dose of aflatoxin, d , as determined with different dose-incidence models. The models for the different curves are as follows: OH, one-hit model; MS, multistage model; W, Weibull model; MH, multi-hit model; MB, Mantel-Bryan (log-probit model)

Source: Krewski and Van Tyzin (1981).

Chapter 4

Toxicology and Epidemiology — Strengths and Limitations

Norton Nelson

TOXICOLOGY AND EPIDEMIOLOGY - STRENGTHS AND LIMITATIONS
Norton Nelson

A contrast of some of the limitations and potentialities of toxicology and epidemiology will be presented in this chapter as well as suggestions as to how each can reinforce the other and a description of where they fall short of fully meeting the desired objective. In the present context, the desired qualitative objective is the indictment or exoneration of particular risk factors as human disease agents; a more elaborate objective is a quantitative assessment of effect generally in terms of confidence limits.

It is rare that toxicology or epidemiology alone is adequate to secure the desired refinement; almost invariably some degree of reinforcement will be required in order to define adequately the degree of association. There is, of course, an ethical consideration here which puts toxicology in a very special category. Toxicology at its best, can be used as a predictive tool which, when successful, avoids the need for waiting for an epidemiological outcome. This is especially relevant to chemicals which are potentially eligible for entering into the realm of human exposure. Obviously, if it is possible to predict dangers in advance through laboratory studies in test animals, this is vastly superior to waiting for the costly, dangerous, and needless wastage of human life.

A classic example of this was the very early demonstration in the Food and Drug Administration of the carcinogenicity of acetyl aminofluorene, which was developed as a possible pesticide (Wilson, De Eds and Cox, Jr. 1941). The identification (date) of its carcinogenicity through animal tests accomplished two things; 1) its banning as a practical pesticide and 2) the emergence of this chemical as one of the major research tools for study of experimental cancer. Both objectives have served mankind well.

The research from the Institute of Environmental Medicine reported in this paper was supported in part by Center Grant No. ES 00260 from the National Institute of Environmental Health Sciences and Center Grant No. CA 13343 from the National Cancer Institute, and Special Institutional Grant No. 00009 from the American Cancer Society, and individual awards identified in each of the referenced papers from the Institute.

The following discussion will be restricted largely to the field of

and background information than other kinds of injury from external agents.

Epidemiology can be viewed as a study of human experience under conditions as well-controlled as the circumstances permit; it thus represents the decision to examine in an orderly and carefully planned manner the experience of humans with respect to one or several potential risk factors through selected observations.

In contrast, the term toxicology is used here to define controlled laboratory experimentation. The distinction is between experimentation on the one hand and observation of experience on the other. Toxicology then is a very broad area which permits controlled experimental study of many kinds in which the variables (at least ideally) can be controlled in advance and examined by planned experimentation. Although toxicology is generally thought of as being exclusively related to the study of non-human test animals, it is not so restricted here. The essential distinction then is controlled experiment rather than partially controlled observation of experience. In this sense, toxicology or laboratory experimentation is not excluded from the possible study of humans. Obviously, there are major ethical considerations which limit such study; nevertheless, there are some circumstances in which such studies can be undertaken without violating ethical considerations or doing injury to the human subjects. Having stated this, however, the greatest bulk of controlled laboratory experiment will be done on non-human animals or isolated systems of a variety of kinds including those obtained from humans for example, as a by-product of surgical procedures. In short, the term toxicology as used

here will deal broadly with controlled experiments aimed at directly or indirectly securing information supportive of the estimation of risk from an external agent or agents.

In this sense, a contrast will be made between some of the limitations and opportunities present in the two techniques of toxicology and epidemiology with emphasis on one of the major obstacles in the use of controlled studies in animals, namely, species differences; and finally a few examples will be given of laboratory studies which in different ways have aided the refinement or verification of risk factors in humans. The theme throughout this discussion will be the balanced useage and reinforcement of laboratory studies to strengthen epidemiological observations underway, or in some cases, to avoid the necessity of prior epidemiological studies earlier carried out. Table 1 contrasts some of the limitations and opportunities in the two approaches.

Since the ultimate concern is with human injury, the relevance of tests on animals or isolated systems can be uncertain because of species differences and other restrictions. Epidemiological studies are, of course, the most directly relevant to human disease.

There is an enormous difference between the ability to control variables in planned laboratory studies as compared to epidemiological studies. In planned laboratory studies, ideally at least, control of variables is straightforward and excellent. This is in strong contrast to epidemiological studies where control of variables can be extremely difficult and uncertain. Thus, exposure can be precisely established in animal studies and only with difficulty and many shortcomings can it be measured reliably in the

observation of humans. Environmental conditions can also be often precisely established for the test animals while one must accept many variables in human studies. Confounding factors again ideally are more readily controlled in the laboratory whereas in epidemiological studies, these require extremely careful planning; complete success is rarely achievable. Identification of causal factors is straightforward in the laboratory, but often very difficult in field studies.

The size of the population for normal laboratory studies of small mammals is severely limited. A study permitting the identification of a one percent effect is extremely costly and has only been tried in very few experimental circumstances. Ideally, human populations can, and indeed have been, in a few instances, studied in very large groups. Obviously, in the studies such as in the test tube using a biochemical or microbiological system, for example, the Ames Test of cell culture, very large test populations are possible. The restriction at issue here primarily relates then to the use of various kinds of small test mammals.

Accordingly, the sensitivity of toxicology studies is routinely poor, and is also frequently poor in epidemiological studies. Under very special circumstances, the use of very large population groups can improve the sensitivity substantially.

Although, it is possible to carry out epidemiological studies with some degree of genetic characterization, this brings with it substantial difficulty. In the case of animal studies with small mammals, genetic diversity can be very carefully controlled so that the test animals represent a very narrow spectrum of genetic diversity. Indeed, it may be argued that sometimes toxicological studies using small mammals suffer from too much restriction in terms of genetic diversity.

In laboratory studies, it is possible with only rare and occasional lapses, to strictly control the presence of pre-existing disease. In epidemiological studies, this is not the case although descriptive information on prior disease can be obtained and there are, of course, techniques for excluding persons with definable illness during the course of epidemiological studies.

Biological mechanisms of action are easily examined in laboratory studies. There are no real restrictions in such studies or those using in vitro systems as surrogates for whole animals. The information so derived may, however, be of uncertain relevance to the human situation. In the case of epidemiological studies, there are, of course, major and often insuperable ethical hindrances to the studies of biological mechanisms. On the other hand, there are surrogate systems using human tissues which permit classification of mechanisms under controlled laboratory studies.

There are similarly in small animals no restrictions on the use of diagnostic tests, whether they be biopsy, sacrifice or the complete range of clinico-chemical observations. Diagnostic tests on humans in epidemiological studies, of course, are severely restricted and indeed are becoming more so as the ethical considerations receive more powerful consideration.

In respect to the sensitivity of epidemiological studies, OSHA has attempted to define what could be regarded as essentially negative study of occupational exposure to suspect carcinogens (Federal Register 1980). This would require that the period of exposure be monitored for not less than twenty years, that the follow-up be at least thirty years. If under those circumstances, the risk ratio is less than 1.5 with a population of such size as to assure that such a statistically significant risk could have been detected, then for practical purposes, this could be regarded as essentially a negative study.

Nelson

The use of surrogate systems involving in vitro biochemical studies of tissues, enzymes and tissue extracts can provide very helpful information on mechanisms. Cell cultures and micro-organisms (Ames, Durston, Yamasaki et al 1973) have achieved remarkable success in informing the experimentalist of potential mutagenicity or carcinogenicity. To some extent, they can replace the use of whole animals for such studies. However, the point where the use of whole animals can be entirely eliminated is still well ahead of us. It might parenthetically be added that comparable success in the development of short-term tests for effects other than genotoxicity has been very limited.

The limitations in the use of short-term tests have to do with two factors: 1) differences in species response, and 2) limitations arising from the removal of the test system from its normal milieu. There are, of course, a number of ingenious ways to improvise in the test tube certain parts of a large system such as S9 mixed function oxidase systems from liver extracts. Nevertheless, a cell or even an organ culture moved away from its normal location is a far cry from an organism with full interplay with other parts of a living organism.

Figure 1 illustrates a somewhat simplified mammal including the usual intake and excretory organs plus a series of functioning, some dependent organs and systems which in life are highly complicated, integrated and interrelated. One can visualize a technique for providing a series of systems, some of them based on extracts of human tissues and cells, made available as a by-product of surgical procedures, which would undertake to systematically examine such relevant factors as the interplay of transport from intake point to the target tissue, entry across membranes and detoxification activation enzymes. (Figure 2). Whether it will, in fact, be possible to devise a practicable, simplified

paradigm which would be of practical utility in selectively mimicking human metabolism remains to be seen. It is already evident that reactivity of human tissues in a limited examination of in vitro studies shows very wide differences among individuals (Eisenbud, Laurer, Rosen et al 1969).

This is not the time to discuss the improvement of exposure assessment in epidemiological studies through the use of biological monitoring in humans and similar procedures. There are, however, some new procedures which have substantial promise in respect to carcinogenic exposure through extremely sensitive techniques for measuring DNA adduct formation in leukocytes secured from circulating blood (Weinstein 1983). These procedures and others suffer from the handicap that the material of importance may be relatively evanescent so that earlier exposures will not be reflected in the measurements of exposed individuals. Long-lived isotopes which are bone seekers such as lead 210 can leave quite a durable record in individuals who have been exposed, for example, to radon daughters (Eisenbud, Laurer, Rosen et al 1969). Exposure assessment in epidemiological studies is one of the major barriers in quantitative estimates of risk.

The following is a brief review of several instances in which laboratory studies at the Institute of Environmental Medicine have served to clarify several epidemiological studies. The first example concerns radiation exposure.

Albert and co-workers (Albert, Pasternack, Shore et al 1975) have had under study for some years a group of persons who as school children were given scalp irradiation for the control of tinia capitis (infectious ringworm) of the scalp (Shore, Albert and Pasternack 1976). It was possible to identify the original machine with which the irradiation was administered. Careful records were maintained of the timing and orientation of the scalp irradiation in a sister department at New York University some years earlier. This made it

Nelson

-7-

practical to reconstruct by laboratory studies involving phantoms, precise estimates of dosage delivered to a number of tissues. In this way, the tissue dosage of ionizing radiation could be characterized for the skin, for the brain, for the bone marrow contained in the skull and indeed, for the thyroid. This made possible an estimate of the tumor incidence in respect to dose received (Table 2). (Harley, Albert, Shore et al 1976)

In this case, laboratory procedures in association with an epidemiological study permitted linking tumor incidence with received dose.

Thyroid adenomas were found in six subjects; the procedure led to some 6 rem dosage of the thyroid. In extending this analysis, Harley examined a series of other studies of human exposure of the thyroid to ionizing radiation and Table 3 illustrates a relatively uniform dose response relationship for tumor incidence of the thyroid from several different studies.

In another example at the Institute of Environmental Medicine, laboratory procedures were used to help identify the probable active agent(s) responsible for chromate-related lung cancer. The data on lung cancer occurrence from the inhalation of chromate compounds is now well-established in a substantial number of studies. However, the nature of the particular compound or compounds responsible has never been adequately identified. On this basis, a series of laboratory studies was undertaken, first by a technique which was developed permitting the implantation of test material in the segmental bronchus of a rat allowing it to remain there for the lifetime of the animals to determine whether or not bronchogenic cancer resulted. One such study explored a series of five suspect agents using this pellet implantation technique (Laskin, Kuschner and Drew 1969) and clearly identified chromic chromate and calcium chromate as positive (Table 4); other compounds failed to produce bronchogenic cancer in the exposed rats. On this basis, a lifetime inhalation study was

undertaken using calcium chromate aerosol to which rats were exposed six hours per day five days a week (Laskin, Kuschner and Drew 1969). These studies were the first to produce identifiable high incidence of bronchogenic cancer in the laboratory. In this case, laboratory studies overcame the frequent weakness of epidemiologic studies in identifying specific causal agents.

Since then, elegant biochemical and cell biology studies have been carried out with the interesting finding that the chromium ion with a valence of 6 readily penetrates cells but valence 3 does not; inside the cell, the chromium 6 is reduced to chromium 3 which appears to be much more reactive with DNA than the chromium 6 itself. This suggests that the ultimately active agent is chromium 3 which is dependent on entry as chromium 6 which is then reduced to the lower valence state (Jennete 1981).

A third example in which laboratory studies preceded and stimulated epidemiological studies is illustrated by work which was done with bischloromethyl ether (BCME). This compound is used as a chemical intermediate in the production of ion exchange resins. The actual compound used is chloromethyl methyl ether which is normally contaminated with up to six to eight percent of the bischloromethyl compound. In this instance, there was suspicion of malignancy in a plant using these materials but there was no confirmed evidence of an excess occurrence nor was any specific agent under suspicion.

VanDurren's work on structure activity relationships among organic carcinogens suggested strongly that bischloromethyl ether might well be an active carcinogen. Studies were then undertaken using skin painting and the work clearly showed that the bischloromethyl ether was a potent skin carcinogen

for the mouse (Van Duuren, Sivak, Goldschmidt et al 1969). In extension of this work, Kuschner and Laskin (1975) exposed rats to inhaled bischloromethyl ether at the very low concentration of 0.1 parts per million. This exposure produced a high incidence of bronchogenic cancer in the rat clearly indicating this particular compound as a bronchogenic carcinogen (Table 5).

An industry-wide study of workers exposed to this compound was then undertaken. While this was underway, a shrewd set of observations by Figueroua from chance remarks of several patients diagnosed with lung cancer, led to a report systematically identifying the occurrence of lung cancer in workers exposed to this compound (Figueroua, Raszkowski and Weiss 1973). Meanwhile, the cohort study, was completed. Exposure in this study was estimated prior to conduct of the epidemiological study by a survey of workers in the plants which characterized the intensity of exposure through job descriptions; it was also possible to establish the time periods over which exposure occurred.

The results illustrated in Tables 6 and 7 (Albert, Pasternack, Shore et al 1975) show the characteristic features of a chemical carcinogen in which increase of either level or duration of exposure leads to increased incidence of cancer. Thus, toxicological studies served to identify the culprit out of the multitude of possible agents that might have been guilty. Similar findings have since been reported in Japan (Sakabe 1973), in Germany (Theiss, Hey and Zeller 1973), and in the United Kingdom (McCallum, Woolley and Petrie 1983).

Subsequent work in a follow-up of these studies has supported the earlier findings and has shown that late appearing tumors are occurring in some of the workers exposed prior to the cleanup of the plants; lung cancers related to BCME increased from twenty-two to thirty-two cases (Pasternack, Wales-Collingwood and Shore 1985).

As noted, it has sometimes been possible in advance of any observed injury, to identify chemical carcinogens and to permit institution of control measures, abandonment of the chemical or identification of substitute materials in order to avoid human cancer. This is a most important application of toxicologic studies. In the former case, epidemiologic studies may be required for follow-up purposes to test the adequacy and effectiveness of controls.

It is thus evident that there are uses of toxicology as adjuncts of epidemiologic investigations. There are many instances where controlled laboratory procedures can answer questions with a refinement and expedition not possible in studies of humans. This partnership can be strongly beneficial and toxicology should be regarded as a major source of assistance in the conduct of epidemiologic studies.

TABLE 1: Toxicology and epidemiology: Contrasts of selected limitations and opportunities

Limitations and opportunities	Toxicology	Epidemiology
Relevancy	Uncertain	Excellent
Control of variables: Exposure Environment Confounding factors	Excellent	Poor
Identifying causal factors	Excellent	Poor
Size of population	Limited	Can be very large
Sensitivity	Poor	Poor
Genetic diversity	Normally deliberately narrow	Broad
Intercurrent disease	Controllable	Not controllable
Study of mechanisms	Easily accessible	Ethical hindrances
	Uncertain relevancy	Directly relevant
Diagnostic tests	Unrestricted (except interview)	Severly restricted

TABLE 2: Scalp irradiation per 1,000 children and associated incidence of tumors of the head and neck

Tumors	Irradiated Rate/1,000 ^a		Controls Rate/1,000 ^a	
	No.	per	No.	per
Neurogenic				
Brain	5		0	
Acoustic neuroma	1	7.2	0	
Neck-schwannoma	2	1.1	0	
Total	8	8.3	0	
Parotid	4	3.6	0	
Skin-epithelial				
Basal cell	11	22.6	1	2.6
Cylindroma	2	1.8	0	
Other	0		2	2.1
Total	13	24.1	3	4.7
Skin-other	10	8.1	5	5.2
Bone-skull and jaw	4	2.5	0	
Mouth-larynx papilloma	5	19.2	2	15.6
Thyroid adenoma	6	5.1	0	
Hemopoietic or lymphopoietic				
Leukemia	4	2.1	1	2.6
Hodgkin disease	4	2.8	2	4.5
Lymphosarcoma- submandibular	1	0.5	0	
Total	5	5.4	3	7.1

^aCumulative incidence for 0 to 29 elapsed years based on life-table methods to correct for number at risk.

Source: Shore, Albert and Pasternack, 1976

TABLE 3: Selected studies of tumor excision incidence and median time interval to tumor excision

Studies	Dose (rad)	Tumor excision incidence (%)	Median time interval (exposure to tumor excision) (years)	Incidence per rad (%) ^a
Hempelmann, Rochester, control	1.9	6/5005(0.12)	25.3	0.06
NYU-tinea capitis	6	6/2215(0.27)		0.05
Adenoma		6	20	
Carcinoma		0		
Hempelmann, Ann Arbor	20	8/786(1.02)		0.05
Adenoma		7		
Carcinoma		1		
Hempelmann, Rochester	88	42/2611(1.61)		0.02
(all other categories)				
Adenoma		32	23.5	
Carcinoma		10	16.9	
Hempelmann, Rochester	401	33/261(12.6)		0.03
Subgroup C				
Adenoma		20	24.2	
Carcinoma		13	21.8	
Conrad, Marshallese children <10 yrs. at irradiation	1500	18/19(94.7)		
Adenoma		17	11	
Carcinoma		1	15	

^aThe incidence per rad is calculated for the total dose that includes 1.9 rad due to background. The expected total tumor incidence of 9/2215 was used for the NYU study.

Nelson

TABLE 4: Carcinomas produced with chromium compounds in rats

Compounds	Animals No.	Squamous cell carcinoma	Adeno- carcinoma	Hepto- cellular carcinoma
Process residue	100	1		1
Calcium chromate	100	6	2	
Chromic chromate	100			1
Chromic oxide	98			
Chromic trioxide	100			2
Cholesterol control	24			

Source: Iaskin, Kuschner and Drew (1970)

TABLE 5: Cancer and induction times seen in 200 rats following limited exposures to 0.1 parts per million bischloromethyl ether

Origin and type of cancer	Cancers total no.	Mean cancer time (days)	Range (days)
Nose: esthesioneuro- epithelioma	17	447	266-853
Malignant olfactory tumor (unclassified)	1	405	405
Ganglioneuro- epithelioma	1	334	334
Squamous cell carci- noma involving turbi- nates and gingiva	1	594	594
Poorly differentiated epithelial tumors	4	462	253-676
Adenocarcinoma (nasal cavity)	2	696	652-739
Lung: squamous cell carcinoma	13	411	215-578
Adenocarcinoma	1	877	877

Source: Ruschner, Laskin, Drew (1975)

TABLE 6: Respiratory cancer deaths among CMME-exposed workers of firm-one by duration and intensity of maximum exposure

Intensity of maximum exposure (score)	Duration of maximum exposure (years)	Distribution of workers by elapsed years since first exposure				Total for ≥ 5 elapsed years	
		0-4	5-9	10-14	15-25	Age-adj. ≤ 5 rate per 1000 years ^a	
Heavy (4-6)	≥ 5	0	1/1 ^b	1/1	1/10	3/12 (25)	23.0
	1-5	0	2/3	4/15	6/73	12/91 (13)	8.7
	<1	0	8 ^c	27	4/153	4/188	1.5
Medium (2,3)	≥ 5	0	0	0	10	1/10	
	1-5	0	11	1	65	0/77	
	<1	0	18	9	85	0/112	
Low (0.1,1)	≥ 5	0	1	1	2	0/4	
	1-5	17	13	2	10	0/25	
	<1	52	28	42	41	0/111	

^aAdjusted to the total exposed group at Firm-One.

^bNumerator is number of deaths; denominator is number of workers.

^cNumber of workers is given in cells with no deaths.

Source: Albert, Pasternack, Shore et al (1975).

TABLE 7: Respiratory cancer deaths among CMME exposed workers of firm-one with $\geq$ elapsed years since first exposure by cumulative exposure experience

Weighted cumulative exposure score ^a	Persons No.	No.	Lung cancer deaths %	Age-adjusted rate per 1000 per year ^b
20-40	40	9	23.1	23.8
12-20	60	3	5.0	2.4
8-12	61	4	6.6	3.5
5-8	57	1	1.8	1.1
3-5	73	1	1.4	1.2
<3	<u>339</u>	<u>1</u>	<u>0.3</u>	0.3
Total	630	19	3.0	

^aExposure score x years of exposure.

^bAdjusted to the total exposed group at Firm-One.

*Chloromethyl methyl ethers

Source: Albert, Pasternack, Shore, et al (1975).

REFERENCES

- Albert, R.E., Pasternack, B.S., Shore, R.E., Lippmann, M., Nelson N. and Ferris, B. (1975) Mortality patterns among workers exposed to chloromethyl ethers - A preliminary report. Environ. Health. Perspect. 11:209-14.
- Ames, B.N., Durston, W.E., Yamasaki, E. and Lee, F.D. (1973) Carcinogens as mutagens: a simple test system combining liver homogenates for activation and bacteria for detection. Proc. Nat. Acad. Sci. USA 70:2281-85.
- Eisenbud, M., Laurer, G.R., Rosen, J.C., Cohen, N. and Thomas, J. (1969) In Vivo measurement of lead-210 as an indicator of cumulative radon daughter exposure in uranium miners. Health Phys. 16:637-46.
- Federal Register January 22. 1980. 45:5001-296.
- Figueroa, G.W., Raszkowski R. and Weiss, W. (1973) Lung cancer in chloromethyl methyl ether workers. N. Engl. J. Med. 288:1096.
- Harris, C.C., Autrup, H., Conner, R., Barrett, L.A., McDowell, E.M. and Trump, E.S. (1976) Interindividual variation in binding of benzo (a) pyrene to DNA in cultured human bronchi. Science 194:1067.

Harley, N., Albert, R., Shore, R. and Pasternack, B. (1976) Follow-up study of patients treated by x-ray epilation for tinea capities: Estimation of the dose to the thyroid and pituitary glands and other structures of the head and neck. *Phys. Med. Biol.* 21:631-42.

Jennete, K.W. (1981) The role of metals in carcinogenesis: Biochemistry and metabolism. *Environ. Health Perspect.* 40:233-52.

Kuschnier, M., Laskin, S., Drew, R.T., Cappiello, V. and Nelson N. (1975) The inhalation carcinogenicity of alpha halo-ethers. 3. Lifetime and limited period inhalation studies with bis (chloromethyl) ether at 0.1 ppm. *Arch. Environ. Health* 30:73-7.

Laskin, S., Kuschnier, M., and Drew, R.T. (1969) Studies in pulmonary carcinogenesis. In: *Inhalation Carcinogenesis, Proceedings of a Biology Division*, (Eds. Hanna, M.G., Jr., Nettesheim, P. and Gilbert, J.R.) Oak Ridge National Laboratory, Conference held in Gatlinburg, TN, October 8-11, 1969 (AEC Symposium Series 18), pp. 321-351, U.S. Atomic Energy Commission, Office of Information Services, Washington, DC.

McCallum, R.I., Woolley, V. and Petrie, A. (1983) Lung cancer associated with chloromethyl methyl ether manufacture: an investigation at two factories in the United Kingdom. *Br. J. Indust. Med.* 40:384-89.

Nelson, N. (1984) Use of Toxicity Test Data in the Estimation of Risks to Human Health. ACS Symposium Series, No. 239. *Assessment and Management of Chemical Risks*. (eds. Rodricks, J.V. and Tardiff, R.G., pp. 13-22.)
Nelson

Pasternack, B.S., Wales-Collingwood K. and Shore, R.E. (1985) Lung cancer mortality among workers exposed to chloromethyl ethers. An Update. (Abstr.) Fourth International Symposium on Epidemiology in Occupational Health, Como, Italy, September 10-12.

Sakabe, H. (1973) Lung cancer due to exposure to bis chloromethyl ether. Ind. Health 11:145-48.

Shore, R., Albert, R. and Pasternack, B. (1976) Follow-up study of patients treated by x-ray epilation for tinea capitis. IV. Resurvey of post-treatment illness and mortality experience. Arch. Environ. Health 31:21-8.

Thiess, A.M., Hey, W. and Zeller, H. (1973) Zur toxicologie von dichlordimethylather verdacht auf kanzerogene wirkung auch beim menschien. Arbeitsmed Arbeitsschuz 23:97-102.

Van Duuren, B., Sivak, A., Goldschmidt, B., Katz, C. and Melchionne, S. (1969) Carcinogenicity of halo-ethers. J. Nat. Cancer Inst. 43:481-86.

Wilson, R.H., DeEds, F. and Cox, A.J., Jr. (1941) Toxicity and carcinogenic activity of 2-acetaminofluorene. Cancer Res. 1:595-608.

Weinstein. I.B. (1983) The monitoring of DNA adducts as an approach to carcinogen detection. Ann. Rev. Public Health 4:409-13.

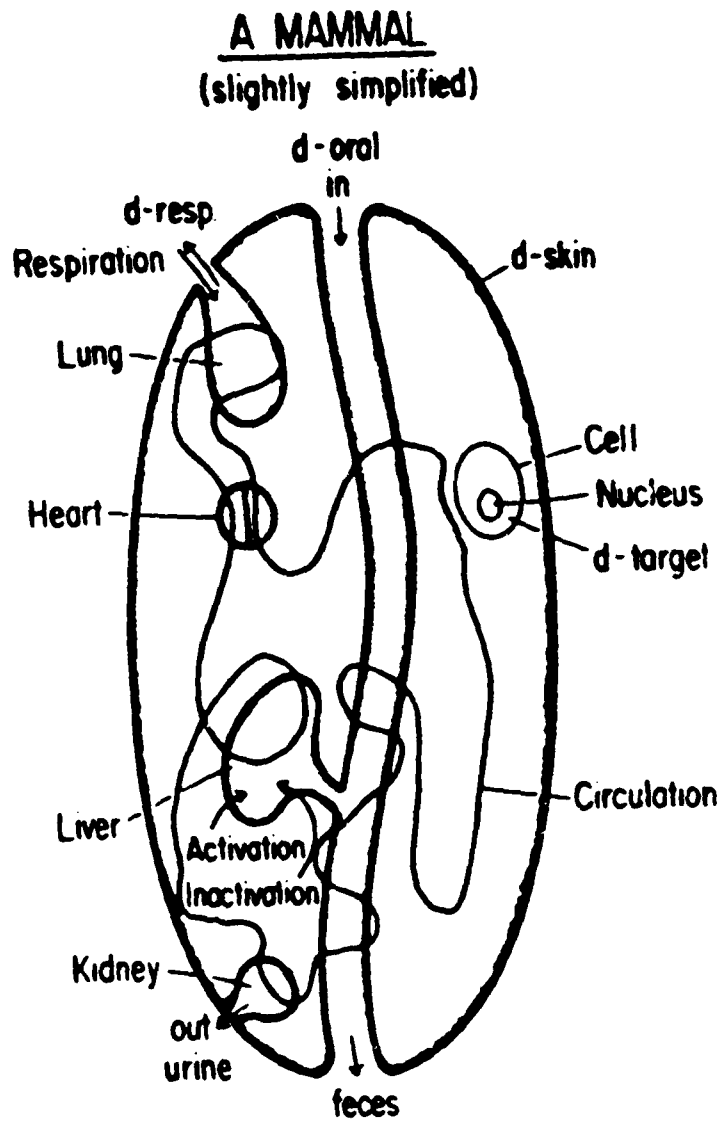


Figure 1. Simplified diagram of a mammal

FIGURE 2

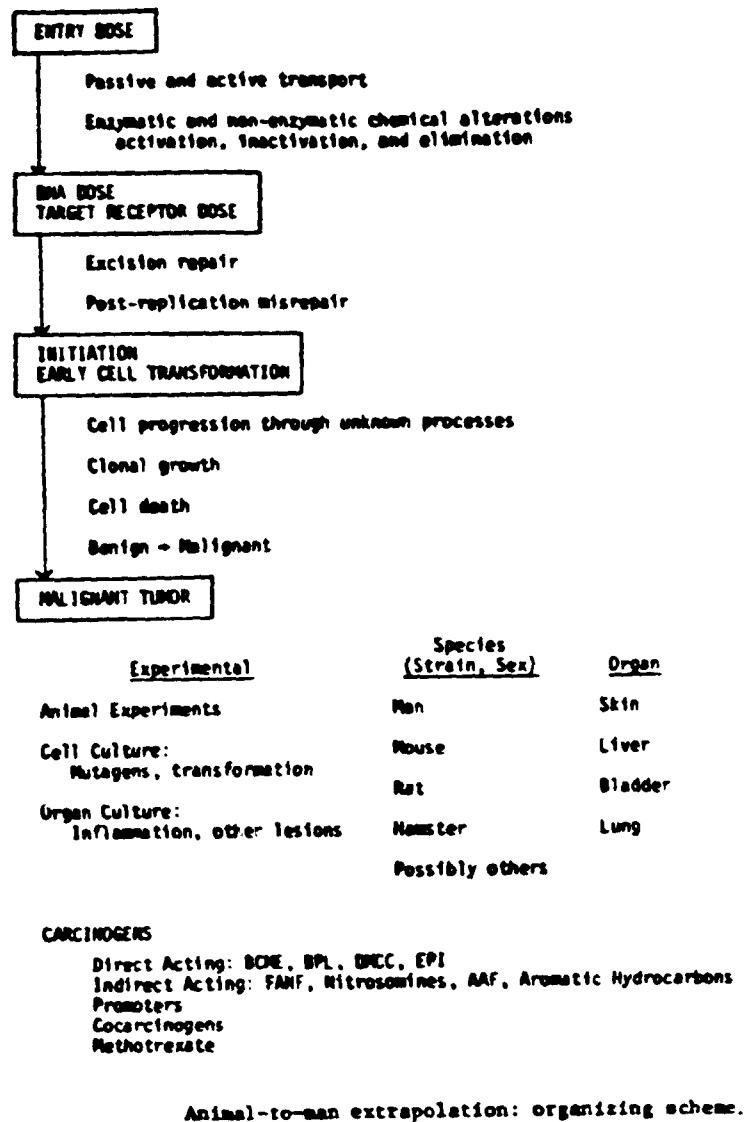


Figure 2. Organizing scheme for animal-to-man extrapolation

Part II

How Has Epidemiology Contributed to
Health Risk Assessment?

Chapter 5

Notes on the Contribution of Epidemiology to Health Risk Assessment

Michel A. Ibrahim

NOTES ON THE CONTRIBUTION OF
EPIDEMIOLOGY TO HEALTH RISK ASSESSMENT

by Michel A. Ibrahim, MD

Epidemiology has a major role in assessing risk and setting policy. Health risk can be measured by case-control, cohort, or experimental studies. The impact of exposure to noxious stimuli is often confounded by environmental factors, genetic endowments, and personal characteristics. Epidemiologic methods are useful in separating the relative effects of these factors.

Epidemiologists are well aware of the criteria, guidelines, or so called rules of evidence that may be used in judging the quality of, and hence the strength of, the inference from epidemiologic investigations. Understanding of these criteria by policy makers is equally important. The rules of evidence may be grouped into 3 areas: the range of epidemiologic studies, biases, and the characteristics of the association. The strength of inference from an observed association depends largely on the study design used to elucidate the relationship. The studies range from anecdotes and clinical hunches on one end of the spectrum to the experiment on the other. Between these two extremes are studies of ecologic correlations and cross-sectional analyses, as well as case-control, historical cohort and prospective cohort designs. The range of inference varies accordingly from speculative to suggestive, moderately suggestive, highly suggestive, moderately firm, or firm. The range of studies also provides a transition from the hypothesis generation design to hypothesis testing as studies move from the more descriptive observational types to the more analytic experimental types (Ibrahim 1985)

Systematic errors arising from poor study design, inadequate data collection or analysis are called biases in the form of information, selection, or confounding. Systematic errors introduced into the study decrease confidence in the validity of the results and inferences and therefore must be discussed thoroughly before conclusions are accepted.

The last rules of evidence set refers to the characteristics of the associations that may be judged against the so called Bradford Hill criteria. Strength of the association, dose-response relationship, temporality, consistency with current knowledge, specificity, and replication are useful in this regard.

The second point relates to the stand that may be taken by scientists in view of the evidence. There are those who advocate action on the basis of slim evidence, thus allowing for actions to stem from emotionalism and panicky behavior. There are the purists who will not act unless solid proof from controlled experiments is available. Ironically these individuals are often in the company of pressure and interest groups. A third group of individuals, which includes many epidemiologists, are those who do not wish to take a stand and who often equivocate by placing an undue number of conditions and qualifications on their findings. They become hypercritical of all studies. This posture by many epidemiologists may produce negative effects. Once a medical student listening to a lecture in which the limitations of epidemiology were outlined, with no mention of its potential, asked whether it's worth teaching a subject that can never produce valid conclusions. In the late 1970s, some government agencies dropped the funding of epidemiologic research, reasoning that it is useless.

A third point may be gleaned from Dr. Goldstein's presentation on the consensus versus confrontation approaches in policy making. In one approach, scientists review the evidence to derive conclusions after reconciling differences by attempting to meet on middle grounds for effecting change. In the other approach, polarization prevails resulting in vigorous discussions for and against a particular action without reaching a consensus. The confrontation approach is often stimulated by poorly designed studies, and publication of negative results or contradictory findings. This is fueled by intense media attention. These factors should encourage epidemiologists to adopt a set of norms or standards. Publications and discussions should be in peer reviewed journals and meetings rather than on the television screen. It was suggested a few years ago, and remains to be a good suggestion today, that a national repository of data should be established where methods of research and analysis are available to other investigators (Ibrahim and Spitzer 1979). When the findings of an investigation are at variance with what is already known, it is incumbent upon the investigator to reconcile these differences as much as possible. Finally, the use of consensus panels to make appropriate conclusions with available evidence should be an effective means for building a foundation for policy making.

The fourth point relates to an issue of classification. An additional dimension may be gained if the modifiability of risk factors and the impact of such modification are included in a classification scheme. This knowledge would be equally helpful to researchers and policy makers. An example of such a classification scheme might be: 1. risk factors which are modifiable and the modification shown to be of proven impact, 2. risk factors which are modifiable

but the impact of the modification is not entirely certain, and 3. risk factors whose modifiability and potential impact are not entirely certain (Ibrahim 1985). Policy makers may opt for setting policy when risk factors are modifiable and impact is proven, and scientists would conduct investigations on the various means of modifying a particular risk factor, and evaluating the impact of such interventions.

A final comment is in reference to a question raised by one of the participants in this conference about the relationship between scientists and policy makers. It is suggested that scientists stay out of policy making so that they keep their "purity". It is debatable whether scientists should be actively involved in policy making. The question is analogous to whether evaluators should also collect the data that will be used in the evaluation process. It is suggested that scientists provide the best information they can muster, serve on policy making committees, and participate in policy making as one of many specialists contributing to policy. But policy has many determinants other than scientific knowledge. Epidemiologists should be content when their information is taken into account in policy making, and should realize that policy making is governed by many factors including economic and political considerations.

In conclusion, epidemiology has a central role in assessing and managing health risks. Epidemiologists and policy makers alike may take note of the importance of applying the rules of evidence to investigations before reaching conclusions. Epidemiologists, after carefully reviewing the evidence, may refrain, whenever possible, from too much equivocation that would impede any action. The consensus approach in making conclusions has been an effective means of resolving differences. The modifiability of risk factors and the impact of such modification may be adopted in the classification of risk factors. Finally, in addition to scientific knowledge, we must consider the role of economic and political factors in determining policy.

Ibrahim

References

- Ibrahim MA: Epidemiology and Health Policy; Aspen Systems Corporation,
Rockville MD 1985
- Goldstein DD: The scientific basis for policy decisions. Presented at the
Epidemiology and Health Risk Assessment Symposium, 14 May 1985, Columbia,
MD
- Ibrahim MA, Spitzer WO: The Case Control Study: The Problem and the Prospect.
J Chron Dis 1979; 32: 139-144

Chapter 6
Estimation of Risk and Inferring
Causality in Epidemiology
Leon Gordis

ESTIMATION OF RISK AND INFERRING CAUSALITY IN EPIDEMIOLOGY

Leon Gordis

A common language for quantitatively expressing risk in epidemiologic studies is essential for discussing epidemiology and health risk assessment. The first part of this chapter will review the definitions of risk which will be referred to in subsequent presentations. This will be followed by a discussion of issues relating to inferences regarding causation, and finally, a number of questions will be raised regarding risk estimation and risk assessment.

I shall begin by reviewing six indices of risk which are listed in Table 1. Each will be briefly reviewed because it is important to keep in mind the differences among these indices as we discuss epidemiology and risk assessment.

The first index of risk is absolute risk. The absolute risk of a disease is the incidence of the disease. It is a very important index of risk in and of itself. For example, a pregnant woman with a possible exposure to rubella who is contemplating the possibility of interrupting her pregnancy, is interested not in the relative risk but in the absolute risk of her having a malformed child.

The second index is the relative risk. The relative risk (RR) is the ratio:

$$RR = \frac{\text{Risk of the disease in exposed individuals}}{\text{Risk of the disease in unexposed individuals}}$$

If the relative risk equals one, the risk of disease in exposed people is the same as that in non-exposed people, so that there does not appear to be any association of the exposure with the disease. (Table 2) If the relative risk is greater than 1, the risk of disease is greater in exposed individuals than in non-exposed individuals indicating an association of the exposure with the disease. For example, the relative risk of lung cancer and smoking is approximately 9 to 10. Finally, if the relative risk is less than 1, the risk of disease is smaller in exposed individuals than in non-exposed individuals suggesting that the exposure may have a protective effect in regard to the disease of interest. Such an observation would be expected, for example, if the exposure in question is an effective vaccine.

Thus, the relative risk is a measure of association. Another measure of association is the odds ratio. In a cohort study in which we study exposed and non-exposed individuals, the odds ratio is the ratio of the odds of developing the disease if the person was exposed, to the odds of developing disease if the person was not exposed. (Figure 1) In a case-control study, in which people with the disease ("cases") are compared to "controls" without the disease, the odds ratio is the ratio of the odds of the cases (those who have the disease) having been exposed to the odds of the controls (those without the disease) having been exposed (Figure 2). Thus, in either a cohort or a case-control study, the odds

ratio is the ratio of the cross-products in a 2x2 table, $\frac{ad}{bc}$. Since in a case-control study we do not have denominators of exposed and unexposed individuals, and therefore cannot directly calculate a relative risk, the odds ratio is often used as the measure of an association in case-control studies. When certain conditions are met (Table 3) the odds ratio is a very good approximation of the relative risk. It should be kept in mind,

however, that the odds ratio is itself a good measure of association of exposure and disease.

Another measure is the standardized mortality ratio (SMR). The SMR is defined as the ratio of the number of deaths observed in a specific population to the number of deaths which would have been expected in that population had the population had the same mortality experience as some "standard" population which is comparable to the population being studied in all crucial respects other than having had the exposure in question. For SMRs, person-years are often used as the denominators. The SMR has frequent applications particularly in occupational studies and its uses and limitations have been the subject of a growing body of literature in recent years.

Aside from the first index, all these measures imply comparison. Another approach to comparison is that of the Proportionate Mortality Ratio (PMR). For example, in a group of exposed individuals, the PMR for cancer is the ratio of the proportion of all deaths in the exposed group which are due to cancer to the proportion of all deaths which are due to cancer in some comparable non-exposed population.

The PMR is useful if denominators cannot be obtained for calculating rates. If a greater proportion of deaths in the exposed group are due to cancer than in the non-exposed group, one is tempted to infer an association of the exposure with cancer. However, several assumptions are involved in using PMRs, many of which are not always met, and the result is often invalid conclusions. (Table 4). In particular, since proportions of deaths are being used, if the proportion of mortality due to cancer is higher in the exposed group than in a standard group, it must be counterbalanced by a difference in

the other direction in the proportions of deaths from other causes. This is not a serious problem when the cause of death in which we are interested is a rare one relative to other causes, but under other conditions it becomes a serious consideration.

The relative risk and related indices are measures of the strength of the association between the exposure and the disease in question and involve a ratio of risks. A second approach to comparing risks uses the differences in the risks. This is called the attributable risk. The concept can be explained simply as follows: Suppose we have two populations, one exposed and one not exposed, and the risk of a disease is greater in the exposed population than in the non-exposed population. (Figure 3). The level of risk in the non-exposed group can be viewed as the background risk, (Figure 4)) a risk which is also shared by the exposed group. If we then want to know how much of the total risk in the exposed group can actually be attributed to the exposure itself (and not to the background risk which all people have) we would subtract the risk in the non-exposed (the background risk) from the total risk in the exposed. (Figure 5) Attributable risk in a group of exposed individuals is therefore calculated by subtracting the risk in non-exposed individuals from the total risk in exposed individuals. Attributable risk answers the question, "How much of the risk in exposed individuals can be attributed to the exposure?" With attributable risk we are also addressing the question: "How much of the risk of disease in exposed individuals could we hope to eliminate if we could reduce or eliminate the exposure?" It is thus a measure of the potential benefit of preventive efforts.

Attributable risk can also be calculated for a total population or a total community which consists of both exposed and non-exposed individuals. Clearly, the potential reduction in disease in a total community by eliminating a certain exposure would be dependent on the proportion of the community which

is exposed at the outset. For example, a smoking cessation campaign, if effective, would be expected to have a far greater impact on mortality from lung cancer in the total community if 95% of the population smokes than if 1% of the population smokes.

Attributable risk is vital in setting public policy. It has also gained considerable importance in the courtroom. In addressing the question of whether an individual's disease is a result of an exposure, the courts have used the test of whether it is "more likely than not" that the individual's disease was caused by the exposure.

CAUSAL INFERENCES

The concept of risk is closely linked to that of causation. Many years ago, Koch and Henle proposed a series of postulates to support a causal relationship in infectious diseases:

1) The organism is always found with the disease; 2) the organism is not found with any other disease and 3) the organism, isolated from one who has the disease, and cultured through several generations, produces the disease (in experimental animals). In 1964, when the Surgeon General's Committee met to assess the available evidence regarding smoking and lung cancer, they found that Koch's postulates could not be applied to a chronic disease such as lung cancer and articulated a set of criteria for deciding whether the relationship was causal:

To judge or evaluate the causal significance of the association between the attribute or agent and the disease, or effect upon health, a number of criteria must be utilized, no one of which is an all-sufficient basis for judgment. These criteria include:

- a) the consistency of the association
- b) the strength of the association

- c) the specificity of the association
- d) the temporal relationship of the association
- e) the coherence of the association"

After a period of over 20 years it may be appropriate to reassess these criteria and I shall only make a few brief points in this connection.

A basic question is what we mean by "cause". Rarely, if ever, do we have the clear situation of a necessary and sufficient cause. The tubercle bacillus, for example, is a necessary cause for tuberculosis but is not sufficient. Radiation is clearly not a necessary cause for leukemia since factors other than radiation may be leukemogenic. Indeed, with most chronic diseases of interest, we are dealing with complex situations in which a given causal factor is neither necessary nor sufficient.

An additional issue is the need to decide whether we are interested in a proximate or remote cause. In the case of cancer which involves a process of initiation and promotion, when we use the word "cause" are we referring to an initiator or promotor? Is genetic susceptibility to the effects of environmental factors to be considered a cause of disease?

Several questions need to be addressed regarding the Surgeon General's criteria. For example, the criterion of a temporal relationship specifies that if a factor is to be considered a cause for a disease, the exposure to it must have occurred prior to the disease. While this is intuitively logical, it may be too simplistic. For example, given what we know about DNA repair mechanisms, are we considering only the total cumulative dose of the carcinogen, or do questions of periodicity of the dose and intervals for DNA repair come into play? That DNA repair mechanisms can play a major role in epidemiologic observations is seen in the next slide (Figure 6) in which we see an earlier age of onset of skin cancer in patients with xeroderma pigmentosum

which is characterized by defective DNA repair mechanisms than in the general population.

The criterion of specificity has always been a problem. Defenders of tobacco have pointed to the lack of specificity of the cigarette smoking - lung cancer association as weakening the case for a causal relationship, since cigarette smoking is also associated with bladder cancer, heart disease, emphysema and other conditions. However, the concept of specificity is obscured because of two factors: First, tobacco itself is a mixture of chemicals and consequently a mixture of potential causal agents. Second, since all cells, regardless of the tissue in which they are found, have common biologic denominators including certain aspects of their intracellular structures and chemistry, it is not unreasonable to expect disease to be produced in different tissues and different sites even by a single agent.

Finally, and perhaps most relevant for a discussion of risk assessment, is the criterion of dose-response relationship. When such a relationship is observed, it clearly strengthens the case for a causal relationship. However, when no such dose-response relationship is seen, does it weaken the case? If a cause-effect relationship is characterized by a threshold, one might expect at times a dichotomous, all-or-none situation, which would not necessarily weaken the evidence for a causal association.

In addition, some new elements may be needed in the criteria. These include the quality of the data, the appropriateness of the study design and the data collected to the question being asked, the power of the study in relationship to the size of the sample studied, the role of interindividual variation in genetically determined susceptibility, adequate attention to confounding and interaction, and proper consideration of biases in study design and in measurement of exposure and of outcome.

Above and beyond these issues, other questions should be raised. Is the use of "criteria" appropriate for a scientific discipline? If so, how should they be used? No specific instructions for use accompanied the listing of criteria by the Surgeon General's committee. In my view, these issues are closely bound up with an assessment of hazard and risk.

Clearly, if a group with a certain characteristic or exposure is found to be at high risk, we would want to know that reducing the exposure would reduce the risk in question. For example, in recent years, some data have suggested that coffee drinking is associated with pancreatic cancer (MacMahon et al 1981). The question is, if this association is real, is the relationship of coffee drinking to pancreatic cancer a causal one or not? There is considerable evidence linking cigarette smoking to cancer of the pancreas. Since we know that cigarette smoking and coffee drinking are associated the observed association could be an indirect one resulting from confounding by cigarette smoking. (Figure 7)

In estimating risk, therefore, potential confounding becomes an extremely important issue and must be diligently searched for. It is important to point out, however, that if a third factor affects an association, the result is not always confounding. For example, the relationship of estrogen dose to thromboembolism is seen to manifest a dose-response relationship. However, blood group clearly has an influence and functions as an effect modifier. (Table 5) Since virtually all the conditions or diseases of interest in the area of risk assessment appear to be multifactorial in origin, this type of relationship is an important one to keep in mind.

SOME ISSUES IN RISK ESTIMATION AND RISK ASSESSMENT

In closing, let us consider some issues in risk assessment many of which will be addressed in subsequent presentations. Two general comments: First,

while epidemiology can effectively address the question of whether a given disease is caused by a certain factor, it has considerable difficulty with the broad question, are there any adverse health effects associated with a certain exposure. Second, epidemiology may have serious problems in detecting very low levels of risk in human populations, and the result is often a need for extrapolation to low dose levels.

Since we must extrapolate to low doses, how do we decide on the appropriate model? Is the relationship likely to be linear, quadratic, or a linear-quadratic combination? Should extrapolation use the best estimate or the upper confidence limit?

How do we deal with the issue of a time-related pattern of exposure? Is there a latent period and if so, how long? This is complicated by the fact that the latent period for a given disease may differ depending on what has caused it. While data on this issue are difficult to obtain, they are most likely to be obtainable for rare and easily recognized diseases which have several discrete causes. A good example is angiosarcoma of the liver. The next slide (Figure 8) shows the latent periods for this condition depending on whether it is caused by vinyl chloride or thorotrast, the difference is clearly seen.

Further issues relate to time: Are the observations made over a long enough time period? Are we interested in life-time risk, such as in the survivors of the atomic bombs at Hiroshima and Nagasaki? Is there a difference in risk related to the age at exposure above and beyond the cumulative exposure or years of exposure? Does periodicity of exposure make a difference?

A critical issue in risk assessment is that of the extent to which findings from a study can be generalized to other populations. The question

involves not only differences in environmental exposures, but also differences in genetic susceptibility.

Finally, how do we interpret findings of no increase in risk. Was the period of observation from initial exposure too short? Was the observation of no increased risk due to chance? Was the measurement of exposure dose adequate? Was the exposure dose too low to determine if the agent is a health hazard under normally encountered circumstances? Was a low level of risk undetectable to start with because of too small a sample size?

We live in a risk conscious society where the expectation of many is a no risk environment. Given the reality, we must refine the epidemiologic methodology so that we can best provide the scientific basis for rational policy and regulatory decisions.

TABLE 1: Indices of risk in epidemiologic studies

Absolute risk
Relative risk (RR)
Odds ratio (OR)
Standardized mortality ratio (SMR)
Proportionate mortality ratio (PMR)
Attributable risk (AR)

TABLE 2: Interpreting relative risks of a disease

- If $RR = 1$ Risk in exposed = Risk in unexposed
(no association)
- If $RR > 1$ Risk in exposed > Risk in unexposed
(positive association; ? causal)
- If $RR < 1$ Risk in exposed < Risk in unexposed
(negative association; ? protective)

TABLE 3: The odds ratio (relative odds) is a good approximation of the relative risk when

1. the "cases" studied are representative of all people with the disease in the population from which the cases were drawn, with regard to history of exposure;
2. the "controls" studied are representative of all people without the disease in the population from which the cases were drawn, with regard to history of exposure; and
3. the disease being studied is not a frequent one.

TABLE 3: The odds ratio (relative odds) is a good approximation of the relative risk when

1. the "cases" studied are representative of all people with the disease in the population from which the cases were drawn, with regard to history of exposure;
2. the "controls" studied are representative of all people without the disease in the population from which the cases were drawn, with regard to history of exposure; and
3. the disease being studied is not a frequent one.

TABLE 4: Some assumptions in using PMRs

1. Overall mortality is equal in populations being compared.
2. Classification of deaths is comparable in both populations
3. Reporting of deaths is complete in both populations.
4. Both populations are comparable in competing causes of death.
5. Data needed for appropriate adjustments are available for both populations.

TABLE 5: Effect of blood group and estrogen dose on R.R. of thromboembolism in women taking combined oral contraceptives*

	<u>ESTROGEN DOSE (g)</u>			
	<u>150</u>	<u>100</u>	<u>80-75</u>	<u>50</u>
Blood Group O	2.5	1.5	1.1	1.0
Blood Groups A, B, & AB	6.8	4.1	3.0	2.7

*Irman, 1970

Figure 1

Odds Ratio in a Cohort Study

	Develop Disease	Do not Develop Disease		
Exposed	a	b	$\frac{a/b}{c/d}$	$\frac{ad}{bc}$
Not exposed	c	d		

Figure 2

Odds Ratio in a Case-Control Study

	CASES (with disease)	CONTROLS (without disease)		
History of exposure	a	b		
No history of exposure	c	d	$\frac{a/c}{b/d}$	$\frac{ad}{bc}$

FIGURE 3

**RISKS IN EXPOSED AND
NON-EXPOSED GROUPS**

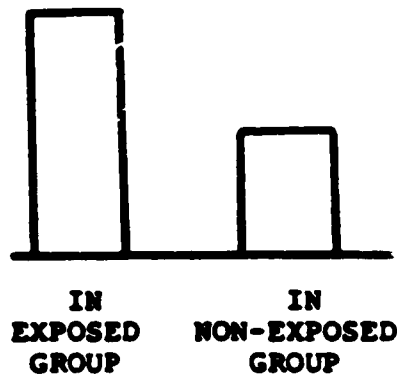


FIGURE 4

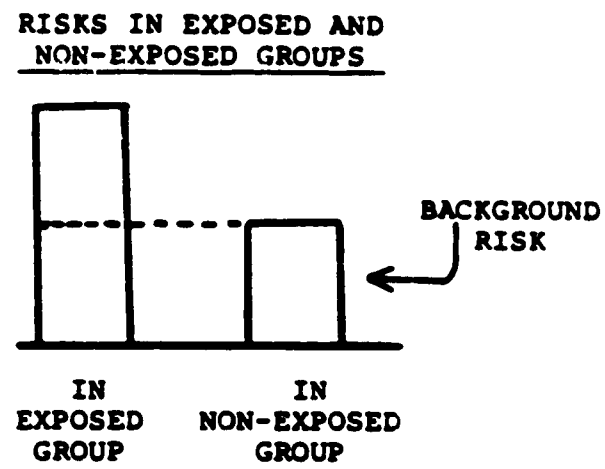
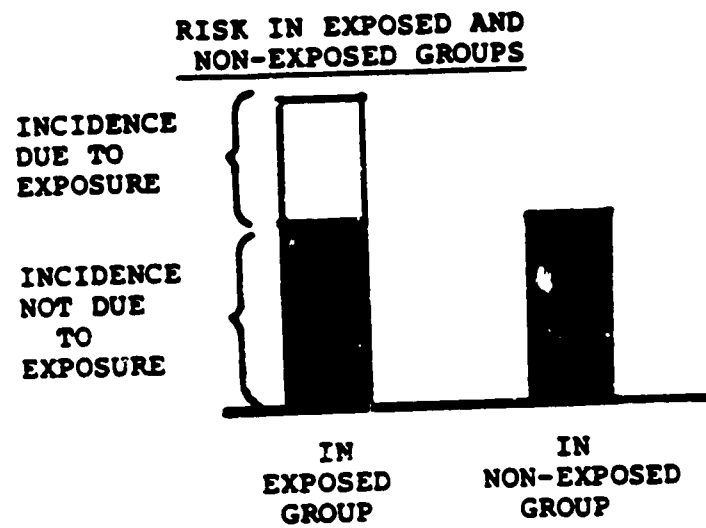


FIGURE 5



**Latent Periods for Hepatic Angiosarcoma
Induced in Males by Vinyl Chloride and Thorotrast**

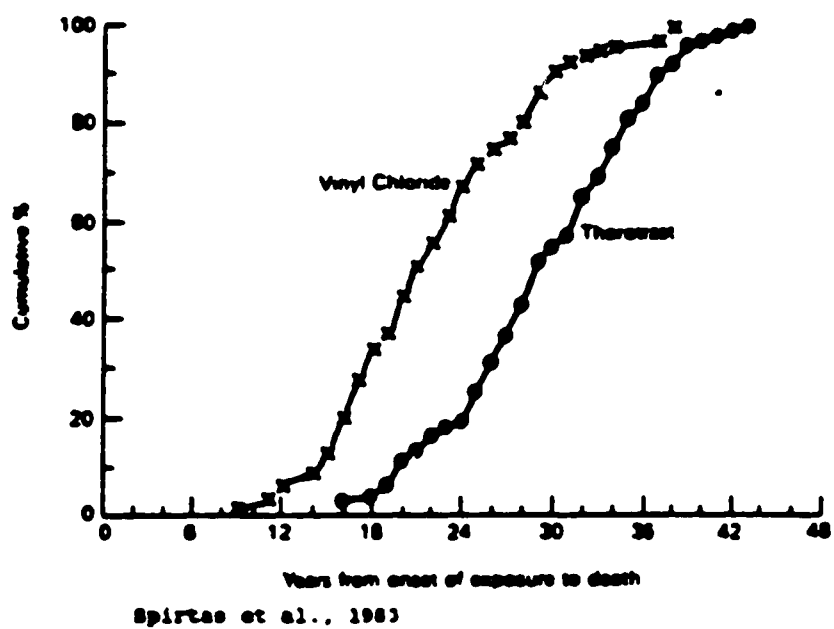


FIGURE 6

Chapter 7

Developmental Risks: Epidemiologic Advances in Health Assessment

Carol J. Rowland Hogue and Marge A. Brewster

DEVELOPMENTAL RISKS: EPIDEMIOLOGIC ADVANCES
IN HEALTH ASSESSMENT

Carol J.R. Hogue and Marge A. Brewster

Introduction

In November, 1984, the Environmental Protection Agency issued "Proposed Guidelines for the Health Assessment of Suspect Developmental Toxicants and Request for Comments". (Federal Register, 1984) These guidelines "describe the procedures that the United States Environmental Protection Agency will follow in evaluating potential development toxicity associated with human exposure to environmental toxicants". They are based on several scientific conferences and publications (U.S. EPA 1982a; U.S.EPA 1980; EPA 1982b). The general format for these guidelines is based on the National Research Council definition of risk assessment, including: hazard identification, dose response assessment, exposure assessment, and risk characterization (Committee on the Institutional Means for the Assessment of Risk to Public Health, 1983). This chapter will focus on the role of epidemiology in the first two steps of risk assessment; hazard identification and dose-response assessment.

There are two goals for this chapter. First, within the context of hazard identification and dose-response assessment, some of the issues which might limit the role ascribed to epidemiology in the quantitative risk assessment of developmental problems will be examined. Second, a few recent advances in epidemiology that have the potential to enhance the status of epidemiologic data for quantitative risk assessment will be considered. The premise of this discussion is that, rather than having a secondary role, epidemiology can and should provide the pregnancy data necessary for human developmental risk assessment of many toxicants, especially those already in the environment.

ISSUES IN EPIDEMIOLOGY

Hazard identification

Developmental risks actually begin prior to conception. Toxicants may inhibit libido, reduce potency, affect sperm quality (number, motility, or shape) or contribute to subfecundity through abnormal development of gonads, ducts and external genitalia; abnormal pubertal development; amenorrhea and anovulatory cycles; delay in conception, or male or female infertility (Bloom and Paul [eds.] 1981). The human reproductive system is extremely complex and can be adversely affected through a multitude of subsystems. For example, for the non-pregnant female, 30 reproductive processes have been identified as susceptible to reproductive toxicants (Clark 1983). For the pregnant female, an additional 18 subsystems have been identified.

Risks to the offspring include spontaneous abortion either with or without morphologic or chromosomal abnormalities, intrauterine growth retardation, preterm delivery, and a broad range of developmental problems, even including cancer. One of the problems with outcome research among offspring is that a large proportion of the adverse outcomes occurs very early in pregnancy, usually before the woman seeks medical care.

Given this tremendously broad range of potential developmental hazards coupled with difficulty in measuring key outcomes such as spontaneous abortions, regulators have turned to experimental animal and in vitro tests. Specific tests for many subsystems have been proposed and developed. The total system, however, is the human. The final assessment for any potential human developmental risk may be the post-marketing assessment of developmental risks in humans. If so, then the role of epidemiology is key in assessing human

developmental risks, especially for potential agents which are already introduced into the environment.

In fact, epidemiology has been useful in identifying developmental risks. For example, the teratogenic effect of methylmercury, polychlorinated biphenyls, thalidomide and diethylstilbestrol were all identified first in humans (Kimmel, Holsan, Hogue et al, 1984). All these discoveries were made as a result of investigating clusters of cases. Despite the value of cluster investigation, there are problems both in separating "false positive" from true problems and in missing epidemics through failure to note the cluster. Investigations of clusters of adverse reproductive outcomes are difficult to conduct, especially when there are few individuals in the exposed environment or when there is no biologically based hypothesis to test.

For example, a cluster of adverse reproductive outcomes among female employees of a school for mentally retarded children and young adults in New York City was reported to the New York City Health Department. In 1982, employees had become concerned about several of their members who had experienced spontaneous abortions, late fetal deaths, and infant deaths. The New York City Health Department investigated the school and found peeling lead paint, which was removed. They could find no other suspected environmental agent and elected to follow the employees to see if the cluster continued. By 1984, it was apparent that the problem had not abated. The Centers for Disease Control was asked to assist in this investigation (Guest, Friede, Berg et al. 1985). An investigation began late in 1984 to determine what elevation in risk was being experienced, and if significant, to recommend further studies of toxicology, environmental modification, and surveillance.

Elevated risks were found of both late fetal death and reported spontaneous abortions for pregnancies occurring while the women were working at the school. However,, it was also found that women employed during pregnancy since 1980

Hogue

-3-

were, on the average, six years older than women not employed and pregnant during the same period of time. Moreover, this increased age among employed women was not seen in the earlier time period. Finally, there was some concern that spontaneous abortions which had occurred prior to 1980 had not been as well remembered as those occurring since then. Surveillance is continuing on this population.

Clustering can occur by chance alone (Landrigan, Melius, Rosenberg et al 1983). For example, it was estimated that for the 7,000,000 video display terminal (VDT) workers in the United States, assuming that they are divided into 100,000 clusters and experience an average fertility rate similar to other women workers, over a 3-year period there would be 2,500 groups who would experience statistically significant excess adverse outcomes (with $p < .05$), 500 at $p < .01$, and 50 at $p < .001$ (Centers for Disease Control). The reported cluster under investigation in the New York City school (unpublished data) may be one of these "expected/unexpected" phenomena. Methods need to be developed to assess whether an observed cluster represents a real problem or merely one of these "expected/unexpected" clusters.

The other side of the use of cluster identification for human hazards assessment is that some clusters can be missed. Epidemics may be hidden because of failure to look for them, or because they are too subtle to be seen in front of the background incidence of adverse effects. The examples of developmental hazards that have been identified through cluster analysis represent either disastrous epidemics such as thalidomide babies and the community methylmercury disaster in Minamata, Japan, or very unusual illnesses, such as vaginal adenocarcinoma among young women associated with diethylstilbestrol exposure in utero. If the outcome is more common, for example low birthweight, then a rise due to introduction of a new developmental

Hogue

toxicant may be overlooked or totally missed. For example, in Love Canal, New York, there was an apparent rise in low birthweight among resident births in the "swale" area, the natural drainage for the chemical dump site in Love Canal. This rise coincided with the active dumping of chemical waste from 1940-1953. Elevated rates of low birthweight were reported in this area for all educational levels of the household and for women who smoked and who did not smoke (Nicholas and Polan 1984). Although there is some doubt about the true health effects of the swale area, if this report represents a true epidemic, it did not come to light until the early 1980's, 30 years after its peak.

Four observations can be made from this example. First, the information for investigating this hypothesized epidemic existed at the time of its occurrence, that is, information on birthweight was available on birth certificates at the time. Birth certificates in the United States include birthweight for almost all live born infants. Thus, a surveillance system based on birthweight could be established for the whole country. Second, low birthweight is the only adverse health effect yet reported for the Love Canal area. One explanation could be that the result occurred by chance alone. Another explanation, however, is that the result is real. Birthweight is a sensitive indicator of adverse reproductive effects. This is true for animal studies, as well. As noted by the EPA, "with the current recommendations in testing protocol being 20 rodents per dose group, it is possible to detect an increased incidence of malformations in the range of 5 to 12 times above control levels, an increase of 3 to 6 times the in utero death rate, and a decrease of 0.15 to 0.25 times the fetal weight. Thus, even within the same study, the ability to detect a change in fetal weight is much greater than for the other end points measured. Consequently, for statistical reasons only,

Hogue

-5-

changes in fetal weight are often observable at doses below those producing other signs of developmental toxicity " (Federal Register 1984). Thus, this epidemic, if real, seems to have affected only the residents living in the swale area. Analysis of community low birthweight rates or county low birthweight rates would not have revealed this epidemic. Thus, a priori geographic areas need to be established based on epidemiologic assessment of potential exposure to developmental toxins before active surveillance of low birthweight is instituted. Fourth, had the increase in low birthweight been noticed and investigated at the time of its peak activity, the level of exposure to various chemicals might have been measurable. Thirty years after the epidemic was over, however, it was not possible to relate the rise in low birthweight to any of the numerous potential developmental toxicants that had been dumped into that site. In an area in which suspected developmental toxicants exist, and in which a rise in low birthweight is observed, it could be possible to do individual monitoring of exposure levels during pregnancy. With the problems of hazard identification through analysis of clusters of adverse events, it is likely that analytic methods for risk identification in epidemiologic research will be both more scientifically valid and more fruitful in separating real from suspected hazards. However, few resources are presently expended on ongoing, carefully designed epidemiologic surveillance. More resources should be invested in carefully designed analytic epidemiologic research on suspected toxicants. With these, the role in epidemiology of identifying and quantifying human risks could be greatly enhanced.

Dose-response assessment

Quantitative risk assessment requires dose-response assessment. General problems exist in human risk quantification because of 1) small numbers both in cluster investigation and in many analytic studies; 2) the difficulty in

Hogue

-6-

quantitating exposure; and 3), the need to examine competing risks rather than just the one under investigation because humans simultaneously expose themselves to a variety of potential hazards. In addition, there are at least seven factors which affect the expression of dose-response in developmental toxicology. These include statistical design, potency of teratogen, types and combination end points, developmental periods and repair, placental transfer, maternal pharmacokinetics, and maternal homeostasis (Kimmel, Holson, Hogue et al, 1984). Three of these will be discussed by illustrating some of the complexities in assessing human developmental risks and the value of combining epidemiologic and animal data in assessment of dose-response relationships.

Estimating fetal dose. One of the few agents for which objective measures of human exposure have been made is methylmercury, carefully studied in 15 mother/infant pairs exposed accidentally in Iraq in 1972 (Kimmel, Holson, Hogue et al, 1984; Amin-Zaki, Elhassanis, Majeed et al. 1974; Koos and Longo 1976). The investigators examined mother/infant pairs in which maternal blood level had been measured to be as high as 50 ng/ml mercury in blood. Blood samples were taken at the first report of symptoms or as soon as the women were located after the outbreak had been reported. These blood measurements were taken from six weeks to ten months after the last exposure to methylmercury which had occurred as a result of eating bread made from contaminated grain.

Quite a bit of experimental work has been done on the relationship between dose, clearance, body burden, an inter-species comparisons for methylmercury exposure (Clarkson 1972). Using this extrapolation model, we estimated the peak dose of methylmercury for both maternal and fetal blood levels (Kimmel, Holson, Hogue et al, 1984). To estimate peak fetal exposure, infant blood levels were used rather than maternal blood levels. Postpartum exposure to methylmercury through breast milk was corrected for. Based on this

Hogue

extrapolation, the data are highly suggestive that, for a given maternal dosage, fetuses were exposed to higher levels of methylmercury than were their mothers.

Experimentation with rats has demonstrated that methylmercury concentrates more in fetal brain than in maternal brain (Null, Gartside and Yei 1973). There was a log-linear relationship with dosage and tissue levels of methylmercury in samples taken on the last day of exposure. The relationship between tissue level and degree of adverse effect has not as yet been established.

This example of combining experimental and human information in an attempt to quantify the risk to the fetus could serve as a model for further research, in that hypotheses raised from studies of one species can be addressed in studies of another. This concept has been proposed (Kimmel, Holson, Hogue et al, 1984); its value lies in making the best use of both human and animal models for assessing human risk.

Sensitivity of risk to the timing of exposure. The probability of an adverse reproductive outcome occurring is highly sensitive to the timing of exposure as well as to the dose. This problem sets developmental risk assessment apart from other areas of human risk assessment, such as cancer. It is well-known among teratologists that there are critical periods of gestation when an exposure will produce an adverse affect. If the mother is exposed before or after this critical period, then the chemical does not produce teratogenic effects. Moreover, during this critical time, a single exposure to a developmental toxicant can be sufficient to produce this teratogenic effect. Thus, chronic exposure may not be necessary for developmental toxicity to occur. The ramifications of these unique characteristics of developmental toxicants for assessing human exposures in occupational settings include the

Hogue

fact that mathematical models based on position and years of employment for example, cannot be used to estimate individual levels of risk for adverse pregnancy outcomes. Specifically, in the investigation of the New York City cluster of reproductive outcomes, the period in gestation during which the woman was employed at the school was examined. It was not sufficient to determine how many years she had worked in school, because the crucial time for consideration of adverse reproductive outcomes was her employment during pregnancy. Moreover, to consider birth defects, exposure during the first trimester was the crucial time.

These problems in quantification and timing of exposure also occur in animal experimentation. The minimum criteria for animal experimentation have been discussed (Kimmel, Holson, Hogue et al, 1984) and proposed guidelines have been developed (Federal Register 1984). In a review of current animal studies, very few were found to have met minimal study design criteria consisting of: 1) three experimental and one concurrent control groups; 2) ten or more litters per group; 3) complete and interpretable presentation of data; and 4) presence of a dose-response relationship.

Dose-response improvements in both human and animal research are needed before human risk assessment for developmental problems reaches its full maturity. Nevertheless, in our review of human and animal literature (Kimmel, Holson, Hogue et al, 1984), we found five compounds for which there is sufficient dose-response data for both human and animals to obtain reasonable estimates of the lowest effect level for each species. Animal species were included which appeared to be the most sensitive to the given developmental toxicant. The exception is the rabbit, for thalidomide, which was included because it is the only commonly-used laboratory species showing anomalies similar to those seen in man. For thalidomide, the dose information for human studies is also

Hogue

somewhat crude, so the information for that toxicant is perhaps less accurate than for the others.

A rather gross comparison, based on qualitative as well as quantitative evaluation of the data, and limited by the data available in the literature, could not take into account the potency of the agent at the doses listed for each species. What the data do show is that the best information available for these agents suggests that the lowest effect levels appear to be within a 10- to 100-fold range of the human effect levels. These data tend to support the use of 100 to 1000-fold safety factors in risk assessment procedures (Jackson 1980). In fact, at least a 100-fold safety factor would have been required for thalidomide, based on data in the rabbit. This analysis supports reviews (Khara 1981; Council on Environmental Quality 1981).

Confounding factors. Variables related to both exposure and outcome are termed confounding factors. These affect our ability to assess quantitative risks and to apportion risk in the individual case (Chase, Kotin, Crump et al. 1985). A major potential confounding factor is maternal smoking which has been related to spontaneous abortion, intrauterine growth retardation, premature rupture of membranes, and developmental delay. As an example of how smoking can confound the relationship between a risk factor and the outcome, low birthweight, consider the issue of coffee drinking. Although drinking seven or more cups of coffee per day has not been significantly associated in every study which has examined it, in the majority of such studies, a relationship has been found. For example, in California, heavy coffee drinking was associated with an 85 percent increase in low birthweight (van den Berg 1977). These women were 90 percent more likely to be smokers than women who drank less than seven cups of coffee a day during their pregnancy. In this study, as in many others, smoking was related to a two-fold increase in low birthweight.

Hogue

-10-

Being related to both exposure and outcome, therefore, smoking was a confounder in this study. After controlling for confounding, the relationship between coffee drinking and low birthweight was still significant, although reduced from an 85 percent excess to a 53 percent excess.

Other common human exposures which can confound studies of potential environmental risks include drugs taken during pregnancy, dietary factors, the number of previous live births, the number of previous still births and abortions, maternal age, maternal x-ray exposure, maternal weight and height, and other factors. In the evaluation of studies assessing the effect of a suspected developmental toxicant, it is important to determine the extent to which the investigator has controlled for such potential confounding factors.

As Chase and colleagues have recently pointed out, it is also important to take these confounding factors into consideration when apportioning risks for the purpose of compensation (Chase, Kotin, Crump et al. 1985). They have developed a mathematical model for apportioning risks for lung cancer and mesothelioma related to asbestos and smoking. For lung cancer, these two risk factors have been found to be multiplicative. Risk factors could also be additive, however, and mathematical models can be developed to apportion risk based on additive models, such as was found for coffee drinking and smoking in the previous example of low birthweight.

Apportioning risks requires epidemiologic data because animal models cannot adequately model the risk involved in human confounding factors. Thus, there is a key role for epidemiology to play in this area of quantitative risk assessment.

ADVANCES IN METHODS

There is obviously a gap between what epidemiology could do for quantitative risk assessment in developmental toxicology and what it is being asked to do. One of the chief reasons for this is the need to apply current, state-of-the-art methodology to new studies and to improve epidemiologic methods to handle the problems raised here, among others. Actually, some progress in improvement of methods can be reported. For example, epidemiologists are becoming increasingly concerned about the effects of misclassification on biasing study results. Moreover, they are developing methods to improve classification of both exposure and outcome. These methodology advances include early pregnancy detection, using of biomarkers, reproductive outcome surveillance with examining more than one adverse outcome.

Misclassification bias. An exposure variable may be misclassified if the measurement technique for the exposure has sensitivity less than 1.0 or specificity less than 1.0; that is, someone who is truly exposed is classified as being not exposed or one who is truly not exposed is classified as being exposed. In epidemiology, there are two types of misclassification: nondifferential and differential. With nondifferential misclassification, in comparing two risk estimates, the sensitivity and specificity parameters are the same for each group. Nondifferential misclassification bias of either the exposure or the outcome will result in underestimating the true risk ratio if it is above 1.0 or underestimating the preventive nature of the protective risk factor if the true risk ratio is less than 1.0. In other words, there is bias towards the null hypothesis.

With differential bias, in comparing two risk estimates, the sensitivity and specificity parameters are not the same for each group. Differential bias occurs, for example, when cases selectively recall an exposure better than do

Hogue

-12-

noncases. This bias can be in any direction. Some investigators have been examining methods to improve the quality of exposure information based on interviews, especially when the respondents may not be the persons who had been exposed, such as spouses of deceased cases (Hogue 1984).

Early Pregnancy detection. Most pregnancy losses occur very early in gestation, usually prior to the initiation of prenatal care. A few pilot studies are currently underway to assess the ability of sensitive pregnancy tests to detect early pregnancy loss through testing daily urine samples. This approach holds much promise, especially for noncontracepting women living or working in areas suspected of having high levels of exposure. Because early pregnancy loss is a common outcome—perhaps as common as 70 percent of conceptions—the monitoring of early loss can be done in a small cohort of exposed women.

Biomarkers. An exciting potential area for improving classification is through the use of biomarkers which could more accurately report not only the presence or absence of exposure but also the amount of individual, biological response to the exposure. Ideally, the tests developed should detect sub-toxic exposures, should reflect increasing body burdens, should be non-specific (reflect exposure to a wide variety of xenobiotics), and should be non-invasive and inexpensive. These eventually could be targeted to specific exposures, as well. Eventually, a total battery of tests should be developed.

Three tests now potentially meet these requirements for exposures during pregnancy. Two of these are based upon the xenobiotic induced alteration of enzymatic activities involved in the metabolism of the xenobiotics. A third involves a highly sensitive test for liver effects of halogenated hydrocarbons. D-glucaric acid is known to be increased upon exposure to compounds (or

metabolites) which enhance the glucuronic acid pathway. Thioethers are known to be increased upon exposure to compounds (or metabolites) which are alkylating agents conjugating to glutathione. Porphyrin pattern is altered upon exposure to halogenated hydrocarbons. These have not been studied in pregnant women to correlate pregnancy outcome with increased levels, but much biochemical and human data would suggest that they meet the criteria for nonspecificity as well as suggesting the chemical type of exposure for additional studies of specific exposures. In other words, if these biomarkers are found to be related to adverse pregnancy outcomes, they can be used to screen populations. Moreover, as these indices are very sensitive, alterations in any one of them do not necessarily indicate the presence of a permanent pathologic state or impending adverse effects upon health. Thus, it would be feasible to use these screening tests in a population suspected of being exposed to environmental agents to test the efficacy of their use in etiologic research. An additional advantage to this battery of assays is that it is well established and relatively inexpensive to administer. The rationale and background for development of this "exposure screen" of human urine samples is given as an appendix to this paper.

Reproductive outcome surveillance. There is an evolution in surveillance of adverse reproductive outcomes. The Center for Environmental Health of the Centers for Disease Control (CDC) has developed a protocol to provide active surveillance of all employees of reproductive age at CDC. For the school employees in New York City, the New York City Health Department and the CDC have begun active surveillance of every pregnancy occurring during 1985. When a woman notices that her menstrual cycle is late, she is to obtain a pregnancy test for verification of the pregnancy. Any spontaneous abortions occurring during the year will be recorded, as well as other adverse and normal pregnancy

Hogue

-14-

outcomes. In addition, fetal samples will be collected, as available, and examined for chromosomal anomalies.

These two examples reflect the philosophy that all reproductive outcomes should be monitored. To do so requires a prospective or historical-prospective research design. If a case-control design is chosen, then cases need to include all defined adverse reproductive outcomes.

Moreover, in a prospective design, investigators can also use early pregnancy test and biomarkers to reduce misclassification of outcome and exposure. These should permit a broader use of epidemiology than is available through cluster analysis alone. In conclusion, with advances in exposure and outcome measurement for developmental epidemiology, the field should soon assume its proper role as vitally important to the quantitative assessment of human developmental risks.

REFERENCES

- Amin-Zaki, E., Majeed, M.A., Clarkson, T.W., Doherty, R.A. and Greenwood M.
(1974) Intra-uterine methylmercury poisoning in Iraq. *Pediatrics*
54:587-85.
- Bloom, A.D. and Paul, N.W. (eds.) (1981) Guidelines for Studies of Human
Populations Exposed to Mutagenic and Reproductive Hazards. *Proceedings of*
Conference held January 26, 1981, Washington, D.C. New York: March of
Dimes Birth Defects Foundation, p. 44.
- Centers for Disease Control, unpublished data.
- Chase, G.R., Kotin, P., Crump, K. and Mitchell, R.S. (1985) Evaluation for
compensation of asbestos-exposed individuals; II Apportionment of risk for
lung cancer and mesothelioma. *J. Occup. Med.* 27:(3):189-98.
- Clark, J.H. (1983) Female Reproduction in *Advances in Modern Environmental*
Toxicology, Vol. III. In: *Assessment of Reproductive and Teratogenic*
Hazards. (eds. MS Christian, WM Galbraith and PV Mehlman). Princeton, NJ:
Princeton Scientific Publishers, Inc., Section II, p. 8.
- Clarkson, T.W. (1972) Recent advances in the toxicology of mercury with
emphasis on the alkylmercurials. *CRC Crit. Rev. Toxicol.* 1:203-34.
- Clarkson, T.W., Amin-Zaki, E. and Atl-Tikriti, S.K. (1976) An outbreak of
methylmercury poisoning due to consumption of contaminated grain. *Fed Proc*
35:2395-9.
- Clarkson, T.W. and Marsh, D.O. (1976) The Toxicity of Methylmercury in Man:
Dose-response Relationships in Adult Populations. In: *Effects and*
Dose-response Relationships of Toxic Metals (ed. C.G. Nordberg).
Amsterdam: Elsevier Scientific Publishing Co., pp. 246-61.

- Committee on the Institutional Means for the Assessment of Risks to Public Health. (1983) Risk Assessment in the Federal Government: Managing the Process. Commission on Life Sciences, National Research Council. Washington, D.C.: National Academy Press, pp. 17-83.
- Council on Environmental Quality. (1981) Chemical hazards to human reproduction. Washington, D.C.: U.S. Government Printing Office, January.
- Federal Register (1984) 49:227 (Nov 23) 43624-31.
- Guest, C., Friede, A., Berg C. and Hogue, C. (1985) Adverse reproductive outcomes for women working in a school for special children, New York City. Presented at the Epidemic Intelligence Service Conference, April 26, 1985.
- Hogue, C.J.R. (1984) Reducing misclassification errors through questionnaire design. In: Reproduction: The New Frontier in Occupational and Environmental Health Research. New York: Alan R. Liss, Inc. pp 81-97.
- Jackson, B.A. (1980) Safety assessment of drug residues. J. Am. Vet. Med. Assoc. 176:1141-4.
- Khara, K.S. (1981) Data presented at the toxicology forum annual meeting. Arlington, Va.
- Kimmel C.A., Holson J.F., Hogue C.J. and Carlo G.J. (1984) Reliability of experimental studies for predicting hazards to human development. NCTR Technical Report for Experiment No. 6015. National Center for Toxicologic Research, Jefferson, Arkansas.
- Koos, B.J. and Longo, L.D. (1976) Mercury toxicity in the pregnant woman, fetus, and newborn infant: A review. Am. J. Obstet. Gynecol. 126:390-409.
- Landrigan, P.J., Melius, J.M., Rosenberg, M.J., Coye, M.J. and Binkin, N.J. (1983) Reproductive hazards in the workplace. Scand. J. Work Environ. Health 9:83-8.
- Hogue

- Nicholas, J.V. and Polan A.K. (1984) Incidence of low birthweight among Love Canal residents. *Science*, 226:1217-19.
- Null, D.H., Gartside, P.S. and Yei, E. (1973) Methylmercury acculations in brains of pregnant, non-pregnant and fetal rats. *Life Sci.* 12(II):65-72.
- U.S. Environmental Protection Agency (EPA). (1982a) Health effects test guidelines, chapter II. Specific Organ/Tissue Toxicity Teratogenicity. Office of Toxic Substances. PB82-232984.
- U.S. Environmental Protection Agency (EPA). (1980) Assessment of Risks to Human Reproduction and to Development of the Human Conceptus from Exposure to Environmental Substances, pp. 99-116. DE82-007897.
- U.S. Environmental Protection Agency (EPA). (1982b) Pesticides Registration: Proposed Data Requirements. *Fed. Reg.* 47:53192-203.
- van den Berg, B.J. (1977) Epidemiologic observations of prematurity: Effects of tobacco, coffee and alcohol. In: *The Epidemiology of Prematurity* (eds. D.M. Reed and F.J. Stanley). Baltimore-Munich: Urban Schwarzenberg, pp. 157-76.

Appendix

Rationale and Background for Development of an
"Exposure Screen" of Human Urine Samples

The multitude of natural and man-made xenobiotic (foreign) chemicals that occur in the environment and in occupations dictates screening methods be developed to detect and monitor single or multiple human exposures to these xenobiotics. Ideally, the tests developed should detect sub-toxic exposures, should reflect increasing body burdens, should be non-specific (reflect, exposure to a wide variety of xenobiotics), and should be non-invasive and inexpensive. The diversity of xenobiotic chemical structures will require that a battery of such tests be developed for maximal coverage of potential exposures.

Three tests of urine potentially meeting these requirements are proposed. Two of these, glucuronic acid and thioethers, are based upon the xenobiotic-induced alteration of enzymatic activities involved in the metabolism of the xenobiotics and of normal endogenous compounds. A third, porphyrin pattern, involves a highly sensitive test for liver effects of halogenated hydrocarbons. Reported effects of individual xenobiotics upon these three urine parameters are summarized in Table 1 (animals and man).

Urinary D-Glucuronic Acid—A reflector of enhanced glucuronidation

Many xenobiotics are biotransformed by conjugation with glucuronic acid, usually having first stimulated the glucuronidation pathway. Glucuronic acid's ability to react with several functional groups (in its active form of UDP-glucuronic acid) and its large tissue supply (from glycogen) make it a major player in detoxification of xenobiotics (Notte and Henderson 1977a). Stimulation of the glucose-to-UDP-glucuronide pathway results in increased

urinary excretion of D-glucaric acid (and, in animals excepting primates and guinea pig, increased excretion of glucaric acid appears to be an enhanced supply of UDP-glucose secondary to decreased glycogen synthesis, although direct induction of one or more enzymes between UDP-glucose and glucaric acid may also play a role (Notten and Henderson 1975). Barbiturate administration to rats was first shown to stimulate the glucuronic acid pathway in 1940 (Longnecker, Fricke and King 1940), then in 1963 was shown by Marsh and Reid (1963) to increase urinary glucaric acid excretion. Since that time, numerous therapeutic drugs have been shown to stimulate this pathway and to concomitantly increase hepatic microsomal drug metabolizing enzymes (Latham 1974; Lecanwasam, Franklin and Turner 1975; Hunter, Carella, Maxwell, et al. 1971; Hunter, Maxwell, Stewart, et al. 1972; Notten and Henderson 1973; Notten and Henderson 1975a; Notten and Henderson 1975b). Hunter, Maxwell, Stewart, et al. (1973) found a significant correlation between the hepatic cytochrome P450 content and the urinary excretion of D-glucaric acid, but there is evidence that enhanced glucaric acid excretion is not always accompanied by induction of hepatic microsomal enzymes (Notten and Henderson 1977a). A well-controlled study of seven healthy males given phenobarbital for 14 days showed up to a seven fold increase in glucaric acid excretion (peaking at day 15 and declining gradually to baseline), and suggested that those subjects with low control values responded most to phenobarbital administration with glucaric acid increase (Roots, Saalfrank and Hildebrant 1975).

Physiologic Alteration and Measurement. The state of pregnancy (Herzberg, Tenenbaum, Fishel et al. 1977) and estrogen therapy (Mowat 1968) reportedly increase glucaric acid excretion, although not significantly. No increase has been found due to cigarette smoking (Fiedler, Schroter and Cramer 1980).

Expressed as D-glucaric acid excretion per day, males have approximately two

Hogue

-20-

fold higher values than females (Simmons, Davis, Dordoni et al. 1974), a sex difference not found when excretion is referenced to creatinine excretion (Fiedler, Schroter and Cramer 1980; March, Turner, Shanley et al. 1974). The effects of age upon excretion of D-glucaric acid have been reported in some papers as no different between children (6-12 years), adult and geriatric ages (March, Turney, Shanley et al. 1974) while others have reported positive age correlation from 18 to 65 (Fiedler, Schroter and Cramer 1980).

There are currently enzymatic (March and Reid 1963; Fiedler, Schroter and Cramer 1980; Simmons, Davis, Dordoni et al. 1974; March, Turner, Shanley et al. 1974) and chromatographic (Ishidate, Matsui and Okada 1965; Gangolli, Longland and Shilling 1974) approaches to measurement of glucaric acid excretion, with widely differing reported reference values, as summarized by Fiedler, Schroter and Cramer (1980). Most studies have employed the enzymatic approach first used by Marsh and Reid (1963); as this approach is highly subject to interference by other urine constituents Fiedler, Schroter and Cramer (1980) advise against this approach. Intra-individual variation has been reported as high as 50 percent, possibly reflecting variation in daily effects of environment, nutrition, or other factors (Roots, Saalfrank and Hildebrandt 1975).

Disease Effects. Disease states known to be accompanied by increased excretion of glucaric acid include alcoholism (Mezey 1976), early stage renal disease in children (Nemeth and Szeleczki 1980) but not in adults with chronic renal insufficiency (Kampf, Roots and Hildebrandt 1980) and liver diseases (Herzberg and Weiner 1978; Sorrel, Burnett, Tuma et al. 1976). Decreased values of glucaric acid excretion have been found to occur in patients with congestive heart failure (Tokola, Pelkonen, Karki et al. 1975), starvation (Sotaniemi, Medzihradsky and Eliasson 1975), severe burns (Ciaccio and Fumcillo 1979), and

Hogue

1979) and in children with favism (Cassimos, Malaka-Zafirium and Tsiuris 1974) and with dystonic reactions to normal doses of antiemetics (Cassimos, Tsiuris, Danielides et al. 1975).

Exposure Effects. Hunter, Maxwell, Stewart et al. (1972) studied glucaric acid excretion of factory workers involved in producing organochlorine pesticides (Aldrin, Dieldrin, Endrin and DDT). Their results show increases in some, but not all, exposed employees and also correlate glucaric acid excretion quantitatively with DDT metabolism. Two studies of Endrin manufacturing plant workers have been performed; Vrij-Standhardt, Strik, Ottevanger et al. (1979) showed higher values after seven working days than after three leave days and Ottevanger and Van Sitter (1979) showed higher values correlated to exposure days and also to excretion of the major Endrin metabolite. Notten and Henderson (1977b) studied glucaric acid excretion in employees from various metal and chemical factories. Their data again show increases in some but not all employees potentially exposed. Styrene exposure apparently does not increase glucaric acid excretion in workers with exposure documented by measurement of urine metabolites (mandelic and phenylglyoxylic acids) (Hotz, Guillemin and Lob 1980). Seutter-Berlage, Wagenaars-Zegers, Hoog-Antink et al. (1979) reported increased glucaric acid excretion in ten workers in a plant packaging a variety of chemicals (insecticides, fungicides, herbicides, fertilizers, etc.). The finding of Ottevanger and Van Sitter (1979) that glucaric acid excretion in Endrin manufacturing workers was normal after a six-week shutdown for maintenance suggested that induction of the glucaric acid pathway was reversible, similar to the response reported after withdrawal from therapeutic drugs.

Notten and Henderson (1977a) studied the effects of numerous compounds including organic solvents, insecticides, fungicides and herbicides in chronic low-dose modes; their results (Table 1) led them to state "it seems justified to assume that an enhanced urinary excretion of D-glucaric acid is useful as a non-specific parameter for exposure to xenobiotic compounds."

Urinary thioethers—A reflector of alkylating agent detoxification

Conjugation of alkylating xenobiotics to the sulfur of glutathione and further metabolism of this xenobiotic-tripeptide conjugate, results in the urinary excretion of S-alkyl-N-acetyl-cysteine. These alkyl derivatives of cysteine are included in the chemical class known as thioethers (structure - R-S-R). Xenobiotics known to undergo this detoxification sequence in animals are listed in Table 1 and include a large variety of compounds (Boyland 1971). Although studies of this detoxification pathway began in 1879, measurement of urinary thioether excretion as a biological parameter of human exposure to potentially alkylating agents was not proposed until 1977 (Seutter-Berlage, Van Dorp, Kosse et al. 1977). In many circumstances, it is a metabolite of the xenobiotic that is reactive (for example, an alkylating agent) rather than the parent compound (Mitchell 1975).

Measurement. All published methods for measuring thioethers in human urine are based on the determination of SH groups released by alkaline hydrolysis, the Ellman method (Ellman 1959). Seutter-Berlage, Selten, Oostendrop et al. (1979) modified the Ellman method to correct for free SH groups and disulfides in urine. Pentz (1978) found no increase in thioether excretion of rats treated with many toxic compounds, employing the method of Seutter-Berlage. Buffoni, Coppi and Santoni (1982) questioned the direct application of the

Ellman method to human urine for estimation of thiol compounds and developed a "highly specific" method for estimating urinary thioethers.

Physiology and disease alterations. Seutter-Berlage, Van Dorp, Kosse et al. (1979) reported approximately 1.5 fold higher values of urinary thioethers in smokers and also demonstrated intra-individual variations of 15-24% (coefficient of variation), using their original unmodified thioether assay. Their report also summarizes the known occurrence of sulfur containing substances in foods (onions, cabbage species, pineapples and bean species) and points out the advantage of reducing their contribution to total thioethers by volatilization. Buffoni, Coppi and Santoni (1982) also report higher (2 fold) thioether excretion in smokers versus non-smokers. In contrast to the glucuronic acid assay, thioether excretion has not been studied with regard to alterations related to age, various disease states, or drug therapies. Endogenous metabolites known to conjugate with glutathione include dihydroxyphenylalanine (DOPA), estrogens, and prostaglandins (Seutter-Berlage, Selten, Oostendrop et al. 1979).

Exposure effects. Seutter-Berlage, Wagenaars-Zegers, Hoog-Antink et al. (1979) found elevation in thioether excretion in production workers in a pesticide packaging factory (as compared to administrators), but found an even greater effect of smoking. They also found higher urinary thioether levels in chemical workers as compared to metal workers (Seutter-Berlage, Van Dorp, Kosse et al. 1979). Kilpikari (1981) showed that rubber and tire workers had urinary thioether concentrations twice that of controls. Vainio, Salvolainen and Kilpikari (1978) showed elevations in urinary thioethers of chemical workers. Operators of chemical waste incinerators likewise appear to have enhanced thioether excretion (Van Door, Leijdekkers, Bos et al. 1981). A recent study

by Jagun, Ryan and Waldron (1982) of urinary excretion of thioethers in nurses handling cytotoxic drugs (mainly cyclophosphamide, vincristine and cytoxan), showed a relationship to such exposure. Mean excretion after five days of such exposure was seven fold that of control nurses, dropping to near control levels three days after cessation of exposure. This finding is consistent with an earlier report of greater mutagenicity of such nurses' urine samples in the Ames test (Falck, Grohn, Sorsa et al. 1979).

Urinary porphyrin pattern—A reflector of halogenated hydrocarbons

The pattern of urinary porphyrin excretion provides a third "index of exposure", reflecting sub-clinical liver effects of xenobiotics that are halogenated hydrocarbons.

Porphyrinogens are precursors of the ring structure in heme which is then incorporated into hemoglobin and other heme proteins. Interruption of this synthetic pathway results in enhanced excretion of porphyrins which vary as to the number of carboxyl groups. Urinary porphyrins of normal individuals include predominantly coproporphyrin (4 carboxyls)>> uroporphyrin (8 carboxyls)>> other porphyrins (7,5,6,3 carboxyls). Severe disturbance of this pathway ultimately results in greatly increased quantities of porphyrins appearing in urine (porphyria) and skin symptoms (porphyria cutanea tarda). Strik and Koeman (1974) summarize the evidence supporting the progression sequence of liver damage leading to secondary coproporphyrinuria then to chronic hepatic porphyria (CHP types A B C), then to latent porphyria cutanea tarda, and finally to overt porphyria cutanea tarda (type D). Each of these stages (coproporphyrinuria and CHP types A through D) can be recognized chemically by pattern and amount of excreted porphyrin (Doss 1979). The earlier stages show alteration in porphyrin ratios whereas later stages also show elevation in quantity.

Hogue

-25-

Physiology and methodology. Increase in total urine porphyrins is known to follow intoxication with heavy metals (Goldberg and Rimington 1962), hormones (Granick 1966), and drugs (Granick 1966; Schmid 1963) as well as with halogenated aromatic hydrocarbons (Ockner and Schmid 1961; Goldstein, Hickman and Jue 1974; Strik 1979a). Halogenated hydrocarbons known to produce chronic hepatic porphyria in man are listed in Table 1 (Strik 1979a). In experimental studies, a chronic exposure is usually needed to evoke porphyria, especially in mammals (Strik 1973). Doss (1979) discusses the chronic hepatic porphyrias resulting from liver diseases, alcohol or estrogens, and from an enzymatic defect in uroporphyrinogen decarboxylase. This last enzymatic defect can be genetic (autosomal dominant) or secondary to certain xenobiotic compounds. The secondary chronic hepatic porphyrias can occur in toxin exposures without an underlying genetic defect, but the defect is required for manifestations of porphyria induced by alcohol or simultaneous presence of liver disease, alcohol or estrogens. Earlier stages of chronic hepatic porphyria appear to be non-symptomatic and potentially reversible once the damaging agent is removed. Presumably, all xenobiotics known to elevate total porphyrins would show the urinary pattern of milder stages at an earlier time, but the porphyrin pattern has not been studied in every porphyrinogenic exposure (Table 1).

The original method of pattern determination was quite laborious, involving multiple extractions and thin-layer chromatography (Strik and Harmsen 1974), limiting the number of samples assessed in exposure situations. A recent method for separating and quantitating individual porphyrins by high pressure liquid chromatography has been reported by Ford, Ou and Ellefson (1981). Effects of age upon these values have apparently not been determined.

Exposure studies. A 1978 study of 115 urine samples from persons exposed to dioxin (TCDD) in the 1976 plant explosion in Seveso, Italy, revealed 84 percent of samples with abnormal porphyrin patterns, and 23 percent also with increased total porphyrins (Center, Strik and Colombi 1979). A 1977 study of Michigan farm families exposed to PBBs (polybrominated biphenyls; Firemaster BP-6) via cattle feed (in an accident which began in 1973), showed 47 percent of 142 samples with abnormal porphyrin patterns, but none had increased total porphyrins (Strik, Doss, Schraa et al. 1979). These two studies detected abnormalities two years after exposure to substances known to be stored in body fat. In contrast, a 1978 study of 16 persons exposed to PCBs (polychlorinated biphenyls) via contaminated rice oils in 1968 did not reveal an abnormal porphyrin pattern—a ten year study lag (Strik, Kip, Yoshimura et al. 1979). Strik (1979) studied 22 samples of persons with occupational exposures to 2,4,5-T (including dioxin), finding 64 percent with abnormal porphyrin pattern, 50 percent also with increased total porphyrins. Of 40 cases with vinylchloride disease, 75 percent had abnormal porphyrin patterns (Lange, Block, Veltma et al. 1976). Of 39 samples of person occupationally "exposed" to hexachlorocyclopentadiene, alkyl chloride, epichlorohydrin and Endrin, no abnormal porphyrin patterns were seen (Nagelsmith, VanVliet, Van de Wiel-Wetzels 1979). These human studies indicate that halogenated hydrocarbon exposure, if still present either from fat or the environment, alters the urinary porphyrin pattern (Strik 1979a).

Summary

These three urinary assessments appear to provide the nucleus of a pattern of non-invasive tests of exposure to xenobiotics. Much remains to be learned about their response to specific environmental chemicals and about their quantitative relationship to body burdens of these agents. Such knowledge can be gained by including this battery of tests in known exposure situations, or alternatively, by extensive investigation of toxin exposures in persons with abnormal tests, with and without adverse outcomes. Inclusion of such tests in epidemiologic studies should vastly increase the power of the studies by more objective assignment of subjects to an exposure group.

TABLE 1: Reported alterations in urine parameters by xenobiotics

<u>Urine Parameter</u>	<u>Effect</u>	<u>Reference</u>	<u>Xenobiotics</u>
Glucaric acid		Notten and Henderson (1977a)	Diuron, Hexachlorobenzene, Heptaclor, Dieldrin, Aldrin, Rhotane, Disulfiram Pentachlorophenol, Perthane, 2-Phenylphenol
		Vrij-Standardt, Strik, Ottevanger et al. (1979); Ottevanger and Van Sitter (1979)	Endrin
		Notten and Henderson (1977a)	Toluene, tetraethyl lead, Arcoclor 1260, Ethanol, n-Hexane, Dodin, Atrazine
		Notten and Henderson (1977a)	Captan, Dimethoate, Nitrobenzene, Aniline, Naphthylacetic acid, Benzene, Rotenone, Binapacryl
		Hotz, Guillemin and Lob (1980)	Styrene

TABLE 1 (Cont'd.): Reported alterations in urine parameters by xenobiotics

<u>Urine Parameter</u>	<u>Effect</u>	<u>Reference</u>	<u>Xenobiotics</u>
Thioethers		Boyland (1971)	Benzene, Naphthalene, Anthracene, Phenanthrene, Benzanthracene, Styrene, Aniline, 2-Naphthylamine, Acetophenetidine, Halobenzenes, 1-Chloronaphthalene, Halonitrobenzenes, Benzl chloride, Phenethyl bromide, 1-menaphthyl chloride, Benzyl acetate, 1-menaphthyl acetate, 3,5-Di-tertbutyl-4-hydroxy-toluene, Iodomethane, Bromoethane, Allyl chloride, 1-Nitropropane, Cyclopentene, Bromocyclohexane, Maleic acid, Isovaleric Acid, Ethyl methanesulphonate, Urethane, Benzothiazole-2-sulphonamide, bis-beta-chloroethyl sulphide, Ethacrynic acid, Arocoline, Allyl acetate
Porphyrin pattern	abnormal	Centen, Strik and Colombi (1979)	Dioxin (2,3,7,8 TCDD)
		Strik, Doss, Schraa et al. (1976)	Polybrominated biphenyls (PBBs)
		Strik (1979b)	2,4,5-T
	quantity*	Strik (1979b)	Polychlorinated Biphenyls (PCBs), PBBs, Hexachlorobenzene, Octachlorostyrene, Dichlorobenzene, Hexabromobenzene, Lindane, Methyl chloride, vinyl chloride

*Porphyrin pattern not determined
Hogue-Appendix, Pt.2

REFERENCES

- Boyland, E. (1971) Mercapturic acid conjugation. In: Handbook of Exp. Pharmacol. (eds. Brodie and Gillette) Springer-Verlag, Berlin, 28:II:584.
- Buffoni, F., Coppi, C. and Santoni G. (1982) Determination of thioethers in urine. Clin. Chem. 28:249-50.
- Cassimos, C., Malaka-Zafirium, K. and Tsiuris, J. (1974) Urinary D-glucaric acid excretion in normal and G6PD deficient children with favism. J. Pediatr. 84:971-72.
- Cassimos, C., Tsiuris, J., Danielides, B. and Malaka-Zafirium, K. (1975) Urinary D-glucaric acid excretion in children with dystonic reactions caused by antiemetic drugs. J. Pediatr. 87:981-82.
- Centen, A.H.J.F., Strik, J.J.T.W.A. and Colombi, A. (1979) Coproporphyrinuria and chronic hepatic porphyria type A found in people from Seveso (Italy) exposed to 2,3,7,8-tetracholodibenzo-p-dioxin (TCDD). In: Chemical Porphyria in Man, (eds. J.J.T.W.A. Strik and J.H. Koeman), Elsevier, N.Y., pp. 75-82.
- Ciaccio, E.U. and Funcillo, R.J. (1979) Urinary excretion of D-glucaric acid by severely burned patients. Clin. Pharmacol. Ther. 25:340-44.
- Doss, M. (1979) Chronic hepatic porphyrias in humans (Endogenic Factors) In: Chemical Porphyria in Man (eds. J.J.T.W.A. Strik and J.H. Koeman), Elsevier, N.Y., pp. 11-26.
- Doss, M. and Schwartz, S. (1978) Diagnosis and therapy of porphyrias and lead intoxication. J. Clin. Chem. Clin. Biochem. 16:25-82.
- Ellman, G.L. (1959) Tissue sulfhydryl groups. Arch. Biochem. Biophys. 82:70-7.
- Falck, J., Grohn, P., Sorsa, J., Vainio, H., Heinonen, E. and Holsti, L.R. (1979) Mutagenicity in urine of nurses handling cytostatic drugs. Lancet I:1250-51.
- Hoque-Appendix, Pt.2

- Fiedler, K., Schroter, E. and Cramer, H. (1980) Glucaric acid excretion. Analysis of the enzymatic assay and findings in smokers and non-smokers. Eur. J. Clin. Pharmacol. 18:429-32.
- Ford, R.E., Qu, C. and Ellefson, R.D. (1981) Liquid chromatographic analysis for urinary porphyrins. Clin. Chem. 27:397-401.
- Gangolli, S.D., Longland, R.C. and Shilling, W.H. (1974) A gas liquid chromatographic method for the determination of D-glucaric acid in urine. Clin. Chem. Acta 50:237-43.
- Goldberg, A. and Rimington, C. (1962) Diseases of Porphyrin Metabolism, C.C. Thomas Co., Springfield, IL.
- Goldstein, J.A., Hickman, P. and Jue, D.L. (1974) Experimental hepatic porphyria induced by polychlorinated biphenyls. Toxicol. Appl. Pharmacol. 27:437-48.
- Granick, S. (1966) The induction in vitro of the synthesis of alpha-aminolevulinic acid synthetase in chemical porphyria: A response to certain drugs, sex hormones, and foreign chemicals. J. Biol. Chem. 241:1359-75.
- Herzberg, M., Tenenbaum, E., Fishel, B. and Wiener, M. (1977) D-glucaric acid and gamma glutamyl transferase as indices of hepatic enzyme induction in pregnancy. Clin. Chem. 23:596-98.
- Herzberg, M. and Weiner, M.H. (1978) Increased D-glucaric acid excretion by jaundiced patients. Clin. Chem. 24:1759-61.
- Hotz, P., Guillemin, M.P. and Lob, M. (1980) Some hepatic effects (induction and toxicity) caused by occupational exposure to styrene in the polyester industry. Scand. J. Work Environ. Health 6:206-15.

- Hunter, J., Carella, M., Maxwell, J.D. and Stewart, D.A. (1971) Urinary glucaric acid excretion as a test for hepatic enzyme induction in man. *Lancet* I:572-75.
- Hunter, J., Maxwell, J.D. Stewart, D.A. and Williams, R. (1972) Increased hepatic enzyme activity from occupational exposure to certain organochlorine pesticides. *Nature* 237:399-401.
- Hunter, J., Maxwell, J.D., Stewart, D.A. and Williams R. (1973) Urinary D-glucaric acid excretion and total liver content of cytochrome P450 in guinea pigs: relationship during enzyme induction and following inhibition of protein synthesis. *Biochem. Pharmacol.* 22:743-47.
- Ishidate, M., Matsui, J. and Okada, J. (1965) Biochemical studies on glucuronic acid and glucaric acid. I. Quantitative chemical determination of D-glucaric acid in urine. *Anal. Biochem.* 11:176-86.
- Jagun, O., Ryan, M. and Waldron, H.A. (1982) Urinary thioether excretion in nurses handling cytotoxic drugs. *Lancet* II:443-44.
- Kampf, D., Roots, I. and Hildebrandt, A.G. (1980) Urinary excretion of glucaric acid, an indicator of drug metabolizing enzyme activity, in patients with impaired renal function. *Eur. J. Clin. Pharmacol.* 18:255-61.
- Kilpikari, I. (1981) Correlation of urinary thioethers with chemical exposure in a rubber plant. *Br. J. Ind. Med.* 38:98-100.
- Lange, C.E., Block, H., Veltma, G. and Doss, M. (1976) Urinary porphyrins among PVC workers. In: *Porphyrins in Human Diseases.* (ed. M. Doss), Karger, Basel, p. 352.
- Latham, A.H. (1974) D-glucaric acid as an index of hepatic enzyme induction by anticonvulsant drugs in man. *J. Pharm. Pharmacol.* 26:284-86.
- Lecamwasam, D.S., Franklin, C. and Turner, P. (1975) Effect of phenobarbitone on hepatic drug-metabolizing enzymes and urinary D-glucaric acid excretion in man. *Br. J. Clin. Pharmacol.* 2:257-62.
- Hogue-Appendix, Pt.2

- Longnecker, H.E., Fricke, H.H. and King, C.G. (1940) The effect of organic compounds upon vitamin C synthesis in the rat. *J. Biol. Chem.* 135:497-510.
- March, J., Turner, W.J., Shanley, I. and Field, J. (1974) Value of urinary excretion of D-glucaric acid by normal individuals. *Clin. Chem.* 20:1155-58.
- Marsh, C.A. and Reid, L.M. (1963) Changes in D-glucaric acid excretion induced by stimulators of ascorbic acid biosynthesis. *Biochem. Biophys. Acta* 78:726-28.
- Mezey, E. (1976) Increased urinary excretion of D-glucaric acid in alcoholism. *Res. Commun. Chem. Pathol. Pharmacol.* 15:735-42.
- Mitchell, J. (1975) Covalent binding: Useful index of the formation of chemically active drug metabolites. *Biochem. Soc. Trans.* 3:622-29.
- Mowat, A.P. (1968) Developmental effects on liver D-glucuronolactone dehydrogenase levels and on D-glucaric acid excretion in urine; hormonal effects on D-glucaric acid excretion in urine. *J. Endocrinol.* 42:585-90.
- Nagelsmit, A., VanVliet, P.W., Van der Wiel-Wetzels, W.A.M., Wielard, M.J., Strik, J.J.T.W.A., Ottevanger, C.F. and Van Sitter, N.J. (1979) Porphyrins as possible parameters for exposure to hexachlorocyclopentadiene, allylchloride, epichlorohydrin and Endrin. In: Chemical Porphyrin in Man, (eds. J.J.T.W.A. Strik and J.H. Koeman), Elsevier, N.Y., pp. 55-62.
- Nemeth, I. and Szeleczki, T. (1980) Increased glucaric acid excretion in children with renal disease. *Int. J. Clin. Pharmacol. Ther. Toxicol.* 18:357-60.
- Notten, W.R.F. and Henderson, P.T. (1975a) Alterations in urinary D-glucaric acid excretion as an indication of exposition to xenobiotics. In: Proceedings of the International Symposium - Environmental and Health, ECE-EPA-WHO, Paris, 1974.
Hague-Appendix, Pt.2

- Notten, W.R.F. and Henderson, P.T. (1975b) The influence of n-hexane treatment on the glucuronic acid pathway and activity of some drug metabolizing enzymes in guinea pig. *Biochem. Pharmacol.* 24:127-31.
- Notten, W.R.F. and Henderson, P.T. (1973) Effect of disulfiram on the urinary D-glucaric acid excretion and activity of some enzymes involved in drug metabolism in guinea pig. *Arch. Int. Pharmacodyn. Ther.* 205:199-208.
- Notten, W.R.F. and Henderson, P.T. (1977a) The Interaction of chemical compounds with the functional state of the liver. I. Alteration in the metabolism of xenobiotic compounds and D-glucuronic acid pathway. *Int. Arch. Occup. Environ. Health* 38:197-207.
- Notten, W.R.F. and Henderson, P.T. (1977b) The interaction of chemical compounds with the functional state of the liver. II. Estimation of change in D-glucaric acid synthesis as a method for diagnosing exposure to xenobiotics. *Int. Arch. Occup. Environ. Health* 38:209-20.
- Notten, W.R.F. and Henderson, P.T. (1975) Alterations in the glucuronic acid pathway caused by various drugs. *Int. J. Biochem.* 6:111-19.
- Ockner, R.K. and Schmid, R. (1961) Acquired porphyria in man and rat due to hexachlorobenzene intoxication. *Nature* 189:449-50.
- Ottevanger, C.F. and Van Sitter, N.J. (1979) Relation between anti-23-hydroxy-Endrin excretion and enzyme induction in workers involved in the manufacture of Endrin. In: *Chemical Porphyria in Man.* (eds. J.J.T.W.A. Strik and J.H. Koeman), Elsevier, N.Y., pp. 123-29.
- Pentz, R. (1978) Influence of diet and xenobiotics on the urinary excretion of bound thiol groups in man and rats. *Naunyn-Schmiedeberg's Arch. Pharmacol.* 302:Suppl. R, 20.

- Notten, W.R.F. and Henderson, P.T. (1975b) The influence of n-hexane treatment on the glucuronic acid pathway and activity of some drug metabolizing enzymes in guinea pig. *Biochem. Pharmacol.* 24:127-31.
- Notten, W.R.F. and Henderson, P.T. (1973) Effect of disulfiram on the urinary D-glucaric acid excretion and activity of some enzymes involved in drug metabolism in guinea pig. *Arch. Int. Pharmacodyn. Ther.* 205:199-208.
- Notten, W.R.F. and Henderson, P.T. (1977a) The Interaction of chemical compounds with the functional state of the liver. I. Alteration in the metabolism of xenobiotic compounds and D-glucuronic acid pathway. *Int. Arch. Occup. Environ. Health* 38:197-207.
- Notten, W.R.F. and Henderson, P.T. (1977b) The interaction of chemical compounds with the functional state of the liver. II. Estimation of change in D-glucaric acid synthesis as a method for diagnosing exposure to xenobiotics. *Int. Arch. Occup. Environ. Health* 38:209-20.
- Notten, W.R.F. and Henderson, P.T. (1975) Alterations in the glucuronic acid pathway caused by various drugs. *Int. J. Biochem.* 6:111-19.
- Ockner, R.K. and Schmid, R. (1961) Acquired porphyria in man and rat due to hexachlorobenzene intoxication. *Nature* 189:449-50.
- Ottevanger, C.F. and Van Sitter, N.J. (1979) Relation between anti-23-hydroxy-Endrin excretion and enzyme induction in workers involved in the manufacture of Endrin. In: *Chemical Porphyria in Man.* (eds. J.J.T.W.A. Strik and J.H. Koeman), Elsevier, N.Y., pp. 123-29.
- Pentz, R. (1978) Influence of diet and xenobiotics on the urinary excretion of bound thiol groups in man and rats. *Naunyn-Schmiedeberg's Arch. Pharmacol.* 302:Suppl. R, 20.

- Roots, I., Saalfrank, K. and Hildebrant, A.G. (1975) Comparison of methods to study enzyme induction in man. *Adv. Exp. Med. Biol.* 58:485-502.
- Schmid, R. (1963) Hepatotoxic drugs causing porphyria in man and animals. *S. Afr. J. Lab. Clin. Med.* 9:212.
- Seutter-Berlage, F., Selten, G.C.M., Oostendrop, S.G.M.L. and Hoog Antink, M.J. (1979) The modified thioether test. In: Chemical Porphyria in Man (eds. J.J.T.W.A. Strik and J.H. Koeman), Elsevier, N.Y., pp. 233-36.
- Seutter-Berlage, F., Van Dorp, H.L., Kosse, H.G.J. and Henderson, P.T. (1977) Urinary mercapturic acid excretion as a biological parameter of exposure to alkylating agents. *Int. Arch. Occup. Environ. Health* 39:45-51.
- Seutter-Berlage, F., Van Dorp, H.L., Kosse, H.J.J., Hoog Antink, J.M.T. and Wagenaars-Zegers, M.A.P. (1979) The estimation of mercapturic acids and other thioethers in urine. In: Chemical Porphyria in Man (eds. J.J.T.W.A. Strik and J.H. Koeman), Elsevier, N.Y. pp. 225-32.
- Seutter-Berlage, F., Wagenaars-Zegers, M.A.P., Hoog-Antink, J.M.T. and Custers, H.A.F.M. (1979) Urinary thioether and D-glucaric acid excretion after industrial exposure to pesticides. In: Chemical Porphyria in Man (eds. J.J.T.W.A. Strik and J.H. Koeman) Elsevier, N.Y., pp. 131-35.
- Simmons, C.J., Davis, M., Dordoni, B. and Williams, R. (1974) Urinary D-glucaric acid assay by an improved enzymatic method procedure. *Clin. Chim. Acta* 51:47-51.
- Sorrel, M.F., Burnett, D.A., Tuma, D.J., et al. (1976) Paradoxical urinary excretion of D-glucaric acid in acute and viral hepatitis. *Clin. Pharmacol. Ther.* 20:265.
- Sotaniemi, E.A., Medzihradsky, F. and Eliasson, G. (1975) Glucaric acid as an indicator of the use of enzyme-inducing drugs. *Clin. Pharmacol. Ther.* 15:417-23.

- Strik, J.J.T.W.A. (1979a) Porphyrins in urine as an indication of exposure to chlorinated hydrocarbons. *Ann NY Acad. Sci.* 390:308-10.
- Strik, J.J.T.W.A. (1979b) The occurrence of chronic hepatic porphyria in man caused by halogenated hydrocarbons. In: Chemical Porphyria in Man, (eds. J.J.T.W.A. Strik and J.H. Koeman), Elsevier, N.Y., pp. 3-9.
- Strik, J.J.T.W.A. (1973) Species differences in experimental porphyria caused by polyhalogenated aromatic compounds. *Enzyme* 16:230-44.
- Strik, J.J.W.A, Doss, M., Schraa, G., Robertson, L.W., VonTiepermann, R. and Hamrjson, E.G.M. Coproporphyrinuria and chronic hepatic porphyria Type A found in farm families from Michigan (U.S.A.) exposed to polybrominated biphenyls (PBB) (In: Chemical Porphyria in Man, (eds. J.J.T.W.A. Strik and J.H. Koeman), Elsevier, N.Y., pp. 29-54.
- Strik, J.J.T.W.A. and Harmsen, E.G.M. (1974) Urinary porphyrin pattern analysis. In: Chemical Porphyria in Man, (eds. J.J.T.W.A. Strik and J.H. Koeman), Elsevier, N.Y., pp. 213-20.
- Strik, J.J.T.W.A., Kip, H., Yoshimura, T., Masuda, T. and Harmsen, E.G.M. Porphyrins in urine of Yusho patients. In: Chemical Porphyria in Man, (eds. J.J.T.W.A. Strik and J.H. Koeman), Elsevier, N.Y., pp. 63-79.
- Strik, J.J.T.W.A. and Koeman, J.H. (1974) The diagnosis and occurrence of chronic hepatic porphyria in man caused by halogenated aromatics (polyprominated biphenyls, poly chlorinated biphenyls and 2,3,7,8-tetrachlorodibenzo-p-dioxin). In: Chemical Porphyria in Man, (eds. J.J.T.W.A. Strik and J.H. Koeman), Elsevier, N.Y. pp. 3-9.
- Tokola, O. Pelkonen, O. Karki, N.T. and Luoma, P. (1975) Hepatic drug-oxidizing enzyme systems and urinary D-glucaric acid excretion in patients with congestive heart failure. *Br. J. Clin. Pharmacol.* 2:429-33.

- Vrij-Standardt, W.G. and Strik, J.J.T.W.A., Ottenvanger, C.F. and Van Sitter, N.J. (1979) Urinary total porphyrin excretion in workers exposed to Endrin. In: Chemical Porphyrin in Man, (eds. J.J.T.W.A. Strik and J.H. Koeman) Elsevier, N.Y. p. 113-121.
- Vainio, H., Salvolainen, H. and Kilpikari, I. (1978) Urinary thioether concentrations of employees of a chemical plant. Br. J. Ind. Med. 35:232-34.
- Van Door, R., Leijdekkers, Ch-M., Bos, R.P., Brouns, R.M.E. and Henderson, P.T. (1981) Enhanced excretion of thioethers in urine of operators of chemical waste incinerators. Br. J. Ind. Med. 38:187-90.

Chapter 8
Asbestos and Cancer
Philip E. Enterline

ASBESTOS AND CANCER

Philip E. Enterline

Human evidence of a cause and effect relationship between asbestos and cancer was available for several decades before it was generally accepted. The reasons for this are partly linked to the manner in which the epidemiological method evolved as it applies to such diseases as cancer. It is now known that asbestos is a causal factor in malignant mesothelioma and lung cancer and probably in cancer of several other sites. A consensus with regard to this was reached only within the last twenty five years, however, despite the fact that some of the earliest observations were made over fifty years ago.

The earliest recorded observation on a relationship between asbestos exposure and lung cancer was by Wood and Gloyne in 1934. They reported two cases of lung carcinoma seen at autopsy in 53 cases of asbestosis. Within the next four years similar case reports appeared in the literature from both the United States and Germany (Lynch and Smith 1935; Egbert and Geiger 1936; Nordman 1938). The Germans quickly arrived at the conclusion that there was a cause and effect relationship and in 1943 the German government issued a decree declaring cancer of the lung, when associated with asbestosis of any degree, an occupational disease. Probably because of World War II, this went unnoticed in the United States and following World War II, German literature and German laws were not very popular outside of Germany and received little attention. The situation was not helped by the fact that Gloyne, who co-authored the original observation of an association, was extremely cautious in interpreting his findings. As early as 1928, Bridge and Henry had proposed conditions that needed to be fulfilled in order that a cancer be classified as industrial in

crigin. These conditions are perhaps not much different than those that would be invoked today. They were 1) the incidence rate in the occupation under review should exceed that in the general population to a significant extent and 2) in the occupation of concern there should be sufficient association of workers with a substance proved experimentally to have carcinogenic properties. Actually in the case of asbestos all of these conditions were not met till 1967 when the first positive animal experiment was reported (Gross, deTreville, Tolker et al. 1967). The lack of positive animal data was not because of a lack of animal experiments. Animal studies using asbestos date back to the 1930's but, for the most part, these were considered to be negative for cancer (Enterline 1978).

Until 1955 nearly all of the published reports were simply case reports of lung cancer and asbestosis seen together at autopsy. Some of the investigators were aware of the selection that is at work in selecting cases for autopsy. For example, Gloyne (1951) was concerned about the fact that the autopsy records at the London Chest Hospital showed 21.3 percent primary lung cancers whereas in the general population at that time only 2.4 percent of all deaths at ages 15 and over were due to lung cancer. Today one would probably say "Berkson's Bias" may have been at work here (Berkson 1946). That is, autopsies are done on a selected group of deaths with selection favoring cases where more than one disease is present.

In 1955, in England, Richard Doll carried out what is probably the first modern epidemiological study of asbestos exposure and cancer. In that study he fulfilled one of the criteria proposed by Bridge and Henry in 1928 by demonstrating that the incidence rate of lung cancer in asbestos exposed workers was considerably higher than the incidence rate in the general

population. There was, however, no accepted positive experimental data and perhaps because of this and because Doll's study dealt with a very small population, many epidemiologists and statisticians were skeptical. For example, in 1956 E. Oyler Hammond of the American Cancer Society and W. Machle in a review of the situation stated that "there are at present too few cases and too little epidemiologic data to establish a significant relationship." In 1958 a second epidemiologic study was carried out consisting of a follow-up of asbestos miners in the province of Quebec and in 1960 a third study was carried out, this time of asbestos workers in California (Braun and Truan 1958, Dunn, Linden and Breslow 1960). In both instances the authors concluded that their findings did not support an association. Thus, as late as 1961 it could be stated in the question and answer section of the Journal of the American Medical Association that "there is no epidemiologic evidence (of an association between asbestos and lung cancer) among American workers " (Johnstone 1961).

The key word here is epidemiological. By 1961, epidemiologists were beginning to have considerable influence and had succeeded in convincing at least some of the medical profession that unless there is a denominator and a comparison population, observations of disease in humans might not be very meaningful. For example, in 1958 the Committee on Occupational Cancer of the AMA Council on Industrial Health concluded that individual case reports contribute little to an understanding of occupational cancers in that "specific increases in age, sex, and type specific cancer incidence above normal incidences for the same age, sex, and site represent the only reliable evidence for the existence of an occupational cancer hazard."

And so in the early 1960's the kind of consensus needed for an aggressive program to control asbestos exposure in order to prevent human cancer had not

been achieved. The turning point apparently came in 1963 when Mancuso and Coulter reported a positive epidemiologic study of workers in an asbestos textile plant followed in 1964 by a report by Selikoff, Chung and Hammond which showed a large excess in respiratory cancer among asbestos insulation workers. The entire situation was pretty much resolved at a conference held by the New York Academy of Sciences in the fall of 1964. However, there still remained some uncertainty because of the inability of investigators to produce cancer experimentally.

There is a lesson to be learned by reviewing the early history of the asbestos-cancer controversy. For one thing, the evolution of modern epidemiology and biostatistics as applied to cancer tended to drive out other forms of observation on human disease. In particular, case reports were numerous and persistent and should have played a larger role in resolving the issue. For another, there should have been more awareness of what was happening elsewhere in the world. Somehow being an American during the 1940's and 1950's was equated to being different and ideas originating outside of America were not readily accepted. That attitude may still persist in occupational epidemiology. Clearly the decision by the Germans to declare asbestosis related lung cancer an occupational disease in 1943, which incidentally was reaffirmed in West Germany in 1950, should have had more influence on attitudes toward the question of asbestos and cancer (Tabershaw 1951). Moreover, in England researchers were fairly certain that there was an association based on the work by Richard Doll already cited, but also because of an earlier report published by the Chief Inspector of Factories in 1949 which showed that in fatal cases of asbestosis 13.2 percent had cancer whereas in fatal cases of silicosis only 1.3 percent had cancer (Merewether 1949).

The story thus far is about the difficulties in finally arriving at a consensus regarding asbestos and cancer. This is, however, only the first step in health risk assessment. There remained a problem of how to quantify risk to human health and how to establish dose-response relationships. Epidemiologists and biostatisticians have a great deal to offer in this regard. There has been monitoring of asbestos dust levels in occupational environments dating back to the 1930's when it was clearly established that levels of asbestos dust and the risk of asbestosis were highly related. In the late 1960's there was an opportunity to examine data from the Johns Manville Corporation relating to dust levels for various jobs in their asbestos products producing plants along with a file of all retirements from the company starting in 1941, including work histories for each retiree. An analysis of these data were undertaken.

Figure 1 illustrates the kind of information that epidemiologic studies can produce and which can be useful in quantifying risk (Enterline, DeCoufle and Henderson 1973). It also raises a number of important issues, including whether a straight or a curved line should be fitted to such data and whether the measure of exposure used, which in this case gives equal weight to the dose rate and the duration of exposure, is appropriate.

If it is possible to gain the cooperation of industry, studies of this type are simple to conduct. An earlier attempt had been used using only records on file at the Social Security Administration (Enterline 1965). Unfortunately, those records contained no information on duration or level of exposure. Nor did they show in all cases when the exposure began. The usefulness of such a data base is therefore limited. Records maintained by labor unions are perhaps better but also have some of the same disadvantages. The best data base for occupational epidemiology is clearly records maintained by the companies whose employees are being studied.

Enterline

--5--

Data were recently updated on retired workers from the Johns Manville Corporation and are presented at this time. Table 1 summarizes the follow-up status of 1,074 white males who retired from the Johns Manville Corporation during the years 1941 through 1967 and who were employed as production or maintenance workers by the company prior to retirement. This table shows data relative to the current update as well as data relative to two previous data sets, one set first published in 1972 dealing with a follow-up for deaths through 1969 and one published in 1979 dealing with a follow-up through 1973 (Enterline, DeCoufle and Henderson 1972, Henderson and Enterline 1979). Follow-up for death started at age 65 or at the age of retirement if it started after 65 years of age. Included are persons retiring prior to age 65 either because of disability or other reasons but who lived to be age 65.

What is perhaps unusual about this study is that nearly everyone is now dead and, insofar as information on death certificates is reliable, the disease responsible for each death is known. Most other studies are highly censored in this regard. That is, in most occupational epidemiologic studies, only a small proportion of the population of interest has died and there is no way of knowing whether the distribution of deaths in this proportion is predictive of the distribution likely to be observed among those still alive. As shown in Table 1 by 1980 about 88 percent of the cohort of retired asbestos workers were known to have died and death certificates were obtained for nearly all of these.

Table 2 shows standardized mortality ratios (SMR) for causes of death that are of particular interest with regard to asbestos exposure. One item of interest is the fact that relative mortality rates for all causes appears to be climbing slightly as the population ages showing perhaps the long term effect

of the exposure which took place many years earlier. All of the excesses shown for the period 1941 through 1980 are statistically significant ($P < .05$) and it is highly likely that these excesses are due mainly to exposure to asbestos.

Some earlier reports did not look at cause of death in any great detail. Since their publication however, there has been an interest in kidney cancer and two studies by others have shown an excess (Putoni, Vercelli, Merlo et al. 1979, Selikoff, Hammond and Seidman 1979). There does appear to be a sizable excess in the death rate for kidney cancer among asbestos exposed retirees. There is a very high SMR for eye cancer which does not seem to have been reported elsewhere (Table 2). However, this is based on only two deaths and may have little or no meaning.

Not shown in Table 2 is the experience of this population with regard to malignant mesothelioma and asbestosis, two important consequences of asbestos exposure. Unfortunately malignant mesothelioma is not identifiable in the Seventh Revision of the International Lists, the revision to which deaths in this retiree study were coded (World Health Organization, 1957). A review of the death certificates, however, shows eight deaths in this cohort of 1074 white males probably due to malignant mesothelioma. Four of these were definite pleural mesotheliomas, one was probably a pleural mesothelioma, two were definite peritoneal mesotheliomas, and for one the site was not given. Asbestosis is codable as a cause of death but SMR's are meaningless in studies of this type due to the rarity of this disease in the general population. In this cohort, of the 86 deaths due to non-infectious pulmonary disease shown in Table 2, 22 were due to asbestosis.

Table 3 shows observed numbers for all causes of death and standardized mortality ratios by exposure levels expressed as in Figure 1, for the three follow-up periods. Clearly there is a good relationship between the cumulative

Enterline

-7-

dust exposure at time of retirement and the subsequent all cause death rate. This relationship is very much the same for each of the three follow-up periods. There is some indication that the death rates in the lower exposure groups have increased over these periods whereas in the high exposure groups death rates appear to be fairly constant. Adjustment has been made in Table 3 and in previous tables for age and calendar time.

Table 4 shows data comparable to those in Table 3 except they are confined to cancer of the lung. Again there is a fairly consistent dose-response relationship among the report periods. Interestingly death rates in the highest exposure group declined slightly over these three report periods while death rates in the two next lower exposure categories appear to have increased slightly. If it is true that death rates in the highest exposure group leveled off while those in lower groups continue to climb, it would support the notion that time to tumor is inversely related to the intensity of the exposure—often called the Druckrey effect described in some detail by Jones and Grendon (1975). Such a phenomena, if it exists in humans, would greatly influence quantitative risk assessment from epidemiologic data.

From data such as those shown in Table 4 it is possible to quantify the risk associated with asbestos exposure. This is shown graphically for asbestos in Figure 2 with a straight line fitted to the data. Clearly the amount of dust to which the retired asbestos products workers in this cohort had been exposed at time of retirement was an excellent predictor of their subsequent mortality experience from cancer of the lung. The linear regression line fitted to the data tells us that one year's exposure to one million particles per cubic foot of asbestos dust, eight hours a day, five days a week, will cause about a half a point increase in the standardized mortality ratio for lung cancer.

Enterline

-8-

At one time the occupational standard for asbestos exposure was expressed in terms of million particles per cubic foot—as presented here. At that time the exposure limit was five million particles per cubic foot. From data presented in Figure 2 it can be easily calculated that forty years exposure at the old standard would increase the standardized mortality ratio by about 100 points, if the standardized mortality ratio in the absence of exposure was 100. That is, if the number of deaths observed were just that which would be expected based on the experience of the United States population of the same age living during the same period of time, then one could say that the old standard of five million particles per cubic foot permitted a doubling of the risk of dying of lung cancer.

The standard for asbestos exposure is no longer expressed in terms of particles, but rather in terms of fibers and conversion of the exposure data from particle counts to fiber counts presents some complications. The purpose of this chapter has been to illustrate how fairly straightforward observations of human populations can produce a quantification of risk to human health, not to discuss some of the inherent complications.

DISCUSSION

The data presented involving a cohort of asbestos workers, nearly all of whom are now dead, seems particularly useful in the context of risk assessments currently being made by the Environmental Protection Agency for various carcinogens. These risk assessments are based on the fact that since we all die, risks expressed as numbers of cases per thousand of deaths are equivalent to numbers of cases per thousand population exposed. Clearly the entire mortality experience of a population is needed in order to implement a concept of this nature.

With regard to the question of how epidemiology has contributed to health risk assessment for cancer a two stage process has been described. First there must be acceptance of a causal relationship between a substance and a disease, and second some quantification of the relationship is needed. Of course the second contributes strongly to the first since perhaps the most important evidence of a causal relationship is the demonstration of a dose-response relationship.

The extent to which the experience with asbestos relates to problems faced today and in the future may be questioned. For one thing a great deal has been learned about how to properly conduct and interpret epidemiologic investigations and such investigations are now being aggressively carried out. In retrospect it would seem that good studies of asbestos workers, like the study carried out by Richard Doll in 1955, could have resolved the issue of asbestos as a cancer hazard considerably earlier than was actually the case. Also, techniques for producing cancer in animals have improved considerably since the early attempts with asbestos and it is unlikely that so many years will pass before true carcinogens can be demonstrated in laboratory animals. Positive animal data should definitely influence the interpretation of epidemiological data. Finally, we are now living in an age when it is believed that most cancer is preventable and this encourages positive interpretations of associations seen in epidemiologic and in other studies.

Looking back on the asbestos literature it is now possible to see positive relationships that the authors then interpreted as negative. Of course in doing this there exists the benefit of all that came after these papers were written. Even today, however, there is still much to be learned about the interpretation of data. Surely whoever looks back in time from, say, the year 2010 will see many things that were seen but misinterpreted in the year 1985.

Enterline

REFERENCES

- Berkson, J. (1946) Limitations of the application of fourfold table analysis to hospital data. *Biometric Bull.* 2:47-53.
- Braun, D.C. and Truan, T.D. (1958) An epidemiological study of lung cancer in asbestos miners. *Arch. Ind. Health* 17:634-53.
- Bridge, J.C. and Henry, S.A. (1928) Industrial cancers. In Report of the International Conference on Cancer, London, July, 1928, William Wood & Co., Baltimore.
- Council on Industrial Health. (1958) Committee on Occupational Cancer. *JAMA* 166:2171.
- Doll, R. (1955) Mortality from lung cancer in asbestos workers. *Br. J. Ind. Med.* 12:81.
- Dunn, J.E., Linden, G. and Breslos L. (1960) Lung cancer experience of men in certain occupations. *Am. J. Public Health* 50:1366-75.
- Egbert, D.S. and Geiger, A.J. (1936) Pulmonary asbestosis and carcinoma: report of a case with necropsy findings. *Am. Rev. Tuberc.* 34:143.
- Enterline, P.E. (1978) Asbestos and cancer, The international lag. *Am. Rev. Respir. Dis.* 118:975-78.
- Enterline, P.E. (1965) Mortality among asbestos products workers in the United States. *Ann. N.Y. Acad. Sci.* 132:156-65.
- Enterline, P.E., DeCoulfle, P. and Henderson, V. (1972) Mortality in relation to occupational exposure in the asbestos industry. *J. Occup. Med.* 14:897-903.
- Enterline, P.E., DeCoulfle, P. and Henderson, V. (1973) Respiratory cancer in relation to occupational exposures among retired asbestos workers. *Br. J. Ind. Med.* 30:162-66.

Enterline

- Gloyne, S.R. (1951) Pneumoconiosis: A histological survey of necropsy materials in 1205 cases. *Lancet*: 1,810.
- Gross, P., deTreville, R.T.P., Tolker, E.B., Kaschak, M. and Babyak, M.A. (1967) Experimental asbestosis: The development of lung cancer in rats with pulmonary deposits of chrysotile asbestos dust. *Arch. Environ. Health* 15:343-55.
- Hammond, E.C. and Machle, W. (1956) Environmental and Occupational Factors. In, Pulmonary Carcinoma: Pathogenesis, Diagnosis and Treatment, (eds. E. Mayer and H.C. Maier), New York University Press, New York.
- Henderson, V. and Enterline, P.E. (1979) Asbestos exposure: Factors associated with excess cancer and respiratory disease mortality. *Ann. N.Y. Acad. Sci.* 330:117-26.
- Johnstone, R.T. (1961) Silicosis and cancer (Letters to the Editor). *JAMA* 178:352.
- Jones, H.B. and Grendon, A. (1975) Environmental factors in the origin of cancer and estimation of the possible hazard to man. *Fed. Cosmet. Toxicol.* 18:251-68.
- Lynch, K.M. and Smith, W.A. (1935) Pulmonary asbestosis III: Carcinoma of the lung in asbesto silcosis. *Am. J. Cancer* 24:56.
- Mancuso, T.F. and Coulter, E.J. (1964) Methodology in industrial health studies. *Arch. Environ Health* 6:210-26.
- Merewether, E.R.A. (1949) Annual report of the Chief Inspector of Factories for the year 1947, HMSO, London.
- Nordmann, M. (1938) Der berufskrebs der asbestarbeiter. *Ztschr. Krebsforsch* 47:288.
- Puntoni, R., Vercelli, F., Merlo, F., Valerio, F. and Santi, L. (1979) Mortality among shipyard workers in Genoa, Italy. *Ann. N.Y. Acad. Sci.* 330:353-77.

- Selikoff, I.J., Churg, J. and Hammond, E.C. (1964) Asbestos exposure and neoplasia. JAMA 188:22-6.
- Selikoff, I.J., Hammond, E.C. and Seidman, H. (1979) Mortality experience of insulation workers in the United States and Canada. Ann. N.Y. Acad. Sci. 330:91-116.
- Tabershaw, I.R. (1951) Report on the industrial hygiene in the Western Zone of Germany. Arch. Ind. Hyg. Occupa. Med. 3:298.
- Wood, W.B. and Gloyne, S.R. (1934) Pulmonary Asbestosis: A review of one hundred cases. Lancet 2:1383.
- World Health Organization (1957) International Classification of Diseases, Geneva.

TABLE 1: Summary of results of three follow-up studies of retired asbestos products workers

	Follow-up Period		
	1941-69 ^a	1941-73 ^b	1941-80 ^c
Total in cohort	1074	1074	1074
Dead	617	782	944
Death certificate found	608	749	910
Death certificate not found	9	18	34
Presumed alive	457	292	130

Source: ^aEnterline, DeCoufle and Henderson (1972)
^bHenderson and Enterline (1979)
^cUnpublished data (1985)

TABLE 2: Standardized mortality ratios (SMR) for selected causes of death for retired asbestos products workers

Cause of death	Follow-up Period		
	1941-1969 ^a (SMR)	1941-1973 ^b (SMR)	1941-1980 ^c (SMR)
All causes	116.5	121.6	123.8
All cancer	154.7	160.8	160.2
Respiratory	259.4	257.5	258.4
Lung	270.8	268.4	270.7
Digestive	136.1	141.4	140.0
Kidney	282.9	322.6	275.8
Eye	2008.5	1666.7	1544.5
Non-infectious pulmonary disease	257.0	240.4	234.9
Heart disease	108.4	109.4	111.6

Source: ^aEnterline, DeCoufle and Henderson (1972)
^bHenderson and Enterline (1979)
^cUnpublished data (1985)

TABLE 3: Numbers of deaths and standardized mortality ratios (SMR) by exposure level and study period for all causes of death of retired asbestos products workers

Exposure Level MPPCF ^d - years	Follow-up Period					
	1941-1969 ^a		1941-1973 ^b		1941-1980 ^c	
	Observed	SMR	Observed	SMR	Observed	SMR
<125	259	108.8	315	110.1	376	111.7
125-249	113	109.2	153	117.9	193	120.3
250-499	146	118.0	194	127.7	239	132.8
500+	99	153.8	12.0	159.3	136	158.7

Source: ^aEnterline, DeCoufle and Henderson (1972)

^bHenderson and Enterline (1979)

^cUnpublished data (1985)

^dMillion particles per cubic foot

TABLE 4: Numbers of deaths and standardized mortality ratios (SMR) by exposure level and study period for cancer of the lung of retired asbestos products workers

Exposure Level MPPCF ^d - years	Follow-up period					
	1941-1969 ^a		1941-1973 ^b		1941-1980 ^c	
	Observed	SMR	Observed	SMR	Observed	SMR
125	16	212.3	19	196.5	23	196.5
125-249	7	186.7	9	177.9	14	217.4
250-499	13	293.6	19	327.0	23	331.3
500+	13	546.2	16	546.1	17	507.5

Source: ^aEnterline, DeCoufle and Henderson (1972)

^bHenderson and Enterline (1979)

^cUnpublished data (1985)

^dMillion particles per cubic foot

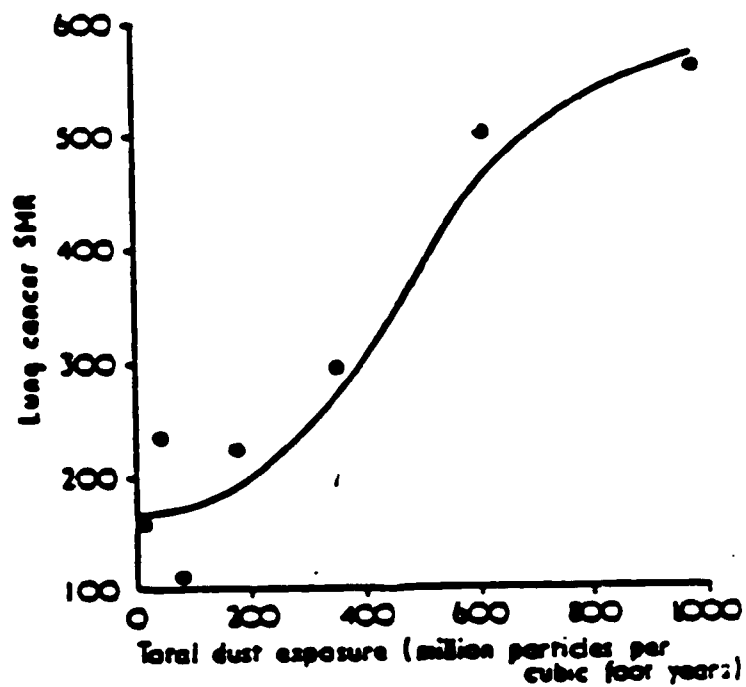


Fig. 1 - Total asbestos dust exposure and lung cancer mortality of retired asbestos products workers

Source: Enterline, DeCoufle and Henderson (1973).

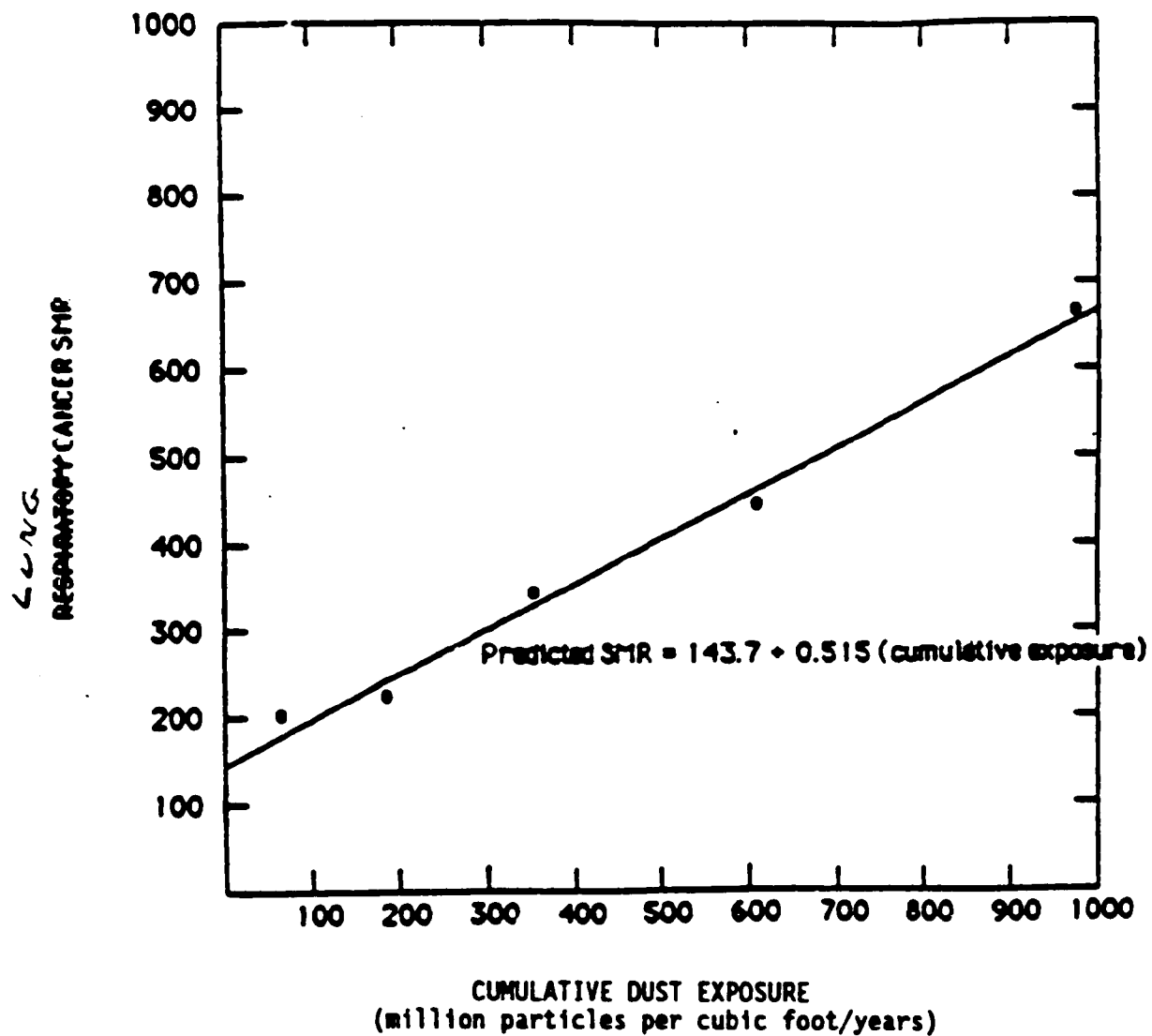


Fig. 2 Total asbestos dust exposure and lung cancer mortality of retired asbestos products workers

Source: Unpublished data (1985)

Chapter 9

Genetic Susceptibility and the Estimation of Risk

Gilbert S. Omenn

GENETIC SUSCEPTIBILITY AND THE ESTIMATION OF RISK

Gilbert S. Omenn

Introduction

Individual differences in susceptibility to potentially hazardous chemicals represent a significant parameter in the characterization of risks to human health from such chemicals. Major environmental laws, such as the Clean Air Act and the Occupational Safety and Health Act, require that standards be set to protect individuals who are highly susceptible to the particular agent being regulated. In fact, the research agenda has not yet responded significantly to this regulatory mandate. Nevertheless, we know that genetic differences in the metabolism of chemicals and genetic differences in tissue responses to chemicals should be analyzed in risk assessments and may provide useful clues to the mechanisms of toxicity of specific chemicals. At the same time, we must emphasize that reduction or elimination of exposures to toxic chemicals on the job and in the general environment must be the objective of risk reduction and disease prevention efforts. Genetic screening, even when reliable and relevant tests might become available, should not be allowed to substitute for continued progress in control of exposures.

Framework for Risk Assessment

Several years ago my colleagues and I outlined a framework for the identification, characterization, and control of risks from potentially hazardous chemicals (Calkins, Dixon, Gerber et al. 1980) (Table 1). The distinctions between risk assessment (first and second step) and risk management (third step) have been discussed extensively in a report from the National Research Council (1983) on Risk Assessment in the Federal Government: Managing the Process. The phrase "characterization of risk" seems more appropriate than the commonly used phrases risk assessment quantitative risk assessment. The term characterization conveys more clearly the need to combine qualitative, subjective, and interpretive information with quantitative extrapolations from animals to humans, from high dose to low dose, from one

route of exposure to another, and from persons of "average" susceptibility to those of greater susceptibility. It is essential to describe the nature and severity of the adverse health effects, the strength of the overall evidence, the choices of empirical data from which extrapolations will be made, the sources and time-dose relationships of exposures, and, finally, the differing susceptibility of various definable subgroups in the population.

Quantitation is desirable whenever the data permit reasonable estimates. Furthermore, it should be noted that tests carried out simply to infer potential hazard from a particular substance depend upon analyses of quantitative data. Often the "hazard-identification" step is inappropriately considered to involve a yes or no decision. Unfortunately for this simplistic view, there is no qualitative difference between "positive" and "negative" results in epidemiological studies, bioassays, or in vitro assays. These terms apply, according to statistical conventions, to the quantitative results in animals or humans in an exposed group, compared with frequencies of the same effects in unexposed or control groups. Obviously, the number of subjects, the frequency in the control group, and the statistical criteria are crucial variables.

It is essential to be explicit in describing whether the estimated risk or estimated incidence is an observed frequency, a best estimate, or an estimate of the upper bound defined statistically by a 95 percent, or 99 percent, confidence limit. The credibility and comprehensibility of such statistical conclusions are enhanced tremendously when the evidence under review is biologically or clinically plausible (Omenn 1984b). The effects should be related to dose; the route and timing of exposures should have reasonable relationships to the sites and kinds of effects observed. However, the most compelling evidence comes from knowledge about the mechanism of action of the

toxic agent, the metabolic pathways for activation and/or inactivation, the target sites in particular organs and within cells, and the basis for variation in host responses. Evidence of mechanisms of toxicity is important not only for carcinogenic effects, but also for mutagenic, teratogenic, organ-specific, and behavioral effects.

Impetus to Bridge Fields of Genetics and Environmental Health

The factor termed "range of susceptibility" under characterization of risk in Table 1 subsumes a variety of potentially important host factors. These include: genetic traits, age, sex, pre-existing diseases, behavioral traits, co-existing exposures, medications and vitamins, and individual protective measures and work practices. Figure 1 demonstrates schematically that the overall population dose-response curve may mask a highly susceptible subgroup, or several subgroups, in the population. In fact, the marked variation among humans is a primary justification for utilizing data from highly sensitive animal species in order to protect susceptible humans.

This chapter will focus on genetic variation as a component of host factors influencing the characterization of risks in human populations. The impetus to bridge the fields of human genetics and environmental health comes from clinical, scientific, and policy objectives.

From a clinical point of view, workers with health complaints attributed or possibly attributed to occupational exposures often ask, "Why me, Doc? Everyone else works in the same conditions, and I don't believe they are any more or less careful than I am." It is good occupational medicine practice to attempt to answer such a question, while more generally exploring conditions of exposure to find ways to decrease risks for everyone. Physicians often cite the shibboleth about "tailoring the treatment to the patient", but most drug

regimens are prescribed routinely. Choosing the right treatment may depend upon knowing the patient's responses, or even the etiologic mechanism. Finally, more and more patients are asking physicians whether some unexplained medical problems, from birth defects to arthritis to mental illness, might be due to particular exposures. In the absence of obvious effects in a high proportion of the population, one must still wonder whether some individuals or subgroups are sufficiently susceptible that such an inquiry may be a useful clue to investigation.

From a scientific point of view, understanding the mechanisms or variation in biotransformation of chemicals can advance our knowledge of the mechanisms of toxicity from those agents. To a geneticist, this approach takes advantage of nature's experiments and capitalizes on the mutant analysis approach which has been so valuable in many organisms, including humans, in revealing hitherto obscure metabolic pathways and intermediates. Recognition of genetic heterogeneity also is a reminder to think of mechanisms in the plural, rather than attempting to force all seemingly similar effects into a single pathogenetic pathway with a single preventive or curative control measure.

Finally, from the point of view of regulatory policy, it must be noted that the Occupational Safety and Health Act (OSHA) requires that standards be set so that "no worker shall suffer any adverse effect", even if exposed at the level of the standard for a full working lifetime... allowing for technical feasibility of the proposed standard. OSHA proceedings generally have neglected individual differences in susceptibility, except for differences due to previous exposures to the same agent, as in the case of lead. In contrast, under the requirements of the Clean Air Act that national ambient air quality standards be set to protect even the "most susceptible subgroups" in the

population, the Environmental Protection Agency (EPA) has given explicit treatment to the identification of groups most likely to be highly susceptible (Friedman 1979). For the ozone (photochemical oxidant) standard, these were individuals with asthma or emphysema (chronic obstructive lung disease); for the airborne lead standard, these were young children (Omenn 1979). The EPA declared that the proposed standards would be protective for such subgroups, and with an adequate margin of safety, though there was little research to confirm these expectations until recently (Koenig, Pierson, Horike et al. 1983). Smaller subgroups of individuals with inherited predispositions to chronic lung disease, such as children with cystic fibrosis (one in 2000 births in the white population) or alpha-1 antitrypsin deficiency (about 1 in 5000 persons), could be investigated, as well. In general, it must be admitted that research on subgroup susceptibility lags far behind regulatory inferences about which groups might be especially susceptible to particular exposures.

The Emerging Field of Eco-Genetics

The relationships between genetic and environmental agents embrace to complementary phenomena: inherited variation in susceptibility to exogenous agents (eco-genetics) and signs of damage to the genetic apparatus from previous environmental exposures (genetic toxicology). Genetic toxicology is a rapidly growing field, with an array of tests and endpoints, both biochemical and chromosomal (Bloom 1981, Omenn 1983, Bridges, Butterworth, Weinstein 1982). This field is likely to be transformed as new work on oncogenes and DNA repair is incorporated into studies of the effects of exogenous agents (Omenn 1984a).

Previous genetic damage may increase susceptibility to subsequent exposures, so genetic toxicology and eco-genetics are clearly interrelated. However, this discussion will focus on genetic traits that are inherited as recessive or dominant single gene effects (Mendelian inheritance). It is believed that polygenic inheritance and genetic/environmental interactions, including the role of nutrition, will prove even more important than single gene traits; for example, we have reviewed genetic predispositions to non-malignant pulmonary diseases (Koenig and Omenn, in press). Such complex interactions are more difficult to analyze at this stage of knowledge, just as effects of interactions of mixtures of chemicals are more difficult to analyze than the effects of chemicals studied alone.

It is important to recognize the uniqueness of individuals, which is readily revealed by profiles of biochemical traits that are genetically polymorphic (at least one percent of the population having a less common variant than the usual). Tools of biochemical and molecular genetics show variation and polymorphism at every level of investigation: blood groups; plasma proteins; plasma and red blood cell (and other tissue) enzymes; HLA antigens (tissue types); chromosomal banding patterns; and DNA sequence analysis.

Genetic predispositions to environmental agents are well known for quite a variety of agents, including drug chemicals, pesticides, inhaled pollutants, foods and food additives, sensory stimuli, allergic and sensitizing agents, and, of course, infectious agents (Table 2). These traits have been reviewed elsewhere (Omenn and Motulsky 1978, Omenn and Gelboin 1984, Omenn 1982). Much more research has been carried out on differences in metabolism of chemicals prescribed as drugs and differences in target tissue sites of action for

specific drugs than on the several other classes of environmental agents. Nevertheless, several of the underlying mechanisms revealed by drugs surely are relevant to differences in susceptibility to non-drug chemicals encountered in general and occupational environments.

The descriptions which follow deal with potentially important eco-genetic traits that are common in the general population, i.e. frequent polymorphisms. Other much less frequent genetic traits are known which predispose to adverse reactions from radiation or chemical-induced carcinogenesis (Shaikh, Warren and Little, in press). Xeroderma pigmentosum is a rare autosomal recessive condition characterized clinically by extreme hypersensitivity to sunlight, leading to multiple cancers, including melanomas, of the exposed skin. These patients and their cells in vitro are unable to repair certain kinds of damage to the DNA, when excision repair of UV light-induced thymine dimers or of chemical cross-links induced by bifunctional alkylating agents, crosslinking agents, or polycyclic hydrocarbons is needed. Sunlight-exposed areas like the anterior eye and the anterior tongue also have higher risks of cancer, but apparently internal sites do not. Heterozygotes with the XP gene and half-normal repair activity seem to show no demonstrable effects in vivo or in vitro. Hereditary retinoblastoma and nevoid basal cell carcinoma syndrome are both autosomal dominant disorders with increased susceptibility to x-irradiation (Strong, 1985). Retinoblastoma patients develop multicentric malignancies in the eye and are at high risk for radiotherapy-induced osteogenic sarcoma (up to 40 percent of patients). Mitomycin-C and UV light induce sister-chromatid exchanges in retinoblastoma cells in vitro. Patients with nevoid basal cell carcinoma syndrome develop multiple basal cell cancers of the skin, plus medulloblastomas and other brain tumors. Mechanisms of these

effects are unknown. Finally, two other rare autosomal recessive disorders are of potential interest. Ataxia teleangiectasia cells are very sensitive in vitro to ionizing radiation; although no epidemiologic evidence suggests that environmental or occupational radiation triggers the leukemias, lymphomas, or other tumors in these patients, radiotherapy can be fatal. Cells from patients with Fanconi's anemia have marked sensitivity to DNA cross-linking agents, though no such exposures have been implicated in the cases of leukemia among these patients.

Potential Genetic Predispositions Due to Differences in Metabolism

1. The acetylator phenotype

A single-gene difference affecting N-acetyl-transferase activity in liver makes individuals slow acetylators (recessive) or rapid acetylators (dominant) of a series of clinically useful drugs: the anti-tuberculosis agent isoniazid (INH), the anti-high blood pressure agent hydralazine, the anti-arrhythmic agent procainamide, various sulfa drugs including dapsone, and an anti-depressant MAO inhibitor, phenelzine. Adverse effects of these drugs tend to occur more frequently, or at common doses only, in the slow acetylators, who build up higher blood levels on a given dose of drug. An important drug/drug/genetic interaction occurs in slow acetylators given INH and already taking phenytoin (Dilantin). Often the prescribing physician is unaware that the patient also is taking phenytoin; therefore, the development of side effects that accompany phenytoin usage (nystagmus, ataxia, somnolence) may be perplexing. The explanation is that in slow acetylators INH reaches concentrations sufficient to inhibit drug-metabolizing enzymes in the liver, including those responsible for phenytoin levels and phenytoin toxicity. Such interactions between chemicals surely occur also in the workplace and between prescribed drug chemicals and environmental exposures.

The N-acetyltransferase is known to be involved in acetylation to inactivate arylamines that are potent bladder carcinogens. These compounds are among the most certain occupational carcinogens. Examples are 2-naphthylamine, benzidine (4,4'-diaminobiphenyl), 4-amino-biphenyl, and 4-nitrobiphenyl, all of which have been used as industrial antioxidants; and chlornaphazine, an anticancer drug. Lower, Hilssan and Nelson (1979) carried out an epidemiologic study in Scandinavia to seek evidence on whether slow acetylators would have a higher risk of bladder cancer, as predicted from their slower rate of inactivation of the agent. In this study, not all bladder cancer cases represented occupational exposures, no occupational data were available; and the role of arylamines would have to be differentiated from other exposures. Any effect of the N-acetyltransferase variation would be diluted by cases due to unrelated etiology. Nevertheless, the study was suggestive that the urban Danish population, presumably enriched for occupational bladder cancers, had an excess of slow acetylators, compared with Danish controls without bladder cancer (67% versus 51%). A rural Swedish population of bladder cancer patients had negligibly higher frequency of slow acetylators than did controls (70% versus 67%). Unfortunately, the control group frequencies were at the extremes of the range observed in Caucasian populations.

There are two major avenues open to improve such a study. First, as is frequently the case, the distribution of phenotypes for genetic polymorphisms differs in different populations. While white and black populations have 50-60 percent slow acetylators, the Japanese have only 10-15 percent. In this population, an enrichment of slow acetylators should be much more readily

detectable. Matsushima, Yagishita and Ando (1981) at the Third International Conference on Environmental Mutagens described an ideal study population, in which long-term follow-up has already been established for workers with high exposures to dyestuff chemicals in past decades. Hopefully, acetylator phenotyping will be incorporated into their protocol.

Meanwhile, the alternative approach of starting with good information about occupational exposures has been exploited successfully (Cartwright, Glashan, Rogers, et al 1982). The prevalence of bladder cancer in Huddersfield, a heavily industrialized town in the Yorkshire region of England, is twice that in the United States and among the highest known in the world. Before 1952, workers in the dye manufacturing trade were exposed to 2-naphthylamine and up to 1960 they were exposed to benzidine. As in the Scandinavian study, there was only a nonsignificant excess of slow acetylators among the total group of Huddersfield bladder cancer patients (67% versus 57% among controls). However, among dye-workers with bladder cancer, 22 of 23 were slow acetylators, in contrast to patients who were clerical workers or engineers using machine oils (40-44%). The only fast acetylator among dye-workers had an atypical type of bladder cancer (adenocarcinoma). There was no excess of slow acetylators among those with T1 or T2 histological types of tumors, whereas slow acetylators predominated among those with the more invasive T3/T4 tumors and carcinoma in-situ. Interestingly, there is no correlation with cigarette smoking, which is an etiological factor associated with the T1/T2 bladder cancers. It would be desirable to attempt to reconstruct exposure histories in order to gain enough information to estimate attributable risks and genetic/environmental interaction factors. Clearly, however, acetylator phenotyping might be useful in determining the frequency of follow-up for workers already subjected to these exposures and for assessment of other workers exposed to other agents which require acetylation for inactivation.

2. Microsomal oxidizing systems/debrisoquine hydroxylation.

A remarkable new drug-metabolizing polymorphism has been investigated in the past few years. Alicyclic hydroxylation of debrisoquine, aromatic hydroxylation of guanoxan, oxidative dealkylation of phenacetin, N-oxidation of sparteine, and oxidations of a growing list of other agents all are affected by a single-gene difference in activity (Omenn and Gelboin 1984, Idle and Smith 1979, Dengler, Eichelbaum and Meyer 1984). This polymorphism is common, with up to 10 percent of the population being extremely poor metabolizers, and the activity difference is large. When expressed as a ratio of unchanged debrisoquine to hydroxylated metabolite, a difference of 200 to 20,000 fold is observed. It is not clear whether a single catalytic site in a single mutated protein of the cytochrome P450 system can account for the broad substrate specificity, but Davies has shown competitive inhibition among debrisoquine, bufuralol, sparteine, phenformin, and guanoxan (Dengler, Eichelbaum and Meyer 1984).

Using cDNA probes prepared from mRNAs for cytochrome P450 isozymes, several investigators have identified and begun to characterize the P450 genes. The P450 responsible for steroid 21-hydroxylation in bovine adrenal cortex (probably corresponding to the enzyme deficient in humans with congenital adrenal hyperplasia) has been mapped in close proximity to the HLA gene cluster on human chromosome 6 (White, New and Dupont); the gene for the debrisoquine and sparteine polymorphism may be linked to the genetic marker Gc on chromosome 4, and the gene for cytochrome P448 is on chromosome 2 (Dengler, Eichelbaum and Meyer 1984). Much more work is needed to combine the DNA, mRNA, protein, substrate, and inhibitor information into a useful understanding of the constitutive and induced levels of drug-metabolizing activities for various agents.

From the occupational and environmental point of view, it is irresistible to inquire what other chemical conversions might be tied to the debrisoquine metabolism polymorphism. Many pro-carcinogenic compounds require such metabolic activation. A report has appeared suggesting that Nigerians with liver cancers possibly induced by aflatoxin exposures include a lower-than-control frequency of cases with poor metabolism of debrisoquine (Idie, Mahgoub, Sloan et al, 1981). This finding fits the hypothesis that poor metabolizers would be less capable of metabolic activation of aflatoxin; however Davies has presented evidence which makes it unlikely that the debrisoquine polymorphism controls aflatoxin activation (In: Omenn and Gelboin 1984).

Much more work of interest can be expected in this system. Larrey, Distlerath, Dannan, et al. (1984) have isolated a rat liver P450 isozyme involved in debrisoquine hydroxylation, as well as progressing toward isolation of the human isozyme (Distlerath and Guengerich 1984). Kupfer has reported a separate polymorphism affecting the aromatic hydroxylation of mephenytoin, with five percent of the Swiss population tested being defective in this oxidation step (Dengler, Eichelbaum and Meyer 1984). Also, we should be alert that genetic associations may differ in different populations. For example, Woolhouse, Eichelbaum, Oates et al (1985) have reported that the regulatory control of debrisoquine/phenformin oxidation seems to be dissociated from control of sparteine oxidation among Ghanaians.

Related to the cytochrome P450 metabolizing systems is a previously described aryl hydrocarbon hydroxylase (AHH). AHH activity is known to be involved in activating many promutagenic and procarcinogenic aryl hydrocarbons, of which benzo(a)pyrene is the prototype. It is clear that genetically

determined differences in susceptibility to important environmental effects can be attributed to a homologous AHH inducibility locus (Ah) in the mouse (Nebert 1978). Kellerman, Shaw, and Luyten-Kellerman (1973) claimed that humans could be differentiated into high, intermediate, and low inducers of AHH activity in peripheral lymphocytes in vitro, that these differences were determined by a single gene effect according to family studies, and that high inducers had a much higher risk of lung cancer, presumably induced by aryl hydrocarbons in cigarette smoke and air pollution. Such results could have had major applications for risk assessment in cigarette smokers and possibly for non-smokers in certain occupations. Unfortunately, the test has been technically complicated and unreliable in others' hands. However, research is moving forward on this type of assay (Omern and Gelboin 1984).

3. Plasma Paraoxonase Activity/Parathion Inactivation

The pesticide parathion and its active metabolite paraoxon are inactivated by a hydrolytic enzyme in normal plasma, paraoxonase (Mueller, Hornung, Furlong, et al. 1983; see Omerns and Gelboin, 1984). Paraoxon inhibits acetylcholinesterase at neuromuscular junctions. About half of the white and black populations in the United Kingdom and in Seattle have a low-activity form of the enzyme, with one-third the activity of those with the high-activity form. These activity differences may be sufficient to cause significant differences in inactivation rates and blood levels for this toxic pesticide and possibly for such congeners as malathion/malaoxon. These pesticides are very widely used throughout the world. Studies of spray pilots, applicators, mixers, field hands, and general population would be desirable. Those at highest routine exposures are applicators and mixers. The hypothesis is that those who have low paraoxonase activity will be predisposed to develop symptoms

and signs of parathion poisoning at lower exposures than those with high paraoxonase activity. It should be possible to measure ambient exposure, personal exposure, inactivation of paraoxon (paraoxonase), subclinical effects of paraoxon (erythrocyte and plasma acetylcholinesterase inhibition, combining activity and immunoquantitation), clinical effects on neuromuscular functions, and rate of recovery. The only missing step at present is an assay with peripheral cells, such as lymphocytes, for the P450-mediated conversion of parathion to paraoxon. Advances in analysis of genes and proteins of the P450 system should be applicable to this problem and permit a sophisticated estimate of attributable risks. The gene for paraoxonase may be linked to the putative gene for the disease cystic fibrosis, which would give great impetus to further biochemical and genetic studies of paraoxonase (White, New and DuPont 1984).

Potential Genetic Predispositions due to Differences in Tissue Sensitivity

1. Hemolysis due to glucose-6-phosphate dehydrogenase (G6PD) deficiency

This X-linked genetic abnormality predisposes affected males to hemolytic crisis from a long list of oxidizing drugs, naphthalene, and fava beans (Motulsky 1976, Beutler 1981). There are striking differences in gene frequency across populations, with distinctly different enzyme variants in persons of Mediterranean, black African, and Asian ancestry. In each population the deficiency gene became frequent because it conferred an advantage in surviving childhood malaria. There is good reason to suspect that such individuals may be predisposed not only to oxidizing chemicals administered as drugs, but also to oxidizing chemicals experienced as occupational chemicals or air pollutants. No systematic studies have yet been done, so carefully planned research must precede any suggestions for screening on a routine basis (Omenn 1984a; Omenn 1982; Office of Technology Assessment 1983).

2. Plasma alpha-1 antitrypsin deficiency

Alpha-1 antitrypsin is a circulating protein making up 90 percent of the alpha-1 globulin fraction of normal plasma. Inherited nearly complete deficiency in this protein is associated with a markedly enhanced risk of pulmonary emphysema, due to destruction of alveolar air spaces. Proteolytic enzymes released from white blood cells and pulmonary macrophages do the damage, unless regulated and inactivated by the alpha-1 antitrypsin. About five percent of the general population are heterozygous carriers for this deficiency and may be at enhanced risk of emphysema if they are smokers. Other irritative pulmonary exposures might also be important triggers of damage.

CONCLUSION

Differences among people in susceptibility to environmental exposures are real and are due to many factors, of which genetic factors are likely to be quite important. Some of the predisposing conditions may be inherited or at least predate occupational or other exposures; others may be caused by, or be exacerbated by, such exposures. The Office of Technology Assessment report The Role of Genetic Testing in the Prevention of Occupational Disease (1983) properly separated surveillance or monitoring for effects on the genetic material (toxicology) from screening for predisposing inherited traits (ecogenetics). In principle, monitoring would be done repeatedly on occupationally-defined populations or other specially-exposed populations, whereas screening would be done only once on individual candidates for employment or current employees. In fact, there is potential overlap between the two categories. Evidence of genetic toxicity might be found upon screening. Predispositions may contribute to variation in monitoring results.

The best way to deal with such definitional problems is to focus on specific mechanisms of toxicity and specific exposures as we seek to characterize and reduce the risks to human health from environmental exposures of many kinds. In any case, recognition of individuals or groups of persons at increased risk for adverse effects from chemical exposures should lead us to redouble our efforts to reduce or eliminate those exposures.

REFERENCES

- Beutler, E. (1978). Hemolytic Anemia in Disorders of Red Cell Metabolism. Plenum Medical, New York.
- Bloom, A.D. (ed) (1981). Guidelines for Studies of Human Populations Exposed to Mutagenic and Reproductive Hazards. March of Dimes Birth Defects Foundation, New York.
- Bridges, B.A., Butterworth, B.E. and Weinstein, I.B. (eds) (1982). Banbury Report 13: Indicators of Genotoxic Exposure. Cold Spring Harbor Laboratory, New York.
- Calkins, D.R., Dixon, R.L., Gerber, C.R., Zarin, D. and Omenn, G.S. (1980). Identification, characterization, and control of potential human carcinogens: A framework for Federal decisionmaking. J. Natl. Cancer. Inst. 64:169-75.
- Cartwright, R.A., Glashan, R.W., Rogers, J.H., Ahmad, R.A., Barham-Hall, D., Higgins, E. and Kahn, M.A.: Role of N-acetyltransferase phenotypes in bladder carcinogenesis: A pharmacogenetic epidemiological approach to bladder cancer. Lancet 2:842-846, 1982
- Dengler, J.H., Eichelbaum, M. and Meyer, U.A. (1984). Report on the Workshop "Molecular basis of polymorphic drug oxidation in man". Eur. J. Clin. Pharmacol. 27:253-57.
- Distlerath, L.M. and Guengerich, F.P. (1984). Characterization of a human liver cytochrome P-450 involved in the oxidation of debrisoquine and other drugs by using antibodies raised to the analogous rat enzyme. Proc. Natl. Acad. Sci. USA 81:7348-52.

- Eiberg, H., Mohr, J., Schmeigelow, K., Nielsen, L.S., and Williamson, R. (1985). Linkage relationships of paraoxonase (PON) with other markers: indication of PON-cystic fibrosis syndrome. Clin. Gen. 28:265-71.
- Friedman, R.D. (1979). Sensitive populations and environmental standards: A legal analysis. The Conservation Foundation, Washington DC.
- Idle, J.R., Mahgoub, A., Sloan, T.P., et al. (1981). Some observations on the oxidation phenotype status of Nigerian patients presenting with cancer. Cancer Lett 11:331-38.
- Idle, J.R. and Smith, R.L. (1979). Polymorphisms of oxidation at carbon centers of drugs and their clinical significance. Drug Metab. Rev. 9:301-07.
- Idle, J.R. and Smith, R.L. (1979). Polymorphisms of oxidation at carbon centers of drugs and their clinical significance. Drug Metab. Rev. 9:301-07.
- Kellerman, G., Shaw, C.R. and Luyten-Kellerman, M. (1973). Aryl hydrocarbon hydroxylase inducibility and bronchogenic carcinoma. New Engl. J. Med. 289:934-37.
- Koenig, J.G., Pierson, W.E., Horike, M. and Frank, R. (1983). A comparison of the pulmonary effects of 0.5 ppm versus 1.0 ppm sulfur dioxide plus sodium chloride droplets in asthmatic adolescents. J. Toxicol. Environ. Health 11:129-39.
- Koenig, J.G. and Omenn, G.S. (In press) Genetic factors in non-neoplastic lung disease. In, Variations in Susceptibility to Toxic Agents in the Air: Identification, Mechanisms, and Policy Implications, (eds.) J.D. Brain, B.D. Beck and Warren), Baltimore Johns Hopkins University Press, Chapter 3.

- Larrey, D., Distlerath, L.M., Dannan, G.A., Wilkinson, G.R. and Guengerich, F. (1984). Purification and characterization of the rat liver microsomal cytochrome P-450 involved in the 4-hydroxylation of debrisoquine, a prototype for genetic variation in oxidative drug metabolism. *Biochem.* 23:2787-95.
- Lower, G.M. Jr, Hilsson, T. and Nelson, C.E. (1979). N-acetyltransferase phenotype and risk in urinary bladder cancer: Approaches in molecular epidemiology. Preliminary results in Sweden and Denmark. *Environ. Health Perspect.* 29:71-79.
- Matsushima, M., Yagishita, T. and Ando, K. (1981). The diagnostic value of urinary cytology in dyestuff workers exposed to aromatic amines. Abstract, Third International Conference on Environmental Mutagens, p 49.
- Motulsky, A.G. (1976). Eco-genetics: Genetic variation in susceptibility to environmental agents. In, Human Genetics, Proc Fifth Intl Cong Human Genetics, Excerpta Medica, Amsterdam, pp 375-385.
- Mueller, R.F., Hornung, S., Furlong, C.E., Anderson, J., Giblett, E.R. and Motulsky, A.G. (1983). Plasma paraoxonase polymorphism: A new enzyme assay, population, family, biochemical, and linkage studies. *Amer. J. Hum. Genet.* 35:393- .
- National Research Council (1983). Risk Assessment in the Federal Government: Managing the Process. National Academy Press, Washington DC.
- Nebert, D.W. (1978). Genetic control of carcinogen metabolism leading to individual differences in cancer risk. *Biochimie* 60:1019-29.
- Office of Technology Assessment (US Congress) (1983). The Role of Genetic Testing in the Prevention of Occupational Disease. US Congress, Washington DC.

- Omenn, G.S. (1984a). Advances in genetics and immunology: the importance of basic research to prevention of occupational diseases. Arch. Environ. Health 39:173-82.
- Omenn, G.S. (1984b). Characterizing risks: utilizing knowledge about mechanisms. In Molecular and Cellular Approaches to Understanding Mechanisms of Toxicity, (ed.) A.H. Tashjian Jr, Harvard School of Public Health, Boston, pp 224-45
- Omenn, G.S. (1983). Environmental risk assessment: relation to mutagenesis, teratogenesis, and reproductive effects. J. Amer. Coll. Toxicol. 2:113-25.
- Omenn, G.S. (1982). Predictive identification of hypersusceptible individuals. J. Occ. Med. 24:369-74.
- Omenn, G.S. (1979). Genetics and epidemiology, medical interventions, and public policy. Soc. Biol. 26:117-25.
- Omenn, G.S. and Gelboin, H. (eds) (1984). Banbury Report 16: Genetic Variability in Responses to Chemical Exposure. Cold Spring Harbor Laboratory.
- Omenn, G.S. and Motulsky, A.G. (1978). Eco-genetics: Genetic variation in susceptibility to environmental agents. In, Genetic Issues in Public Health and Medicine, (eds. B.H. Cohen, A.M. Lilienfeld and P.C. Huang), Springfield Illinois, pp 83-111.
- Strong, L. S. (1985) Mutational models of cancer etiology. In, Genetics in Clinical Oncology, (eds. Chaganti, R.S.K. and German, J.L.), Oxford University Press.

- White, P.C., New, M.I. and DuPont, B. (1984). Cloning and expression of cDNA encoding a bovine adrenal cytochrome P-450 specific for steroid 21-hydroxylation. Proc. Natl. Acad.Sci. 81/7:1986-90.
- Woolhouse, N.M., Eichelbaum, M., Oates, N.S., Idle, J.R. and Smith, R.L. (1985) Dissociation of co-regulatory control of debrisoquine/phenformin and sparteine oxidation in Ghanaians. Clin. Pharmacol. Therap. 37:512-21.

TABLE 2: GENETIC VARIATION IN HUMAN SUSCEPTIBILITY TO ENVIRONMENTAL AGENTS

Drugs

Metabolism:	N-acetyl-transferase
	Plasma pseudocholinesterase
	C-oxidation
Sensitivity:	G6PD deficiency
	Methemoglobin reductase deficiency

Inhaled pollutants/pesticides

- 1-antitrypsin deficiency
- Aryl hydrocarbon hydroxylase inducibility
- Metabolic conversion of nicotine
- Plasma paraoxonase activity

Foods

Lactose intolerance	Fava beans
Wheat gluten/celiac disease	Goitrogens
Saturated fats/atherosclerosis	Catecholamines

Food additives and stimulants

Iron, ethanol, caffeine

Physical agents

Cold weather, motion sickness, metal poisoning, color vision, ultraviolet radiation, X-radiation

Infectious agents/autoimmune disorders

- Malaria: Duffy blood group; sickle hemoglobin, thalassemia, G6PD deficiency
- Predispositions due to impaired antibody production or cellular immunity
- Predispositions associated with histocompatibility phenotypes

Source: Omenn and Motulsky (1978)

TABLE 1: Framework for decision-making in the United States for potentially hazardous chemicals

IDENTIFICATION	Epidemiology Animal tests In vitro screening
CHARACTERIZATION	Relative potency Exposures Range of susceptibility
CONTROL	Information Regulation Substitution

Source: National Research Council (1983)

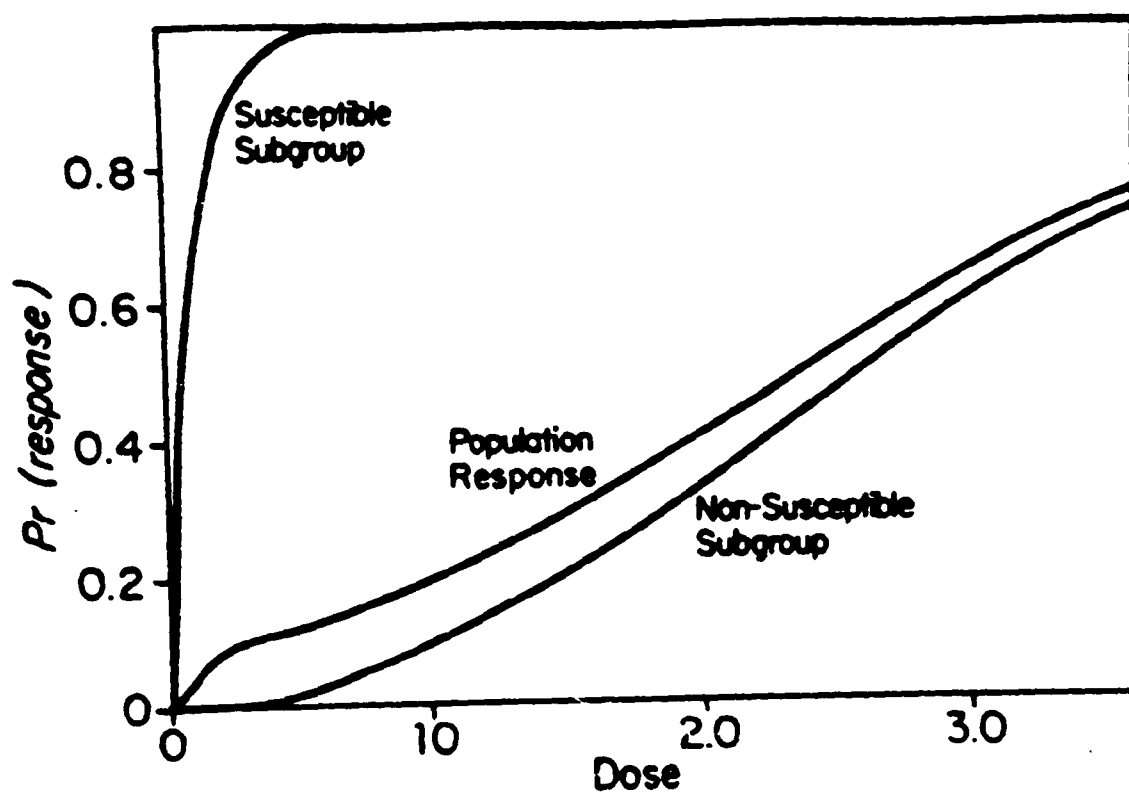


Fig. 1. Scheme to illustrate subgroup heterogeneity in responses to exposure to a given chemical.

Part III

Refining Epidemiologic Approaches
to Assessment of Exposure

Part III - Section A

Measures in the Total Environment

Chapter 10
Issues in the Measurement of Exposure
Genevieve M. Matanoski

ISSUES IN THE MEASUREMENT OF EXPOSURE

Genevieve M. Matanoski

Introduction

Risk assessment involves the prediction of the level of health effects which would result from a measured exposure to a toxic agent. Epidemiologic data has been used for this purpose in radiation risk assessment. For many agents, controversies have developed over interpretation of epidemiologic data relating these substances to human disease and these questions have influenced regulators to turn to animal data to assess risk in humans. The answer to this dilemma lies not in the abandonment of human data, but rather in a more thorough examination of the available human information to identify the flaws in the existing designs and methodologies and then to implement correction so that eventually risk estimates can be based on data from the true reference population, the human.

Role of epidemiologic studies in risk assessment

Epidemiologic studies attempt to associate the risk of disease with exposure to an agent. Therefore, it would seem reasonable that such studies could be used for risk assessment but often they are not. The discrepancy stems from the fact that the needs of epidemiologists are met using study designs, exposure measurements and analysis which will not address the issues raised by risk assessors. This chapter will examine why epidemiologic studies have not satisfied these issues. It will then address the problems with measurements of exposure and finally, the problems of assessing the detailed interrelationships of exposure and outcome will be examined. The latter may be the most important challenge facing epidemiologists today.

Focus of epidemiologic studies

The purpose of etiologic studies in epidemiology is to determine whether there is an association between an exposure and the risk of disease. In most cases, the epidemiologist attempts to demonstrate an all-or-none relationship; that is, exposure to cigarette smoking increases the risk of lung cancer in smokers compared to non-smokers. If dose data are used in epidemiologic studies, the categories are mutually exclusive and based on a life time exposure level; they include only a few large categories encompassing wide dose variations; and they have open-ended categories at the extremes of exposures. Although epidemiologists may frequently attempt to demonstrate an association of increasing risk with increasing levels of exposure, they use these data primarily to support a probable causal relationship between the agent and the disease. They are not concerned with a so-called "safe dose." In the past, epidemiologists have proposed that the logical preventive step, assuming an association does exist between an agent and a disease, is to completely remove the agent to avoid this potential cause of disease. The epidemiologist would recommend that cigarettes be withdrawn from human consumption.

The needs of the risk assessor can best be met if the epidemiologic study is viewed differently. The results of the study must determine the dose response relationship between the disease outcome and the exposure variable. The risk assessor requires this information because many of the agents of concern cannot be banned from human consumption or exposure. For example, it may be necessary to compare risks from exposures to alternative energy resources in order to assess which alternative resource carries the least risk and what level of containment of any resulting pollution may be necessary in order to assure the minimum hazard to the population. This type of risk

assessment is an area where epidemiologic involvement is still new. Most epidemiologists have not been concerned with permissible levels which may be delivered to the general public or to sensitive groups. They have not considered risks in which the referent population is worldwide and yet this may be the arena in which risk assessors are predicting the outcomes of various energy alternatives. Epidemiologists typically consider the risk only in relation to the local population from which the cases originate. If epidemiologists are to address the problems faced by risk assessors, they must design studies, measure exposures and analyze results with a considered view of this specific use. This will require the development of new perspectives on the measurement of exposures such as biomarkers and better methods for estimating exposures. Epidemiologists must address more precisely the exact agent to which the population has been exposed. They must then use better methods of relating dose and outcome than they have in the past.

Testing for the agent

The ideal situation for assessing risk from exposure to an agent would be one in which the population under study had only been exposed to a single agent and a comparison population identical in every respect except for the exposure was also studied; each of these populations would have complete follow-up over a life-time; and measurements of the agent would be available from the time it was first introduced using the same reliable and valid technique. In any study assessing risk, it is necessary to examine how far the collected data and the study design deviate from the ideal.

The first question which must be asked is what agent may be causing a risk? This is a difficult question to answer for most exposures in the general environment. There is never a single agent in the ambient air, at the dump site or in the workplace. There are multiple agents in these environments,

Matanoski

-3-

many of which are complex mixtures which have varied in the past from one time period to the next. Drugs are probably the purest exposure which individuals may experience but even these change by dose level and by composition over time so that, for example, oral contraceptives today do not represent the same exposure they did in the past. The problem is magnified for other exposure measurements. In the general environment it is important to remember the possible interactions between agents within the environment so that what is measured at the source may not represent the material to which the individual is exposed.

The characteristics and level of exposure may change with calendar time. Let us consider the case of particulates measured in air. The level of particulates may have changed over time but these levels have been measured using similar testing procedures for many years. Thus, the method of measuring particulates using the procedures of "total suspended particulates" (TSP) in the United States or British smokes (BS) in the United Kingdom have involved consistent test methods for many years. However, even when testing procedures appear to measure the same agent, as in the case of particulates, the size of the particle, the materials adhering to their surfaces and other factors may all cause differences in the relevant exposure dose to the human and the specific agents which may be delivered to the susceptible tissues. If one considers the example of particulates, the levels in the environment have decreased over time but it is not clear whether the proportional size of the particles has changed. The particulates measured today may not actually represent the same proportion of respirable particles as did the same measure twenty or thirty years ago. Thus dose response curves based on these values may result in errors since they are not relevant to a true measure of tissue exposure to particulates, that is those particles which are of the size which can reach the lung fields.

Matanoski

-4-

The true composition of materials in an exposure situation may change not only by time but also by place. This is especially true of the complex mixtures which are involved in air pollution in the general environment or at the work site. The work environment is repeatedly referred to since it is an important source of information on the risks from exposure to toxic agents. Compared to measurements of population exposures in the general environment, workers may have higher exposures to a pollutant under more controlled conditions in which measurements of the ambient air may have been taken for several years. This makes them an important group when assessing risks. It is also important to recognize the limitations of information from working populations. If one wishes to measure population effects from exposures to decomposition of organics, information gathered from worker exposures in certain smelter settings may not represent similar exposures to organics at each site. Each site may represent an exposure to a completely different mixture of polynucleated aromatic compounds (PACs). None of these exposures may be similar to those experienced by the general population. Organic mixtures of all substances may represent exposures that have the maximum variation according to local site.

In addition, it is important to consider the consistency of testing procedures over time. Radiation is probably the simplest example of a single agent for which external exposures have been quantified in populations over time and yet there is concern about the equivalency of testing procedures from one time period to another. Questions arise concerning the standardization of the dose measurement and recommended procedures for wearing badges or dosimeters. The error rate attached to a measurement becomes very important. If that error rate differs according to the dose and if that error rate has a consistent bias in measurements by dose so that lower doses are overestimated

or higher doses underestimated, the shape as well as the slope of the dose response curve can be altered.

Despite the fact that these issues are of crucial importance in evaluating the effect of an agent on the risk of disease, a close look at epidemiologic data often indicates that, despite differences in exposure and variations in measurements over time and other potentially important considerations, the results of studies in different populations may still yield consistent information regarding dose effects. If this is the case, the consistency of the information would lend support to the probability that the dose response assessment is correct. The data on effects from radiation fit this situation quite closely. The risk of leukemia in relation to exposure to radiation dose is quite consistent from study to study. Therefore, it is important that issues relating to imprecision in test measures not be introduced as the prime reason for discarding epidemiologic data in preference to the controlled animal experiment.

In summary, several issues which arise when one attempts to use epidemiologic data for risk assessment focus on measurement of the agent. There is a need for more attention to specification of the exact agent which may be related to the outcome. Complex mixtures need to be characterized more precisely especially in situations with exposure to organics. Environmental interactions may change the actual agent to which the human is exposed regardless of the agent measured at the source. The characteristics of the agent may change by time and place so that even though measured values may appear the same, the biologically effective dose may have changed. Finally, variations in test procedures may not only alter the capacity to measure certain agents but may also influence the levels of agent which are detectable.

Design and analysis

Epidemiologists must share some of the blame for the non-use of

epidemiologic data in risk assessment. Risk assessors have found that many epidemiologic studies are missing important information, such as the measure of dose in relation to the risk of disease. As epidemiologists have become aware of this deficit, a burgeoning literature is arising which describes new methods for more adequately assessing dose response relationships than in the past.

Early studies in the epidemiologic literature focused on the case-control design. Such studies are unable to evaluate dose or the related measure of latency both of which are needed for dose risk assessment. The only other design which was frequently used in the past, for example, for measuring air pollution risks, was the cross-sectional prevalence study which is similar to the case-control study. In this type of study, although one can compare the rates of disease among areas which have different levels of ambient exposures, dose response curves and latency of disease cannot be determined. Assessment of general environmental risks using dose-response curves derived from occupational studies where exposures are higher and past measurements or estimates of dose are more readily calculable has replaced to some extent the earlier types of studies. In these situations it is usually necessary to decide on the appropriate model to apply to the dose-response curve in order to extrapolate to the low exposure levels of the general population. Specific drug exposures and the accidental exposure of populations to pollutants are types of studies which also have been used in epidemiology to assess dose response relationships. In the case of drug use however, the populations exposed may be unusual because of the disease which prompted their treatment.

In order to evaluate the association between a dose and an outcome in an epidemiologic study, one is faced with a problem of interrelated time variables. Time is frequently the measurement by which one estimates the

exposure of the individual. In occupational settings the duration and the calendar time period of the period of work are part of the estimation of dose. Duration of residence in a specific area or house may define a "dose" from general environmental exposures. Estimation of dose by these methods causes the dose measure to overlap with other time variables such as the latent period of disease and the aging of the population under study. This creates many problems in the analysis of epidemiologic studies.

Epidemiologic studies such as those in occupational settings have often focused on the high dose end of a dose response curve. The shape of the dose response curve at low doses is estimated from these high dose data. This is a difficult task. In most cases, the shape of the curve at lower levels is unknown. There are rarely any populations large enough to estimate risks at low doses so that one can develop points on the lower end of the exposure curve, but this problem is not unique to human studies. It is inherent in animal experiments as well since there, too, one is constrained by population size. The best measure one may be able to obtain may be an estimate of the upper limit of the risk at the lower end of the dose curve. This situation will always be a limiting factor in animal and human data.

The problems of relating precise dose to outcome in longitudinal studies are being examined thoughtfully by both epidemiologists and statisticians.

Dose

In determining the dose of an agent, it is important to consider the action or characteristic response of an agent in the host. The usual method used for calculating the dose of an agent is the cumulative dose, that is, an addition of the dose for each time period over the total time exposed. This is the classical dose measurement for external radiation from X-rays or gamma rays.

This method of measuring exposure works well for agents which enter the host, may cause an effect on some cells and then disappear. Some chemicals as well as radiation may be best measured by this model.

Other agents may need different methods of measuring dose. If one is studying infectious exposures from a workplace environment, it may not be necessary to calculate a cumulative dose of the etiologic agent but only a measure of the probability of exposure. Once an individual has been infected, if the host environment is permissive, the agent will multiply to produce its own dose which is effective in producing disease. Repeated exposures to an infectious source in the environment may increase the probability of disease to some point and thereafter, the probability of ever causing disease in the remaining individuals may actually be reduced compared to some general population because they are not susceptible through previous subclinical infection and immunity or because of some other form of resistance.

The cumulative dose measure should be considered an additive measure of instantaneous exposures. It does not seem to be an appropriate measure for agents which are retained in the body for long periods after initial exposure. Investigators have discussed this problem in relation to silica and asbestos since they do reside for long periods in body tissues but it may also be a consideration in regard to lead and other metals which also create a long-term body burden. If these materials are retained in the tissues then they may continue to insult the body for longer periods than the moment when the exposure took place. In these cases it may be appropriate to analyze the data using an integration of dose in which the dose or fraction of dose retained from a previous year will be added to the level from the present year. This is a more accurate representation of tissue dose than it is a cumulative dose.

Such dose measures have been proposed for silica (Jahr 1974) and for asbestos (Acheson and Gardner 1981). To calculate these doses it is important to know the half life of retention of the agent in the body and the biological effectiveness of the material during that period. Thus, the deposition in body compartments, metabolism, storage and excretion are some of the important measures needed.

Again considering radiation, it is known that the important variable is not just the external exposure but the tissue dose and the biological effectiveness of that dose. Consideration of these variables has increased the precision of dose effect measurements for radiation. Such precision should be introduced for other substances. For asbestos one needs to examine the dose on the basis of how much is inhaled; how long is it retained in the lungs; where does it reside in the lungs, in the central portion or in the periphery, and for how long; into what other tissues does the asbestos migrate and how long is it retained at these other sites? These are important questions which should be considered in determining the true dose of asbestos which is associated with risks of cancer at various tissue sites. The epidemiologist is challenged to pull together such information from multiple sources in order to estimate risk. Epidemiologic studies of the past have not focused on such specific issues as tissue dose and the appropriate method of assessing the true biologically effective dose.

This level of precision may not be needed for all epidemiologic studies. What is needed is an awareness that when assessing studies, one must evaluate whether characteristics of the exposure as well as characteristics of the population may influence the risk of disease due to changes in ambient exposure factors as well as retention of the agent in the host, distribution in various host tissue sites and variations in biological effectiveness of the agent.

Such factors might explain the observed differences in results from one study to another. A study of the risks of cancer of the colon from inhalation of asbestos in an occupational setting would not be the same as ingestion of asbestos in drinking water. The amount and form of asbestos which gets to the colon may be different. Studies of asbestos in different occupational settings may not be equivalent depending on the type of asbestos, the fiber length and even the work and personal characteristics of the population. These factors should be considered in determining the true dose response curve for an agent and in evaluating variation in results from different studies.

Chemical agents present a different problem of assessment of risk in relation to dose. Physical agents have similar and easily definable characteristics regardless of the study. Their effects are likely to be the same across species. Important mechanistic factors to be considered in relation to body dose such as particle size may be known for physical agents. Chemical agents, however, are more of a challenge. The agent itself may be altered by the body. Some of the chemical, as in the case of heavy metals, may be retained in the tissue for varying time periods. Their actions may be enhanced or inactivated in the tissue in which they reside. Thus, although cumulative environmental dose has been useful in formulating dose response curves in the past, this measure should be scrutinized more carefully in the future to determine how it may be modified to better represent the true tissue dose. In some cases, methods of integrating total dose or other methods of modelling tissue dose may be more appropriate.

The dose rate may be another factor to consider in assessing effects although methods of describing this measure in relation to health effects are limited in most epidemiologic studies. A heavy dose at one point in time may cause more serious health effects than multiple small doses which allow periods

of recovery. Of course, the opposite situation could also hold. A single large dose at one point in time may interfere with the clearance system of the lungs or other tissues which will allow more damage or retention of the agent than would be possible otherwise. Dose rates, timing of the doses and age of the subjects may be important factors in relation to disease manifestations.

Latency

The timing of the dose is intimately associated with the measurement of the latent period of the disease. In addition, it is impossible to calculate the true dose unless the latency of a chronic disease is known. If an individual is exposed to an agent until the onset of disease it is obvious that there is a period during which dose of the agent may continue to accumulate despite the fact that there have been irreversible changes which eventually will lead to overt disease. Any dose accumulated during the latent period which was not necessary for the development of disease must be subtracted from the total dose. The period used for correcting or lagging the calculated dose level is usually an empirical measure and often represents the minimal latent period which has been reported for that disease. More consideration must be given in the future to appropriate methods such as modelling which will furnish the best estimate of this lag period.

In the past, most studies assessing dose response relationships have not considered the problem of latency. As long as the focus of epidemiologic studies was on case-control study designs it was difficult if not impossible to estimate the latent period of disease although there have been attempts to do this. Case-control studies are essentially cross-sectional in nature. In this situation if the specific agent of interest causes a difference in the latent period compared to that associated with the usual disease in the

population—that is, if the latent period changes with increasing dose—the distribution of latency in cases may not reflect the true latent period for the disease associated with the exposure. This was pointed out in a review of methods of calculating latent period in bladder cancer (Matanoski and Elliott 1981). Only if the exposures and latent periods of disease remain constant in a population over decades will the case-control study represent a true picture of disease. Since most investigators believe that environmental factors play important roles in the etiology of chronic disease, it is impossible to conceive of our environmental, personal, social or general characteristics being that static. If the environment changes, the characteristics of diseases it causes, would also change.

To illustrate the problems in calculating latency in the population, let us assume that an agent was used in an industry for a only few years. The agent appears to shorten the latency of a disease compared to that which occurs in the natural history of that disease. Let us also assume that the higher the dose of the agent, the shorter the latent period. Most of the workers will have been exposed for the first time at a young age. If one conducts a case-control study in the total population a short time after the earliest onset of exposure to the agent and within the shortened latent period of the industrially caused cases, the correct risk from exposure in the industry would be apparent as long as appropriate young age groups are included. If use of the agent is discontinued and the population is studied later, the attributable as well as the relative risks will be perceived as being low and the dose causing that risk will be erroneously underestimated. In the later period, the number of residual cancer cases will represent only those few cases with a low dose and a long latency because of this low dose (Figure 1). If one examines the general population or even the industry group for dose response

relationships and one does not have the total group of exposed individuals for analysis, the dose response relationship can be misleading if dose is related to latency (Figure 2). If the population is followed from different points after the initial exposure as indicated on the graph, the risk of disease will be based only on the survivors at the first point of contact forward. The apparent latent period of disease could be different depending on when that cohort starts. To be specific (Figure 2) if one conducts a study beginning at point A on the curve and collects cases for several years thereafter, the study group will include cases with high, medium and low exposures and risks by dose will be more representative of the true risk from the past industry exposure than it would if the study was started at point C. At the later point, only the few low exposure cases will remain and all cases will report a very long latency from the time of first exposure. One can also see from Figure 2 that if one studies the population between points A and B versus B and C; that is, the study of the population is terminated for both the start of the true cohort and the period of follow-up. Most of the cases in the first interval will come from the true high and medium level exposed group, whereas in the second interval they will come from the true medium and low level exposed group. If the observed cases are divided into a high medium and low group, they will be artificial divisions which may give erroneous estimates of the true latent period for the disease. This will not only change the true risk of disease in the population but it should also distort the apparent shape of the dose response curve. These principles become important in interpreting what is seen in populations.

Issues in outcome

Most of the previous discussion has focused on outcomes of disease which are permanent. This has been the most common type of investigation in the past. The measurement of cancer or mortality in occupational settings or among survivors of an accidental exposure is a typical example of this assessment of risk. This is not the only type of disease outcome currently being assessed in dose response analyses. Continuous changing and repetitive outcomes provide even more of a challenge in determining the dose associated with a risk.

Attempts have been made to assess the level of mortality rates at each point over a long period of time in relation to the level of air pollution at that time (Gore and Shaddick 1958; Martin and Bradley 1960; Schimmel and Greenberg 1971; Holland et al. 1979). While it is true that many attempts have been made in the past to correlate these two measurements of outcome and exposure on the basis of daily changes as if there were an instantaneous cause and effect, many confounding variables were not considered. Recent attempts to relate these values rely on stronger statistical methods than those of the past but questions of interpretation remain (Mazumdar, Schimmel and Higgins 1982; Ware, Thibodeau, Speizer et al. 1981). The problems which have arisen in these analyses are related to evaluation of events which change by seasonal cycles, by calendar time trends, as well as on a daily basis. These outcome measures must be related to pollution levels which also have patterns of change reflecting season, annual trends and daily levels. After correcting for patterns and autocorrelations among events the statistical analysis must show that any relationship between the agent and the mortality rate are not simply random events.

Studies have attempted to measure longitudinal changes in populations from a relapsing chronic disease or from physiologic alterations in function. These are even more difficult to associate with the dose of an agent. Attempts may

Matanoski

be made to measure episodic periods of increased dysfunction during the course of progressive disease such as chronic pulmonary disease. There have been efforts to relate measures of reduction in the frequency of chronic disease episodes to reduction in levels of the agent in the environment in order to arrive at a safe exposure value (Iawther, Waller and Henderson 1970; Whitmore and Korn 1980). Some studies try to test whether the rate of normal growth of a physiological function or the rate of change of such a function is altered by variation in the level of a pollutant (Ferris, Chen, Puleo et al. 1976, Avol, Linn, Shamoo et al. 1983; Ware, Dockery, Spiro et al. 1984; Remijn, Fischer, Brunekreef et al. 1985; Rona, Chinn and du Ve Florey 1985). The methods of assessing the exposure vary. A single agent may be selected as the marker of pollution in the analysis or several agents may be tested against the changing level of disease and whichever one shows a significant relationship is selected for study. The latter type of analysis has concerned statisticians because multiple examinations of the same data set may uncover significant associations by chance.

In the air pollution field lung function is most often examined. In children this would involve measurements to test deviation from a so-called normal rate of growth of lungs; in adults, the evaluation might refer to measurements to test deviation from the usual rate of deterioration of lung function. These studies have created new challenges for statistical analysis of longitudinal data. The biological interpretation of such data also create new questions. There is the problem of autocorrelation of both the outcome as well as the pollutant measures which may cause both factors to change in the same pattern over time. Events on subsequent days are related to the events of the previous day. This is true of asthmatic attacks, bronchitis and infections and obviously lung function. The characteristic growth of lungs in children

may be related to the measurements of function at the start of the study. Intermittency of repeated testing may cause a problem especially since this too may be a patterned and not random event. Participation may vary from one time period to another and the methods of handling these problems are difficult when one is evaluating a continuous biologic measure. Certain individuals may always be tested on specific days of the week at a time that relates to the same pattern of air pollution. Further consideration should be given to the fact that the dose of agent on one day is usually related to the individual's lung function several days later instead. This would be the case if there is a delay in the manifestation of lung change due to the exposure. These are very difficult problems to solve.

The previous discussion has emphasized the problems which have limited the use of epidemiologic data for risk assessment. The iteration of this long list should not be taken to imply such data should not be used for risk assessment. Indeed, the conclusion that should be reached from this discussion is the exact opposite. There are human data which can be used for risk assessment. There are appropriate methods of analyzing these data which are developed or are being developed (Breslow, Lubin, Marek et al. 1983; Cook and Ware 1983; Korn and Whitmore 1979). Methods for measuring exposures for epidemiologic studies still need further research and development for many substances but the work is progressing. Clearly all of these areas will require more serious thought by the scientific community. None of this discussion is meant to imply that the animal data are not important as well. The laboratory scientist may provide important leads relating to biological mechanisms of action as well as dose response of agents. Tying together appropriate information from all sources will strengthen our knowledge and lead to correct interpretations of biologic effects of environmental substances.

Matanoski

-17-

Summary

Despite the problems which risk assessors have perceived related to the use of dose response data from epidemiologic studies, it seems reasonable that these data may be more relevant for assessment of human risk than are animal data. Several issues arise because the epidemiologists and risk assessors have had different reasons for using the information. As improved analytic techniques are applied to the data and as epidemiologists become aware of the needs of risk assessors, the usefulness of the dose response data on human populations can be enhanced. Communication and understanding of the areas of interest for both groups is essential. There needs to be a strong effort on the part of governmental agencies to foster the growth of interdisciplinary research in quantitative human risk assessment and to further the ties between laboratory scientists and epidemiologists so that data from both fields can be used more effectively in assessing risks.

REFERENCES

- Acheson, E.D., Gardner, M.J. (1981). Exposure limits - the scientific criteria
In, Recent Advances in Occupational Health (ed: J.C. McDonald), Churchill
Livingston, NY. Chapter 23.
- Avol, E.L., Linn, W.S., Shamoo, D.A., Venet, T.G., Hackiney, J.D. (1983)
Acute respiratory effects of Los Angeles smog in continuously exercising
adults. J. Air Poll. Control Assoc. 33:1055-60.
- Breslow, N.E., Lubin, J.H., Marek, P. et al. (1983) Multiplicative models
and cohort analysis. J. Am. Stat. Assoc. 78:1-12.
- Cook, N.R., Ware, J.H. (1983) Design and analysis methods for longitudinal
research. Ann. Rev. Public Health 4:1-23.
- Ferris, B.G., Jr., Chen, H., Puleo, S., Murphy, R.F.H., Jr. (1976) Chronic
non-specific respiratory disease in Berlin, N.H., 1967 to 1973. Am. Rev.
Resp. Dis. 113:475-85.
- Gore, A.T., Shaddick, C.W. (1958) Atmospheric pollution and mortality in the
country of London. Br. J. Prev. Soc. Med. 12:104-13.
- Holland, W.W. et al. (1979) Health effects of particulate pollution:
re-appraising the evidence. Am. J. Epid. 110:525-659.
- Jahr, J. (1974). Dose-response bases for setting a quartz threshold limit
value. A new simple formula for calculating the "Lifetime Dose" of
quartz. Arch. Environ. Health 29:338-40.
- Korn, E.L., Whittemore, A.S. (1979) Methods for analyzing panel studies of
acute health effects of air pollution. Biometrics 35:795-802.
- Lawther, P.J., Waller, R.E. and Henderson, M. (1970) Air pollution and
exacerbations of bronchitis. Thorax 25:525-39.

- Martin, A.E., Bradley, W.H. (1960) Mortality, fog and atmospheric pollution -an investigation during the winter of 1958-59.
- Matanoski, G.M. and Elliott, E.A. (1981). Bladder cancer epidemiology. Epidemiol Rev. 3,203-29.
- Mazumdar, S., Schimmel, H., Higgins, I.T.T. (1982): Relation of daily mortality to air pollution: An analysis of 14 London winters 1958/59-1971/72. Arch. Environ. Health 37:213-20.
- Remijn, B., Fischer, P., Brunekreef, B., et al. (1985) Indoor air pollution and its effect on pulmonary function of adult non-smoking women (Parts I-III). Inter. J. Epid. 14:215-230.
- Rona, R.J., Chinn, S., du Ve Florey, C. (1985) Exposure to cigarette smoking and childrens' growth. Inter. J. Epid. 14:402-409.
- Schimmel, H., Greenberg, L. (1972) A study of the relation of pollution to mortality, New York City, 1963-68. J. Air Pollution Control Assoc. 22:607-16.
- Ware, J.H., Dockery, D.W., Spiro, A. III, Speizer, F.E., Ferris, B.G., Jr. (1984) Am. Rev. Respir. Dis. 129:366-74.
- Ware, T.H., Thibodeau, L.A., Speizer, F.E., et al. (1981) Assessment of the health effects of atmospheric sulfur oxides and particulate matter: Evidence from observational studies: Environ. Health Persp. 41:255-76.
- Whittemore, A.S., Korn, E.L. (1980) Asthma and air pollution in the Los Angeles area. AJPH 70:687-696.

Chapter 11

Exposure Measurement for Air Pollution Epidemiology

Benjamin G. Ferris, Jr., James H. Ware and John D. Spengler

EXPOSURE MEASUREMENT FOR AIR POLLUTION EPIDEMIOLOGY

Benjamin G. Ferris, Jr., James H. Ware and John D. Spengler

Introduction

Conventional air pollution epidemiology originated in response to the severe air pollution problems of the 1950's and early 1960's. To assess the health effects of chronic exposure to high levels of fossil fuel air pollution, investigators compared the respiratory health of populations living in communities or geographic areas with sharply contrasting air quality. Since that time, however, regulatory action, economic changes, and new technology have combined to reduce the severity of air pollution in the most severely affected communities. At the same time, tall stacks and long range transport of air pollutants have affected air quality in small communities and rural areas, especially in the eastern United States. Progress in understanding the chemistry and particle size composition of air pollution, along with the capacity for more sophisticated measurement, have further undercut the validity of a simple dichotomy between polluted and clean cities. The contemporary setting of relatively low air pollution exposures and multidimensional measurement of air quality require fundamental changes in the design of epidemiologic studies. Chief among these changes is the growing emphasis on the measurement of air pollution exposure of individuals rather than communities and the associated emphasis on the measurement of total human exposure rather than exposure due solely to outdoor sources. This chapter reviews the developments leading to these historic changes and examines the

Supported by the National Institute of Environmental Health Science ES-0002 and ES-0118, Electric Power Research Institute Contract No. RP1001, and Environmental Protection Agency Contract No. CR 811650-01.

implications of these shifts for the design and analysis of epidemiologic studies.

This discussion is oriented toward the measurement of air pollution exposure in studies designed to provide information about the exposure-response relationship between air quality and human health. Human risk assessment also requires information about the distribution of air pollution exposure in the population at risk (National Academy of Science, 1982). Although the issues discussed here bear directly on the characterization of that distribution, that problem will be discussed only indirectly.

On the use of outdoor air quality measurements in air pollution epidemiology

As noted previously, the health effects of long-term exposure to air pollution have traditionally been studied by comparing the health status of populations living in communities with different air quality. The earliest studies collected few if any air pollutant measurements. When measurements were obtained they were often collected at sites chosen for regulatory purposes and no claims were made that they were representative of the community as a whole. Later studies sought to characterize the heterogeneity of air quality within study communities by gathering data at multiple stations. The Nashville study was one of the earliest investigations to use this approach (Zeidberg, Schueneman, Humphrey et al. 1961). In the Harvard Study of Air Pollution and Health, a network was established of ten or more indoor and outdoor monitoring stations at residences in each of the six study communities (Spengler, Ferris, Dockery et al. 1979). These sites provided measurements of sulfur dioxide (SO_2), nitrogen dioxide (NO_2), respirable suspended particulates (RSP), and the sulfate component at six day intervals. Outdoor measurements collected at

Ferris

-2-

these stations showed little within city variation in respirable sulfate and particle mass concentrations except in Steubenville, Ohio (Spengler, Dockery, Turner et al. 1981). Spatial variation of outdoor concentrations of NO_2 and SO_2 was seen in all cities except Topeka, Kansas (Spengler, Ferris, Dockery et al. 1979). More importantly, these studies identified several limitations in the strategy of estimating personal exposure only from data collected at outdoor stations.

First, some pollutants have patterns of spatial variation that can be characterized only by sophisticated outdoor monitoring programs. Automobile emissions, for example, vary temporally and spatially near major roadways, but can be relatively uniformly distributed away from these roadways. Photochemical and chemical transformations further complicate this picture. Second, individuals move about within their communities, so that site-specific measurements are only indirectly related to personal exposure. This is less of a problem for studies of children, who spend most days at home or at school. Third, and most fundamental, the growing literature on human activity patterns and personal air pollution exposure has shown that most people spend 70 percent or more of their time indoors, that personal exposure is strongly influenced by indoor air quality, and that indoor environments vary substantially among homes within the same community.

For pollutants with outdoor sources, such as SO_2 , sulfates and ozone, this variation is determined primarily by differences in air exchange rates and the presence or absence of sinks indoors. Data from our indoor-outdoor network show that SO_2 concentrations are generally lower indoors than outdoors, that the size of the difference depends upon home characteristics and varies between

as well as within communities, and that annual mean indoor concentrations of SO_2 vary substantially between homes.

For pollutants with indoor sources, variation in indoor concentrations is even more substantial, and indoor concentrations can exceed outdoor values particularly in homes using gas for cooking. NO_2 concentrations were more variable indoors than outdoors in the homes participating in our indoor-outdoor network (Spengler, Ferris, Dockery et al. 1979; Spengler and Soczek 1983). Figure 1 represents the composite 24 hour annual indoor NO_2 concentrations measured in over 100 gas-cooking homes in Portage, Wisconsin. The mean outdoor concentration was 15 $\mu\text{g}/\text{m}^3$ yet the mean indoor concentration for gas-cooking homes was approximately 65 $\mu\text{g}/\text{m}^3$ with 3 percent of the homes exceeding 100 $\mu\text{g}/\text{m}^3$, the National Ambient Air Quality Standard (NAAQS) (Sexton, Letz and Spengler 1983). Similarly, indoor RSP concentrations were highest in homes with cigarette smokers. Each cigarette smoker added an estimated 30 micrograms per cubic meter to the annual 24 hour average indoor concentration of RSP (Dockery and Spengler 1983).

Not surprisingly, outdoor measurements are of little value in predicting personal exposure to pollutants with indoor sources. In a study of personal exposure to NO_2 among 300 children and adults living in Portage, Wisconsin, outdoor NO_2 concentration predicted only one percent of the variation in weeklong integrated personal exposure measurements collected with lapel samplers (Table 1).

Reliance only on outdoor measurements weakens epidemiologic studies in another way. Since each station represents exposure for a group of exposed individuals, assessment of the association between air pollution exposure and

health status requires comparisons between these groups rather than between individuals. This can raise difficulties of interpretation in studies comparing several communities. In the Harvard Study, for example, bronchitis rates varied substantially between study communities. Rates tended to be elevated in communities with higher concentrations of total suspended particulates, but also varied for reasons unrelated to TSP concentration, as indicated by the width of the confidence bands around each estimate of city-specific bronchitis rates (Figure 2). Although this unexplained variation may be reduced by refinement of the air pollution measurements, some variability will remain, and can be described as the intrinsic city to city variation in respiratory health. This variability affects the generalizability of the results from a study of several cities, since the application of these results to the population of cities is more appropriately measured by the distance between the points and the regression line than by the precision of the points themselves as estimates of the city specific rates.

Alternatives to outdoor measurement

New approaches to exposure measurement in air pollution epidemiology are required in order to determine more precisely potential relationships to specific components of the air pollution mix. One new approach, the use of home characteristics as predictors of exposure, has proved to have significant limitations. Since the presence of a gas stove is an important predictor of indoor NO₂ concentrations, for example, one might hope that stove type would serve as a useful surrogate for indoor or personal measurements of NO₂ concentration. Because indoor concentrations of NO₂ are highly variable within homes of a specified stove type, stove type predicted only 41 percent

of the variation in personal exposure measurements collected in the Portage study (Table 1). Although the improvement relative to models using only outdoor concentration is substantial, the remaining unexplained variation in personal exposure seriously compromises the sensitivity of epidemiologic studies aimed at assessing the effects of NO₂ exposure.

This experience suggests the need for more comprehensive approaches to the measurement of air pollution exposure for epidemiologic research. These approaches should recognize that exposure varies among individuals living in a single community, and that, while outdoor concentrations are one determinant of exposure, indoor environments and personal activity patterns play a fundamental role. Specifically, epidemiologic studies should be designed to provide individualized estimates of total human exposure to the pollutants under study. From this perspective, each individual, i , has an exposure to pollutant, j , denoted E_{ij} . Epidemiologic studies should be designed to estimate the dose-response relationship between exposure, E_{ij} , and health status. Although studies may rely on outdoor measurements or on measurements at a limited number of indoor sites, such measurements are regarded as surrogates for personal exposure measurements and, potentially, as inputs to estimation of personal exposure.

In discussing the role of total human exposure in air pollution epidemiology, a distinction should be made between measurements of exposure and dose. For this discussion, air pollution exposure is defined as some time-weighted function over a specific time interval of air pollutant concentrations in an individual's immediate breathing environment. Dose is defined as the amount of air pollutant, its constituents, or its by-products, delivered to the site of biologic or physiologic activity. In most situations, dose is neither well defined nor measureable, hence the focus of the subsequent discussion is on the estimation of air pollution exposure.

Ferris

-6-

Estimating personal air pollution exposure

Although individuals frequently have a complex temporal history of air pollution exposure, historical measurements are rarely available, and the time-weighted function of that history which produces the appropriate measure of total exposure is rarely known. The discussion here is on the simpler problem of producing an estimate of total exposure to an air pollutant under study during the period in which measurements are obtained. Duan (1982) has described two approaches to measurement of personal exposure

The first approach, called direct estimation, requires the personal monitoring of the environment of each study participant. Although instruments for personal exposure measurements are improving rapidly, the requirement that each participant carry or wear the instrument limits the application of this approach to pilot studies and special situations where environmental concerns produce strong commitments by study participants. For epidemiologic studies of general populations, less intrusive approaches are required. Duan's second approach, indirect estimation, is based on the notion that individual environments can be classified into a relatively small number of microenvironments, with relatively homogeneous air quality. In this approach, the i th individual's exposure can be represented as

$$E_i = \sum_j t_{ij} c_j$$

where t_{ij} is the fraction of time spent by the i th individual in the j th microenvironment, and c_j is the concentration of the pollutant of concern in that microenvironment.

The indirect approach is attractive when relatively few microenvironments are involved, when activity patterns are stable over time, and preferably, over

individuals, and when the most important microenvironments, that is those with the highest concentrations or exposure times, can be studied. To implement the indirect approach, one must develop estimates of the sojourn times, t_{ij} , for each individual, and the microenvironment concentrations, c_j .

An important limitation to this approach is the unestimated covariance of activity time and intensity, for example respiratory rate and depth, in the microenvironment's concentration. Furthermore, exposures to passive smoking and gas cooking need to be included and in adults occupational exposures have to be estimated as well. It may be advisable to study non-smokers who are not exposed to passive smoking and not exposed to dust at work.

Data collected in the Harvard Study suggest that the indirect approach can be of value, particularly for estimating exposure to pollutant with indoor sources. In the Portage, Wisconsin, study of personal exposure to NO_2 , indoor measurements in bedrooms, used in conjunction with outdoor measurements, explained 67 percent of the variation in week-long integrated personal exposures of adults and children. (Table 1). For children the predicted model explained in excess of 80 percent of the variance. Similar data have shown the importance of indoor measurements of particle concentrations. To pursue this strategy, a study has been initiated of the respiratory health effects in children of exposure to particulate air pollution and NO_2 which will involve indoor measurements in homes and at school for a projected sample of 1,800 school children (300 per city).

Study design

The study design is based on observations that third and fourth grade school children spend all of their time indoors at home, indoors at school, or outdoors, in that order. Moreover, activity patterns vary less among children

Ferris

-8-

than among adults, and should also be relatively stable over time. Since activity patterns and home air exchange rates differ in the winter and summer months, microenvironments have been further differentiated by season, producing a total of six microenvironments (Table 2). Our study design calls for collection of two one-week-long samples of NO₂ and fine particulate concentrations in each home during the winter and again during the summer months. Water vapor and home air exchange measurements will be made in conjunction with the pollution measurement. The size fractionated particle samples (2.5 um) will be analyzed for cigarette smoke and metals. In the initial study three rooms of the home will be monitored for NO₂. A single location is selected for particle measurement. Diaries will be kept by the parents of the daily respiratory symptoms and illnesses of the children. Outdoor concentrations and schools will be monitored continuously during the period of study.

The measurement of indoor environments promises to improve dramatically the understanding of the health effects of environmental exposure to NO₂. An analysis based on 1,800 distinct measurements of exposure and response will significantly refine present estimates of the exposure-response relationship for NO₂ at concentrations characteristic of present indoor and outdoor environments. In analyzing these data it will be important to recognize that indoor NO₂ exposure is accompanied by other constituents of gas stove emissions, and outdoor NO₂ is often associated with other components of automobile emissions.

The interpretation of indoor measurements of particulate concentrations will be even less straightforward. Cigarette smoking is a major contributor to indoor particulate burdens, and the chemical and physical characteristics of

cigarette smoke are very different than those of fossil fuel combustion products. Thus, the estimation of personal exposure to fine particulates must be approached more cautiously. Moreover, if the objective is to assess the health effects of environmental air pollution not due to personal health habits, the analysis and interpretation of the study data will be further complicated. The development of reliable markers for cigarette smoke in particulate samples is critically needed and work toward that objective is underway. The availability of indoor samples of particulate matter will also increase the need for methods of elemental analysis aimed at source attribution, and more measures of acidity. Thus, the effort to measure personal exposure to particulate matter will bring added importance to understanding the chemical and physical characteristics of particulate air pollution.

Implications for air pollution epidemiology

This chapter suggests that we are approaching the end of an era in air pollution epidemiology. Traditional approaches to exposure measurement seem unequal to the challenge of health risk assessment in a time of widespread and low-level exposure to air pollution. Since personal exposure and microenvironmental measurements are expensive, studies oriented toward measurements of total exposure will be smaller and more intensive. Since such studies allow analysis of individuals rather than groups, however, they need not be less powerful.

Implications for human risk assessment and regulation

A shift in focus toward emphasis on total human exposure will require fundamental changes in current approaches to the estimation of exposure

distributions in exposed populations. More importantly, however, such a shift would raise difficult issues about the shape of the Environmental Protection Agency's (EPA) regulatory domain. If exposure to NO_2 proves to be harmful, how should one attribute risk to indoor and outdoor exposure? One could make a proportional allocation of health consequences, but it is believed that the EPA should regard outdoor exposure as occurring in the context of exposure from other sources. Thus, the effects of regulatory action would be to alter exposure to NO_2 at the margin. The potential effects of regulation of outdoor air quality on health would be equal to the proportion of total NO_2 exposure originating outdoors only if the exposure-response is linear.

The regulatory issues are even more difficult for particulate air pollution. Although wood stoves and other commercial products may be subject to regulation, responsibility for the regulation of cigarette smoking would not be welcomed by the EPA. Still, the contribution of cigarette smoke to particle exposure must be considered in the interpretation of data from studies based on the concept of total human exposure. The potential difficulties make a return to traditional approaches to air pollution epidemiology seem inviting. Unfortunately, traditional approaches are no longer adequate to the scientific needs of health risk assessment. Thus, there is a need to address these issues directly and develop strategies for utilizing studies of total human exposure as the data become available.

REFERENCES

- Dockery, D.W. and Spengler, J.D. (1982). Indoor-outdoor relationships of respirable sulfate and particles. *Atmos. Environ.* 15:335-43.
- Duan, N. (1982). Models for human exposure to air pollution. *Environ. Int.* 8:305-9.
- National Academy of Science (1982). Committee on Risk and Decision Making. Risk and decision making: perspectives and research. National Academy Press, Washington, D.C.
- Sexton, K., Letz, R. and Spengler, J.D. (1983). Estimating human exposure to nitrogen dioxide: An indoor/outdoor modeling approach. *Environ. Res.* 32:151-66.
- Spengler, J.D., Dockery, D.W., Turner, W.A., Wolfson, J.M. and Ferris, B.G. Jr. (1981). Long-term measurements of respirable sulfates and particles inside and outside homes. *Atmos. Environ.* 15:23-30.
- Spengler, J.D., Ferris, B.G. Jr., Dockery, D.W. and Speizer, F.E. (1979). Sulfur dioxide and nitrogen dioxide levels inside and outside homes and the implications on health effects research. *Environ. Sci. Technol.* 13:1276-80.
- Spengler, J.D. and Soczek, M.L. (1983). Evidence for improved ambient air quality and the need for personal exposure research. *Environ. Res.* 32:151-66.
- Zeidberg, L.D., Schueneman, J.J., Humpfrey, P.A. and Prindle, R.A. (1961). Air pollution and health: General description of a study in Nashville, Tennessee. *J. Air Pollut. Control Assoc.* 11:289-94.

Ferris

-12-

TABLE 1: Comparison of estimated week-long (one in winter, one in summer) integrated exposures to nitrogen dioxide (NO₂) versus measurements using personal monitors. Portage, Wisconsin study, approximately 300 adults and children.

Model	R2	Root mean square error (ug/m3)
1. Ambient outdoor nitrogen dioxide (NO ₂) concentration	.01	14.9
2. Model 1 and gas versus electric stove (yes/no)	.41	11.8
3. Model 2 and bedroom concentrations	.67	8.8

TABLE 2: Percent time spent by third and fourth grade school children in selected microenvironments by season.

Microenvironments	Winter	Summer
Home indoors	55	70
School indoors	30	0
Outdoors	15	30

FIGURES

Fig. 1. Measured frequency distribution of indoor annual average nitrogen dioxide (NO_2) concentrations and fitted gamma distribution.

Fig. 2. Average rate of bronchitis and one SD error bars, adjusted for sex and age in several communities^a for a child's first three visits versus Total Suspended Particulate concentrations at those times.

^aPO- Portage

WA- Watertown

SL- St. Louis

TO- Topeka

KH- Kingston-Harriman

SB- Steubenville

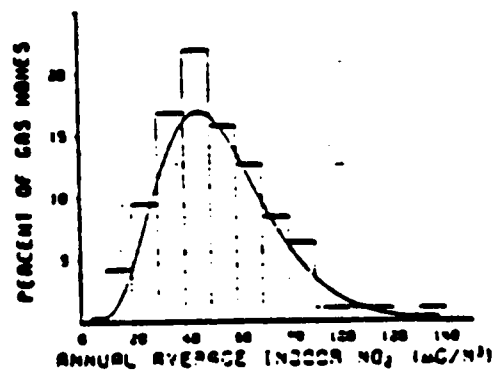


Fig. 1 Measured frequency distribution of indoor annual average NO₂ concentrations and fitted gamma distribution

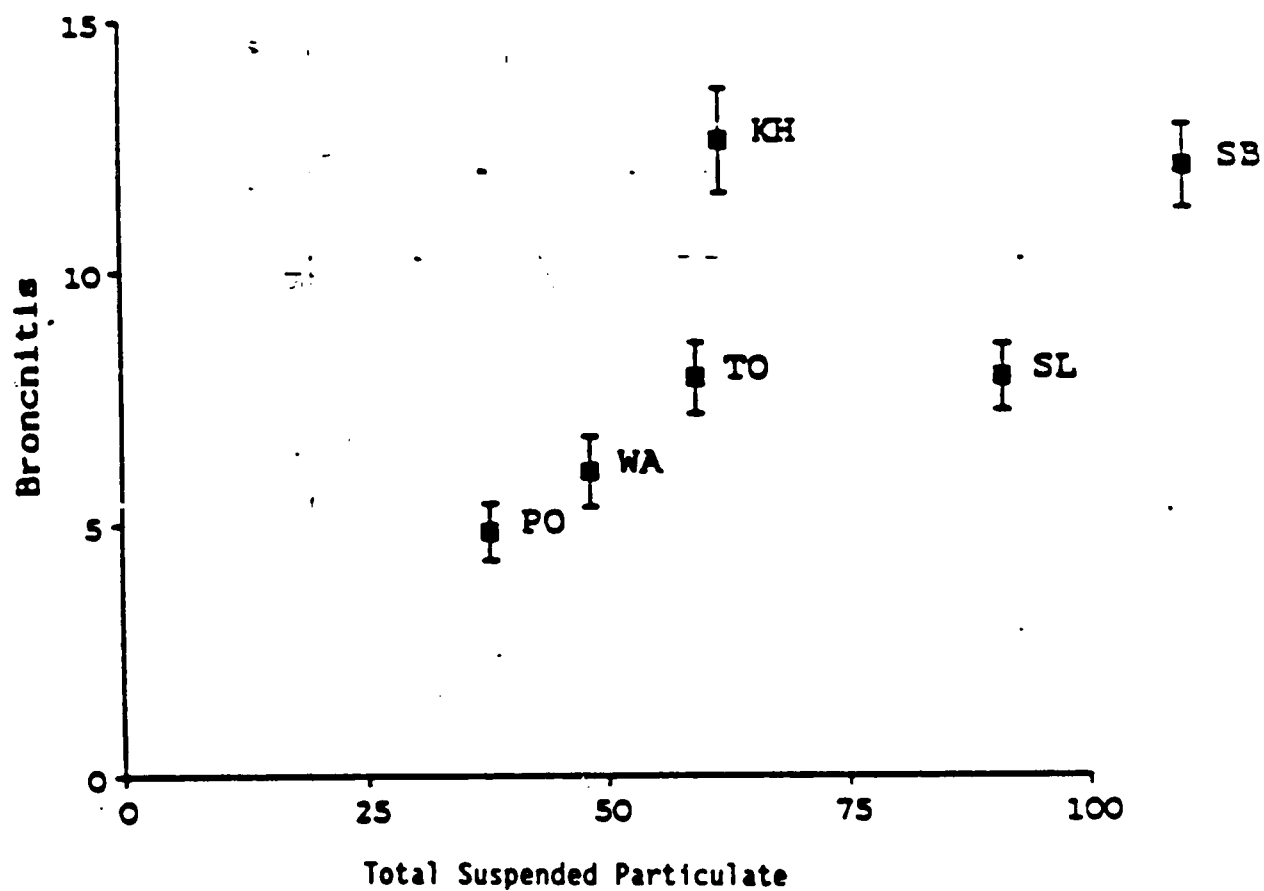


Fig 2: Average rate of bronchitis, adjusted for sex and age- in several communities for children's first three visits versus total suspended particulate concentrations at those times

Chapter 12

Measurement of Exposure to Arsenic in the Occupational Setting

Ian Higgins, Kathleen Welch and Mary Oh

MEASUREMENT OF EXPOSURE TO ARSENIC

IN THE OCCUPATIONAL SETTING

Ian T.T. Higgins, Kathleen Welch and Mary Oh

Measurement of exposure in the occupational environment would seem to be easier than the measurement of exposure in the general community. In an occupational setting, one has to deal with people performing more or less the same tasks each day who are easily accessible for study and who presumably are sufficiently interested in their own welfare to want to participate. However, while the difficult problem of assessing past exposures is common to both situations, relevant measurements are more likely to be made in the industrial setting than in the community setting. The following is a description of an experience in assessing exposure to arsenic trioxide in a cohort of copper smelter workers. The Anaconda Reduction Plant in Montana provides one of the best sources of data on exposure to arsenic. Hence it has been a focus of keen interest to those who would like to define as precisely as possible the quantitative relationship between exposure to arsenic and risk of disease, particularly the risk of lung cancer.

Earlier studies

The cohort was originally defined by Lee and Fraumeni (1969) of the National Cancer Institute (NCI). It comprised 8,047 smelter workers who had worked for twelve months or more before December 31, 1956 and who were originally followed up through 1963. Exposure in this initial study was based on three qualitative categories: heavy, medium and light, depending on the area in the plant where the worker had had his heaviest exposure. Departments were rated on a scale of 1 to 10, 8 to 10 being heavy, 4 to 7 medium and 1 to 3 light, from measurements

made in 1969 (Table 1). The cohort was also categorized by length of employment and period of employment.

Quantitative estimates of exposure were first applied to the NCI exposure categories by H. F. Morris (1975), Senior Hygienist at the Anaconda smelter. Morris derived average exposures for the heavy, medium and light categories of 11.270, .580 and .290 mg/m³ respectively.

Present studies

The Anaconda Company was interested in exploring the lower end of the dose response curve and the mortality data was updated to January 1, 1980 for this purpose. The Anaconda Company provided all of the available industrial hygiene data. These included reports by the State of Montana Board of Health and by the Industrial Hygiene Division of the Company, lists and charts of sampling locations in various departments, exposure factor surveys and Sampling Record Sheets in which arsenic concentrations measured between 1943 and 1965 were tabulated.

Eight hundred and twenty-six measurements of arsenic in air had been made during this period, 186 by the State and 640 by the Company. These measurements had been made at 66 locations, which could be aggregated into 17 of the 35 departments indicated in the individual work records. However it was not possible to match individual workers' exposures to the specific locations where the samples had been collected. The work records gave only the departments in which a man had worked but not the locations within those departments. Investigators were, therefore, faced with the problem of assessing department concentrations from concentrations measured at a number of sampling locations within each department. Locations were originally set up in 1943 by the State after a preliminary survey of the plant "to acquaint the

investigators with the processes involved and to determine the location of the hazards." There were different numbers of locations in each department. About half the departments were not sampled at all because arsenic was not thought to be a problem in them. The locations represented different areas in a department and often different operations. The original sampling locations were later changed to provide more detail or a more logical grouping within departments. The greatest changes were made in the reverberatories, the copper roaster and the converters.

Sampling for arsenic was not evenly distributed over time, and became more infrequent over the entire time period (Figure 1). Samples from the lime rock crusher and ferro-manganese departments were taken only one time for each department (1958), on the local tram two times (1943 and 1944) and in the masons departments and zinc roaster they were taken three times for each department (1943, 1944 and 1945). When the concentrations in the various departments are examined for each year decreases in the arsenic concentrations are apparent (Figures 2-4). Because of the difficulty in determining the significance of such decreases over time, the average values of all of the samples over the entire period were used to develop estimates for each department.

Exposure factors

Exposure factors were developed by the Anaconda Industrial Hygiene Division in order to reflect the proportion of the work shift spent in a contaminated atmosphere. They are simply defined as: hours exposed and hours worked. Sometimes exposure factors were calculated for specific jobs based on time and motion. Often they were directly estimated for particular sampling locations, based on the best judgement of the Industrial Hygienist. Exposure factors

ranged from .06 to 1.0. As might be expected, a smaller proportion of the work shift was spent in departments where the arsenic concentrations were estimated to have been high. There were problems in reconstruction of exposure factors as given in the sample record sheets from information recorded in special surveys. Thus, the validity of exposure factors is somewhat questionable. Even if exposure factors could have been unambiguously assigned to each sampling location, there is reason to question the appropriateness of their use. Given that men were working in departments contaminated by arsenic, the actual concentrations rather than those concentrations reduced by a somewhat arbitrary factor at locations within those departments perhaps better reflects their exposure. These difficulties necessitated analysing the data with and without exposure factors.

Mean concentrations for various locations within departments are shown by departmental category of exposure (very high, high, medium, or low) in Tables 2-5. Note the varying number of locations and samples in each department, the exposure factor in the very high departments is usually less than 0.5 but about 1.0 in the low departments. Estimates of departmental mean arsenic concentrations and Weighted Average Exposure Concentrations (WAEc) from these location samples are shown in Table 6. These arsenic concentrations were calculated as follows: In departments where arsenic concentrations had been measured, the mean of all concentrations at a given location within the department for all years was calculated (1943-1965). This location mean was then multiplied by the appropriate exposure factor to obtain the location WAEc. The mean of all the location concentrations was then taken to obtain the department ambient concentration. Similarly, the mean of the location WAEc was taken to obtain the department WAEc.

In departments where no arsenic measurements had been made, the following procedures were used. For maintenance, surface, shops and unknown departments it was assumed that these employees worked all over the plant and had exposures which were high at times and low at other times. Their average exposure was taken as the average of the heavy, medium and low departments. Lead shop, lime processing, engineering, power and industrial hygiene departments were assumed to have received little arsenic exposure. They were assigned the mean value of all measured low departments. Other unmeasured departments were assigned the same exposure as those measured departments which were similar to them. Thus, the brick yard and coal pile were similar to outside; concentrator similar to lab sample mill, general office and research similar to acid plant and company farms and foundry similar to the town of Anaconda.

Average departmental concentrations and average WAECs were thus available for each department. Indices of individual arsenic exposure were then estimated for each man from work records which were available from the personnel department. These records listed each department in which a man had worked with his dates of starting and stopping. Cumulative lifetime exposure was obtained by multiplying the estimated departmental arsenic concentrations by the number of years worked in each department and summing the quotient for all departments. Precisely the same procedure was used to obtain cumulative lifetime WAEC.

Indices other than cumulative lifetime exposure were also used in these studies. For example, individual exposures were estimated as a time weighted average by dividing cumulative lifetime exposure by the total number of years of exposure. The significance of peaks of exposure were also assessed by noting the highest exposure that a man had experienced for 30 days or more.

These were referred to in an earlier publication as "ceiling" exposures (Welch, Higgins, Oh et al. 1982).

Table 7 and Figure 5 shows the relationship of cumulative lifetime arsenic exposure and mortality from respiratory cancer. Standardized Mortality Ratios (SMRs) comparing the plant cohort experience with that which would be expected for white males in the State of Montana for each exposure category, is shown. There was a clear dose-response relationship from about 50 percent above expectation in the lowest exposure group to a four fold excess in the highest exposure category.

Once individual exposure estimates are obtained, it is possible to explore dose-response relationships in various ways: relative or absolute risk, and the influence of age, age at hire, period of hire, latency and years of work. The object of this discussion has been to indicate the kind of problems encountered in the Anaconda study, the assumptions made and the estimates reached about arsenic exposure. Clearly, others may disagree with what was done and there is no doubt that further analyses could usefully be made. However, different ways of assessing exposure lead to broadly similar conclusions. The important conclusions, however, from the standpoint of risk assessment and public policy are that exposure estimates from available data depend on many assumptions and are likely to be debatable, any dose-response curves drawn from them, are somewhat tentative, and, therefore, any safe levels of arsenic in the work place which may be deduced, are subject to a rather wide range of uncertainty.

In answer to the question, "How can we do it better?", the importance of direct measurement of exposure should be stressed. Indirect assessment of exposure based on integrating a worker's movements with concentrations measured

at fixed sites is likely to be very inaccurate. Instead, random samples of workers should be selected, their exposure during the work shift should be measured sufficiently frequently to obtain reliable estimates of the lifetime occupational exposure experienced in those jobs. Only in this way will the epidemiologists of the year A.D. 2010, if there are any, be provided with the sort of information they will need for their cumulative lifetime exposure-response curves.

REFERENCES

- Lee, A.M. and Fraumeni, J.F. Jr. (1969). Arsenic and respiratory cancer in humans: Follow-up of copper smelter workers in Montana. J. Natl. Cancer Inst. 42:1045-52.
- Morris, H.F. (1975). Statement at OSHA Hearings, Washington, D.C.
- Welch, K., Higgins, I.T.T., OH, M. and Burchfield, C. (1982). Arsenic exposure, smoking, and respiratory cancer in copper smelter workers. Arch. Environ. Health, 37:325-35.

TABLE 1. Exposure categories for 8,047 copper smelter workers by department, Anaconda Reduction Plant, Montana, 1956-1963

Department	Scale	Qualatative Level	Morris's Average (Mg/m ³)
Arsenic roaster Stack cottrells Arsenic kitchen	8 - 10	Heavy	11.278
Converter Reverberatory furnaces Ore roaster Acid plant	4 - 7	Medium	.580
All other	1 - 3	Light	.290

Source: Lee and Fraumeni (1969); Morris (1975)

TABLE 3: Mean arsenic concentrations (Mg/m³) by sampling location in departments with high exposures, Anaconda Reduction Plant, Montana, 1943-1965.

Department	Location	Number Samples	Mean	Exposure Factor	Weighted Average Exposure Concentration
Copper					
roaster	II	1	.001	.50	.0005
	IV	15	.181	.50	.091
	V	43	.091	.40	.036
	VI	3	.123	.13	.016
	VII	8	.244	.50	.122
	VIII	3	133.43	.06	8.006
Reverberatory					
furnaces	I	45	.850	.50	.425
	II	50	2.562	.50	1.281
	III	59	1.646	.50	.823
	IV	59	1.173	.50	.586
	V	24	.963	.40	.385
	VI	18	1.338	.40	.535
	VII	39	.803	.40	.321
	VIII	37	.593	.40	.237
	IX	7	.079	.88	.069
Materials crushing (Lime rock crusher)					
	II	2	.820	.50	.410
	IV	2	8.450	.50	4.225
	V	1	.220	.50	.110
Masons					
	I	7	2.937	.88	2.59
	II	9	2.298	.20	.460
	III	6	20.572	.20	4.114
	IV	6	54.242	.06	3.255

TABLE 2: Mean arsenic concentrations (Mg/m³) by sampling location in departments with very high exposures, Anaconda Reduction Plant, Montana, 1943-1965.

Department	Location	Number Samples	Mean	Exposure Factor	Weighted Average Exposure Concentration
Arsenic					
roaster	I	14	68.348	.44	30.073
	II	33	2.798	.38	1.063
	III	13	81.969	.31	25.410
	IV	6	168.750	.13	21.938
Arsenic					
refinery	I	7	470.70	.13	16.612
	II	19	50.60	.62	10.024
	III	10	143.10	.37	26.173
	IV	5	23.540	.13	1.347
	V	9	25.900	.25	1.180
	VI	8	5.890	.75	1.252
	VII	5	47.100	.06	1.488
	VIII	4	10.42	.50	2.646
Stack					
cottrells	I	32	.126	.25	.032
	II	25	90.435	.25	22.609
Main Flue	I	11	52.992	.38	20.137
	II	9	2.298	.20	.460
	III	6	20.572	.20	4.114
	IV	6	54.242	.06	3.255

TABLE 4: Mean arsenic concentrations (Mg/m³) by sampling location in departments with medium exposures, Anaconda Reduction Plant, Montana, 1943-1965.

Department	Location	Number Samples	Mean	Exposure Factor	Weighted Average Exposure Concentration
Local tram	I	5	.494	1.00	.494
New					
converters	I	19	.135	1.00	.135
	II	5	.144	1.00	.144
	III	12	.582	1.00	.582
	V	27	.582	1.00	.112
	VI	1	.01	.50	.005
	IX	7	.261	.50	.130
	X	4	.525	1.00	.525
Zinc					
roaster	II	7	.076	.25	.079
	III	6	.094	.40	.038
	V	1	.390	.38	.148
	VI	1	1.69	.06	.101

TABLE 5: Mean arsenic concentrations (Mg/m^3) by sampling location in departments with low exposures, Anaconda Reduction Plant, Montana, 1943-1965.

Department	Location	Number Samples	Mean	Exposure Factor	Weighted Average Exposure Concentration
Casting	IV	12	.164	1.00	.164
	VII	3	.039	1.00	.039
	VIII	6	.030	.67	.020
Errromanganese	III	3	.035	1.00	.035
	V	3	.040	1.00	.040
	VI	6	.089	1.00	.089
	X	1	.30	1.00	.30
Phosphate Roaster	V	7	.049	.5	.024
	X	3	.080	.75	.060
	XIII	3	.057	1.00	.057
	99	7	.004	1.00	.004
Lab sample mill	I	1	.060	.25	.015
	II	21	.126	.25	.032
	III	6	.068	.75	.051
Acid-plant	XI	2	.012	1.00	.012
Outdoors	I	1	.08	1.00	.08
	II	1	.05	1.00	.05

TABLE 6: Selected mean arsenic concentrations by exposure category for selected departments, Anaconda Reduction Plant, Montana, 1943-1965.

Exposure Category	Department	Mean Arsenic Concentrations (Mg/m ³)	
		Ambient	WAEC
Very high	Arsenic roaster	80.466	19.621
	Stack cottrells	32.691	11.320
	Arsenic refinery	45.281	7.590
	Main flue	32.526	6.991
High	Masons shop	2.937	2.585
	Copper roaster	22.345	1.379
	Materials crushing	3.163	.807
	Reverberatories	1.112	.518
	Maintenance	NA	.546
	Surface	NA	.546
	Shops	NA	.546
	Unknown	NA	.546
Medium	Tram	.494	.494
	Converter	.258	.233
	Zinc roaster	.563	.077
Low	Ferrumanganese	.116	.116
	Casting	.078	.074
	Outside	.065	.065
	Packyard	NA	.065
	Coal pile	NA	.065
	Lead shop	NA	.048
	Lime processing	NA	.048
	Engineering	NA	.048
	Power	NA	.048
	Industrial hygiene	NA	.048
	Phosphate plant	.048	.038
	Lab sample mill	.118	.033
	Concentrator	NA	.033
	Acid plant	.012	.012
	General office	NA	.012
	Research	NA	.012
	Anaconda town	NA	.00045
	Company farms	NA	.00045
	Foundry	NA	.00045

TABLE 7: Relationship of cumulative lifetime exposure to arsenic and mortality for respiratory cancer, 8044 male copper smelter workers at Anaconda Reduction Plant, Montana followed from 1938-1980

Arsenic Exposure (UG/m ³ yrs)	Respiratory Cancer		
	Observed	Expected	SMR
<500	51	35.9	142*
500-	67	37.5	179**
2000-	122	46.0	265**
12000 & over	98	24.8	396**

NUMBERS OF SAMPLES OVER THE YEARS

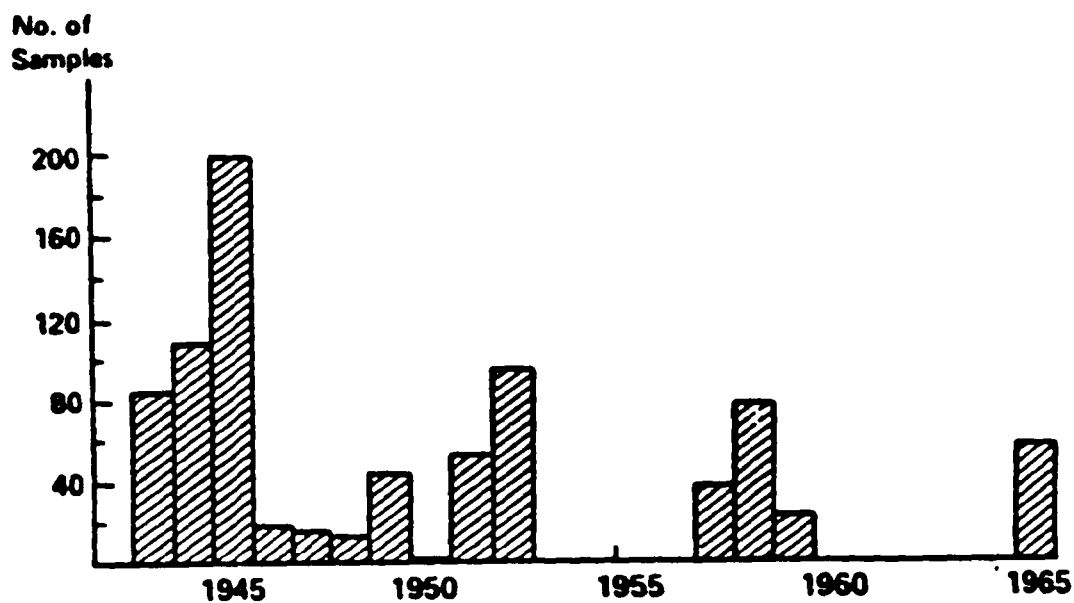


Figure 1

Numbers of Samples for Arsenic From Selected Departments
in the Anaconda Reduction Plant, Montana, 1943-1965

Numbers of Samples Over the Year

Higgins

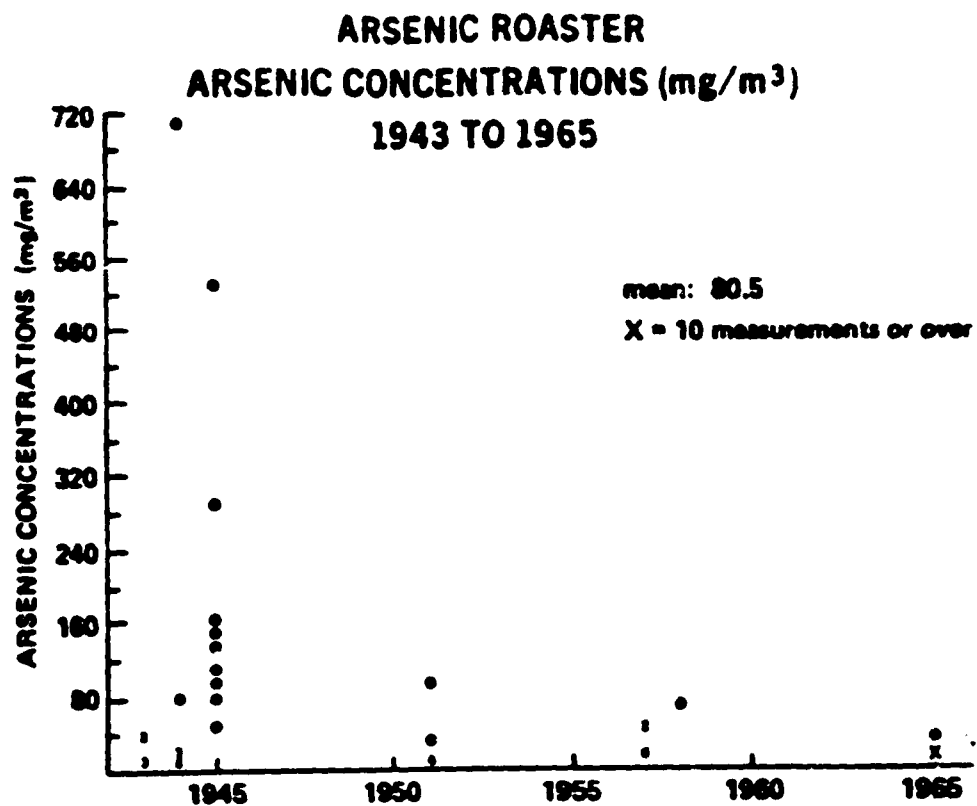


Figure 2

Arsenic Concentrations (mg/m³) from the Arsenic Roaster
Department, Anaconda Reduction Plant, Montana, 1943-1965

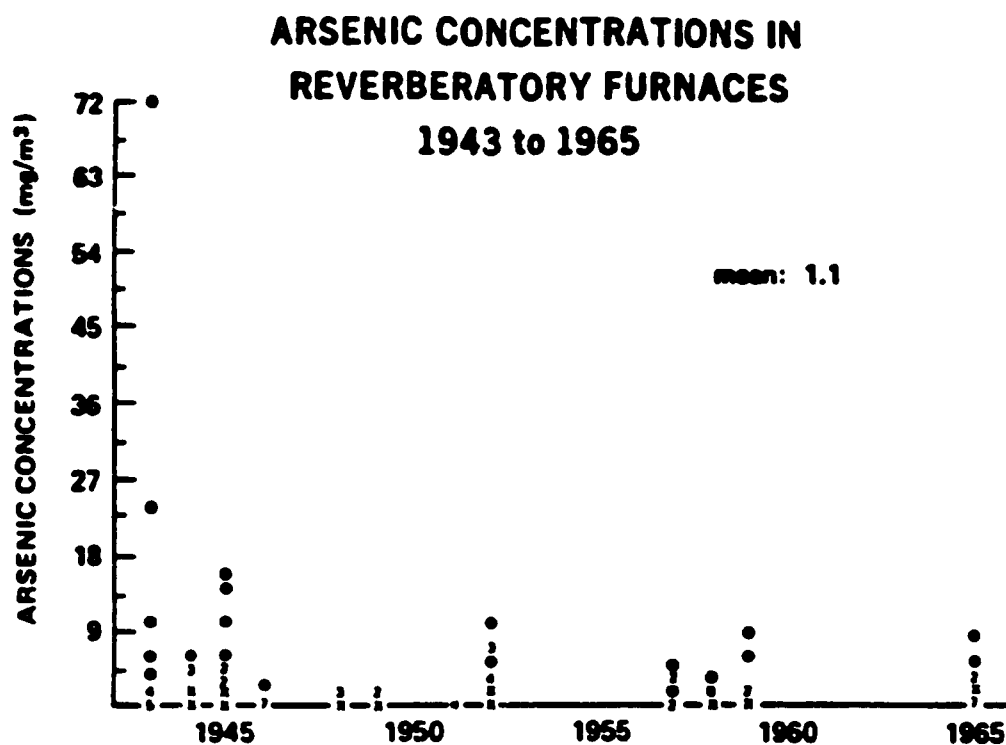


Figure 3

Arsenic Concentrations (mg/m³) in Reverberatory Furnaces
Anaconda Reduction Plant, Montan, 1943-1965

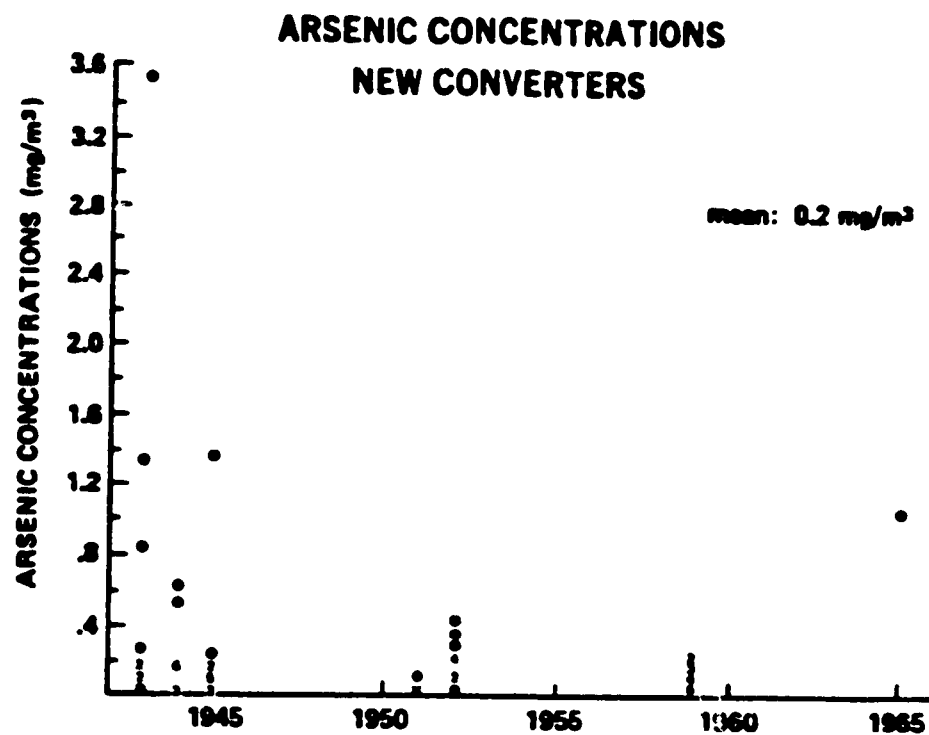


Figure 4

Arsenic Concentrations (mg/³) in New Converters
Anaconda Reduction Plant, Montana, 1943-1965

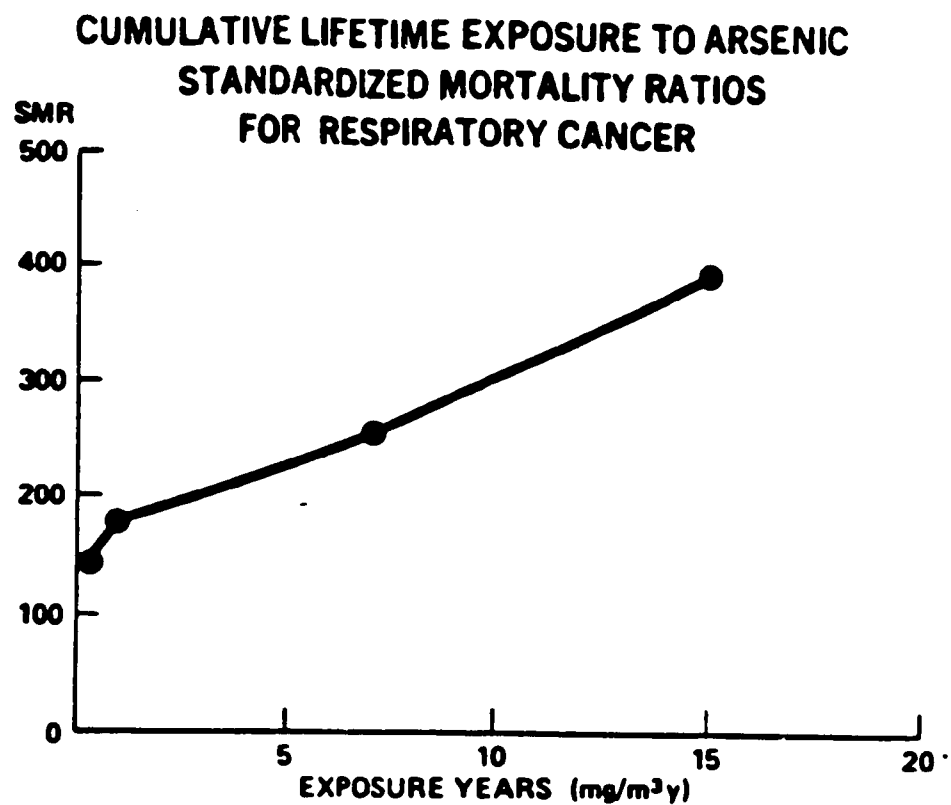


Figure 5

Cumulative Lifetime Exposure to Arsenic and Standardized Mortality Ratios for Respiratory Cancer

8044 Male Copper Smelting Workers at Anaconda Reduction Plant, followed up from 1938-1980.

Part III - Section B
Measures in the Individual

Chapter 13

Relation of Body Burden Measures
to Ambient Measures

Philip J. Landrigan

RELATION OF BODY BURDEN MEASURES TO AMBIENT EXPOSURES

Philip J. Landrigan

INTRODUCTION

The body burden of a toxin is the total amount of that material in the human body. The body burden represents the difference between cumulative lifetime absorption and total excretion.

Measurement of body burden provides a precise and individual assessment of exposure (Nordberg 1976; Wielopolski, Rosen, Slatkin et al. 1983). Such measurements are of considerable value in epidemiologic risk assessment; they complement the measurement of ambient exposure. They also provide a means to study quantitatively the relations between ambient exposures and absorption. Body burden measurements are free of recall and misclassification biases which may confound efforts to assess exposure through oral history (Najarian and Colton 1978; Greenberg, Rosner, Henneken et al. 1985). The measurement of body burden provides a means for assessing toxic exposures which occurred before other means of environmental assessment became available (Ericson, Shirahata and Patterson 1979; Needleman, Gunnoe, Leviton et al. 1979).

The purpose of this chapter is to illustrate the application of body burden measurement to epidemiologic analysis. The principles underlying such applications shall be discussed as will the inherent limitations of body burden measurement. The ethical implications of measuring body burdens shall be presented as well.

Principles underlying assessment of body burden

The body burden of very few substances is stable over time. Further, most toxins are not distributed homogeneously in the human body. Instead, toxins are typically dispersed across several physiologically distinct compartments, which differ from one another in their size and relative accessibility. It is

important to recognize that these compartments are defined solely by the dynamic behavior of the toxin contained in them. There is not necessarily a one-to-one correspondence between the physiologically defined compartments and specific anatomical locations (Hattis 1981). These compartments and their interrelations may be portrayed by metabolic models (Rabinowitz, Wetherill and Kopple 1976; Bernard 1977). Lead provides a classic example (Landrigan, Froines and Mahaffey 1985) (Figure 1).

Lead

Lead, when it is initially absorbed by any route produces a rapid increase in the blood lead level. Thus, measurement of the blood lead level has been used to follow body burdens of new entrants to the lead industry (Tolz, Hernberg, Asp et al. 1973). Also the blood lead level has been used to monitor lead burdens in children (Landrigan, Baker, Feldman et al. 1976). The following case-study illustrates the use of blood lead measurement as an index of high-dose lead exposure in children.

To assess body lead burden in children living near a large lead smelter in Kellogg, Idaho, venous blood samples were obtained in 1974 from a statistical sample of one to nine year-olds in towns adjacent to the smelter (Landrigan, Baker, Feldman et al. 1976). Blood lead levels indicative of increased lead absorption (≥ 40 ug/dl) were found in 385 (41.9 percent) of 919 children. In 1974, an excessively high blood lead level was defined as a blood lead concentration ≥ 40 ug/dl (Surgeon General, U.S. Public Health Service 1971). Today the corresponding definition is a blood level ≥ 25 ug/dl (Centers for Disease Control 1985). Within 1.6 kilometers of the smelter, 170 (98.8 percent) of 172 children had elevated blood lead levels, and 38 had levels ≥ 80 ug/dl, the highest being 164 ug/dl (Figure 2). The prevalence of increased lead absorption declined with distance, and in the most distant control area,

only one of 89 children had an elevated level. These data demonstrate clearly that there existed extensive increased body burden of lead in children in Kellogg and that this increased burden was associated geographically with the smelter (Landrigan, Baker, Feldman et al. 1976).

Air lead levels were also elevated near the smelter (Landrigan, Baker, Feldman et al. 1976) (Figure 3). Immediately adjacent to the plant the mean air lead level in July 1974, one month prior to the blood survey, was 19 $\mu\text{g}/\text{m}^3$. To assess the relationship between air lead and blood lead levels in this population, followup studies were undertaken. A positive dose-response relationship was found. It was also found that the relationship between air lead and blood lead levels was approximately linear and that each 1 $\mu\text{g}/\text{m}^3$ increase in air lead concentration produced a mean increase in the blood lead level of 1.27 $\mu\text{g}/\text{dl}$. Similar studies in other populations of children have shown that each 1 $\mu\text{g}/\text{m}^3$ increase in air lead level produces an increase of 1 to 3 $\mu\text{g}/\text{dl}$ in the mean blood lead level (Yankel, VonLindern and Walter 1977). These observations provided a major basis for the epidemiologic risk assessment which underlies the EPA's ambient standard for airborne lead (U.S. Environmental Protection Agency 1985).

Long continued exposure to lead results in a different pattern of accumulation. Chronic exposure produces sequestration of lead in various tissue compartments, principally in bone. Thus in persons with chronic lead exposure, such as lead workers, 99 percent of absorbed lead is ultimately deposited in bone (Rabinowitz, Wetherill and Kopple 1976). In such persons, the blood lead level is only an imperfect indicator of current exposure, or of total body burden, because the leaching of lead from bone can exert a dominant influence on the blood lead level (Baker, Landrigan, Barbour et al. 1979). This point is illustrated by the many cross-sectional studies of long-term lead

workers which have found only poor correlations between air lead and blood lead levels (U.S. Department of Labor, Occupational Safety and Health Education 1978). The influence of body lead stores on blood lead levels is illustrated also by followup studies of lead workers who had been laid off employment by a strike (Hattis 1981). In these workers blood lead levels required many months to fall to within the normal range. The half-life in the body of the least readily mobilized lead stores is estimated to be as long as 25 years (Rabinowitz, Wetherill and Kopple 1976).

In summary, these data illustrate that in the early phases of exposure to lead, the blood lead level rises rapidly and provides a sensitive indicator of body lead burden. Also under equilibrium conditions, such as pertain in the general population, the blood lead level is a good indicator of recent absorption and serves therefore as a sensitive surrogate for measurement of total body burden (Hernberg 1980). However, in persons with long-term heavy exposures to lead, the blood lead level correlates only poorly with total lead burden. In such circumstances, it may be appropriate in the future to measure bone lead burden directly by X-ray fluorescence analysis (Wielopolski, Fosen, Slatkin et al. 1983).

Polybrominated biphenyls (PBBs)

The following case-history illustrates the use of a body burden measure to document past exposure and to provide a basis for future epidemiologic study (Landrigan, Wilcox, Silva et al. 1979).

In 1973, several hundred pounds of PBB were introduced into cattle feed as the result of a shipping accident in Michigan. Thousands of dairy animals died or were destroyed, and widespread human exposure resulted. In October 1973, a survey of randomly selected samples of human breast milk showed that 50 (36 percent) of 53 samples from the lower peninsula of Michigan

contained detectable levels of PBB (Brilliant, Van Amburg, Isbister et al. 1978).

To determine the effects of PBB exposure on human health, a statewide epidemiologic cohort study was undertaken. Invited to join the cohort were residents of contaminated farms, persons who had directly received contaminated farm products, and chemical workers in the plant which produced PBB. Volunteers were also admitted, but were kept in a separate category.

All persons enrolled in the study were visited in their homes in 1976 and 1977. A venous blood sample was requested from each participant, and 3,639 were obtained. Samples were analyzed for PBB content.

Data on serum PBB levels showed a range of values from 0-1900 parts per billion (ppb); 106 values were greater than 100 ppb. Children aged ten years or less tended to have the highest concentrations. Males had significantly higher values than females ($\chi^2 = 256.5$; $p < 10^{-10}$). PBB levels showed a strong tendency to cluster by families (intra-class coefficient of correlation = 0.85).

Highest PBB levels were found in chemical workers and their families followed by residents of quarantined farms (Table 1). In the chemical worker group, the workers themselves had a mean serum PBB level of 108.7 ppb (range, 0-1, 240 ppb) significantly higher than that of their family members. Volunteers had the lowest PBB levels.

To assess time trends in PBB body burden, paired serum samples, one in 1974 and the other in 1977, were collected from 148 persons. Analysis of these samples indicates that serum levels were generally stable over the three year period. The mean change in serum PBB level was only 16 ppb, and the coefficient of correlation (r) between the two sets of values was 0.96.

Because PBB is lipophilic, its body burden is concentrated principally in adipose tissue. To assess this burden, simultaneous paired specimens of blood and adipose tissue were obtained on 221 Michigan residents. A high coefficient of correlation (r) was seen between these values (0.951), and in the 132 instances where both values were above detection limits, the ratio of adipose to serum concentration was 362.1 (Figure 4)

The body burden data from this cohort confirm the occurrence of PBB exposure in Michigan residents. In the absence of any direct information on environmental exposures, these data provide a unique basis for future epidemiologic assessment.

Such assessment will be necessary, because persons exposed to PBB may be at risk of delayed reactions, particularly cancer. PBB fed to rats has produced neoplastic liver nodules (Kimbrough, Burse and Liddle 1977). The Michigan Department of Public Health and the federal health agencies plan to follow the health status of the Michigan PBB cohort for at least another 10 to 15 years.

Limitations of body burden measures

Measurement of body burden can provide an extremely useful foundation for epidemiologic risk assessment. However, the body burden of few toxins is stable over time or homogeneous from tissue to tissue. The measured body burden of any substance reflects a dynamic equilibrium between absorption and excretion. Measurement of the wrong compartment at the wrong time or measurement of any compartment without an understanding of the underlying toxicokinetics can produce misleading results.

Recognition of time trends is essential in understanding body burdens. Volatile organic solvents, such as styrene, remain in the body for only a few

hours (Lauwerys 1984). The blood lead level reflects exposures which have occurred over the past several months (Rabinowitz, Wetherill and Kopple 1976). The bone lead level (Rabinowitz, Wetherill and Kopple 1976) and the serum level of PBB are, on the other hand, extremely stable over many years (Landrigan, Wilcox, Silva et al 1979). Further, the composition of a body burden can change over time. For example, following the absorption of commercial mixtures of polychlorinated biphenyls (PCBs), the less highly chlorinated PCB molecules have been found to be metabolized and excreted within weeks after exposure, whereas the more highly chlorinated species decay only after months or years (Kimbrough 1980). This differential metabolism is of importance in any attempt to fingerprint exposures through chemical analysis of the body burden.

Ethical constraints on body burden measures

Measures of body burden fail generally to provide adequate information on short-term changes in exposure. Except in the case of very rapidly excreted toxins, the body burden measure responds only sluggishly to environmental fluctuations. For that reason, sole reliance on measurement of body burden in populations with ongoing exposure, such as industrial workers, is neither scientifically defensible nor ethical. In such groups, measurement of the body burden must be coupled with environmental monitoring, which is aimed at identifying and controlling short-term excursions in exposure. Measurement of body burden in such populations, without measurement of the environment, is tantamount to using the exposed persons as environmental monitors.

A second ethical constraint on the measurement of body burden stems from its potential to label individuals who may be at heightened risk of disease. Senator Albert Gore, Jr. said of genetic screening that it has the "potential to serve as a marvellous tool to protect the health of workers or as a terrible vehicle for insidious discrimination" (Rothstein 1984). The same maxim applies to measurements of body burden.

If body burden measures are truly to serve as a basis for disease prevention, the following precepts must be observed in their application:

- Measurement of body burden cannot ethically be undertaken without the informed consent of the persons to be evaluated.
- Measurement of body burden can proceed ethically only in an open atmosphere, in which the persons assessed are informed promptly of their results and of the meaning of those results. Possible uncertainties in interpretation must be conveyed.
- Measurement of body burden is unethical unless it is intended to lead to preventive action, such as medical removal protection (U.S. Department of Labor, U.S. Occupational Safety and Health Administration 1978).
- Finally, medical removal protection, which is based on measurement of body burden, must be linked inextricably to provisions for job protection and for rate retention. For example, under the lead standard of the Occupational Safety and Health Administration, a worker with an elevated blood lead level who is medically removed to a "safe" job, is entitled to maintain the wages and the benefits of his or her former job for as long as 18 months (U.S. Department of Labor, Occupational Safety and Health Administration 1978). This provision prevents the dismissal from employment of lead-burdened workers and their replacement by new employees. Such an approach results in the cleanup of the workplace rather than in the involuntary turnover of the workforce (Rothstein 1984). This approach is consistent with the highest ethical traditions of industrial hygiene, occupational medicine and public health.

REFERENCES

- Baker, E.L. Jr., Landrigan, P.J., Barbour, A.G., Cox, D.H., Folland, D.S., Ligo, R.N. and Throckmorton, J. (1979). Occupational lead poisoning in the United States: Clinical and biochemical findings related to blood lead levels. *Brit. J. Industr. Med.* 36:314-22.
- Bernard, S.F. (1977). Dosimetric data and metabolic model for lead. *Health Physics* 32:44-6.
- Brilliant, L.B., Van Amburg, G., Isbister, J., Humphrey, H., Wilcox, K., Eyster, J., Bloomer, A.W. and Price, H. (1978). Breast milk monitoring to measure Michigan's contamination with polybrominated biphenyls. *Lancet* 2:643-46.
- Centers for Disease Control (1985). Preventing lead poisoning in young children: A statement from the Centers for Disease Control, US Public Health Service, Atlanta.
- Ericson, J.E., Shirahata, H. and Patterson C.C. (1979). Skeletal concentrations of lead in ancient Peruvians. *New Engl. J. Med.* 300:946-51.
- Greenberg, E.R., Rosner, B., Hennekens, C., Rinsky, R. and Colton, T. (1985). An investigation of bias in a study of nuclear shipyard workers. *Amer. J. Epidemiol.* 121:301-08.
- Hattis, D. (1981). Dynamics of medical removal protection for lead — a reappraisal. Cambridge: Massachusetts Institute of Technology, Center for Policy Alternatives, Report no. CPA-81-25.
- Hernberg, S. (1980). Biochemical and clinical effects and responses as indicated by blood lead concentration. In: *Lead Toxicity*, (eds. R.L. Singal and J.A. Thomas), Urban and Schwarzenberg, Baltimore, (pp 367-99).

- Kimbrough, R.D. (1980). Halogenated biphenyls, terphenyls, naphthalenes, dibenzodioxins and related products. Elsevier Scientific Publishing Co., Amsterdam.
- Kimbrough, R.D., Burse, V.W. and Liddle, J.A. (1977). Toxicity of polybrominated biphenyl. *Lancet* 2:602-03.
- Landrigan, P.J., Baker, E.L. Jr., Feldman, R.G., Cox, D.H., Eden, K.V., Orenstein, W.A., Mather, J.A., Yankel, A.J. and Von Lindern, I.H. (1976). Increased lead absorption with anemia and slowed nerve conduction in children near a lead smelter. *J. Pediatr.* 89:904-10.
- Landrigan, P.J., Froines, J.R. and Mahaffey KR (1985). Body lead burden: Summary of epidemiologic data on its relation to environmental sources and toxic effects. In: *Health Consequences of Dietary and Environmental Exposure to Lead*, (ed. K.R. Mahaffey). Elsevier/North Holland Biomedical Press, Amsterdam.
- Landrigan, P.J., Wilcox, K.R., Silva, J. Jr., Humphrey, M.E.B., Kauffman, C., and Heath, C.W. Jr. (1979). Cohort study of Michigan residents exposed to polybrominated biphenyls: Epidemiologic and immunologic findings. *Ann. N.Y. Acad. Sci.* 320:284-94.
- Lauwerys, R. (1984). Styrene. In: *Biological Indicators for the Assessment of Human Exposure to Industrial Chemicals*, (eds. L. Alessio, A. Berlin, M. Boni and R. Foi), Commission of the European Communities, Luxembourg.
- Najarian, T. and Colton, T. (1978). Mortality from leukemia and cancer in shipyard nuclear workers. *Lancet* 1:1018-20.
- Needleman, H.L., Gunnoe, C., Leviton, A., Reed, R., Peresie, H., Maher, C., and Barrett, P. (1979). Deficits in psychologic and classroom performance of children with elevated dentine lead levels. *New Engl. J. Med.* 300:689-95.

- Nordberg, G. (1976). Effects and dose-response relationships of toxic metals.
In: (ed. GF Nordberg): Effects and Dose Response Relationships of Toxic Metals, Elsevier Scientific Publishing Co., Amsterdam.
- Rabinowitz, M.B., Wetherill, G.W. and Kopple, J.D. (1976). Kinetic analysis of lead metabolism in healthy humans. J. Clin. Invest. 58:260-70.
- Rothstein, M.A. (1984). Medical screening of workers. Washington Bureau of National Affairs.
- Surgeon General, U.S. Public Health Service (1971). Medical aspects of childhood lead poisoning. Pediatrics 48:464-68.
- Tola, S., Hernberg, S., Asp, S. and Nikkanen, J. (1973). Parameters indicative of absorption and biological effect in new lead exposure: A prospective study. Brit. J. Industr. Med. 30:134-41.
- U.S. Department of Labor, Occupational Safety and Health Administration (1978). Occupational exposure to lead: Final standard. Federal Register 43:52952-3014.
- U.S. Environmental Protection Agency (1985). Air quality criteria for lead. U.S. EPA Publication No. EPA-600/8-83-028A.
- Wielopolski, L., Rosen, J.F., Slatkin, D.N., Vartsky, D., Ellis, K.J., Cohn, S.H. (1983). Feasibility of non-invasive analysis of lead in the human fibia by soft X-ray fluorescence. Med. Phys. 10:248-51.
- Yankel, A., VonLindern, I. and Walter, S.D. (1977) The Silver Valley lead study: The relationship between childhood blood lead levels and environmental exposure. J. Air Pollut. Control Assoc. 27:763-67.

TABLE 1: Serum PBB^a levels by selected enrollment groups
Michigan PBB Study, 1976-77

Group	Number	Range	Mean	Median
Quarantined farm residents	1750	0-1900	26.9	4.0
Farm product recipients	1114	0-659	17.1	3.0
Chemical workers and families	216	0-1240	43.0	4.5
Pilot study control participants	44	1-13	3.5	2.0
Low-level PBB farm residents	242	0-24	3.5	2.0
Volunteers	273	0-111	3.2	1.0
Total study population	3639	0-1900	21.2	3.0

^a Polybrominated biphenyls (PBB) levels in part per billion (ppb)
Source: Landrigan, Wilcox, Silva et al. (1979)

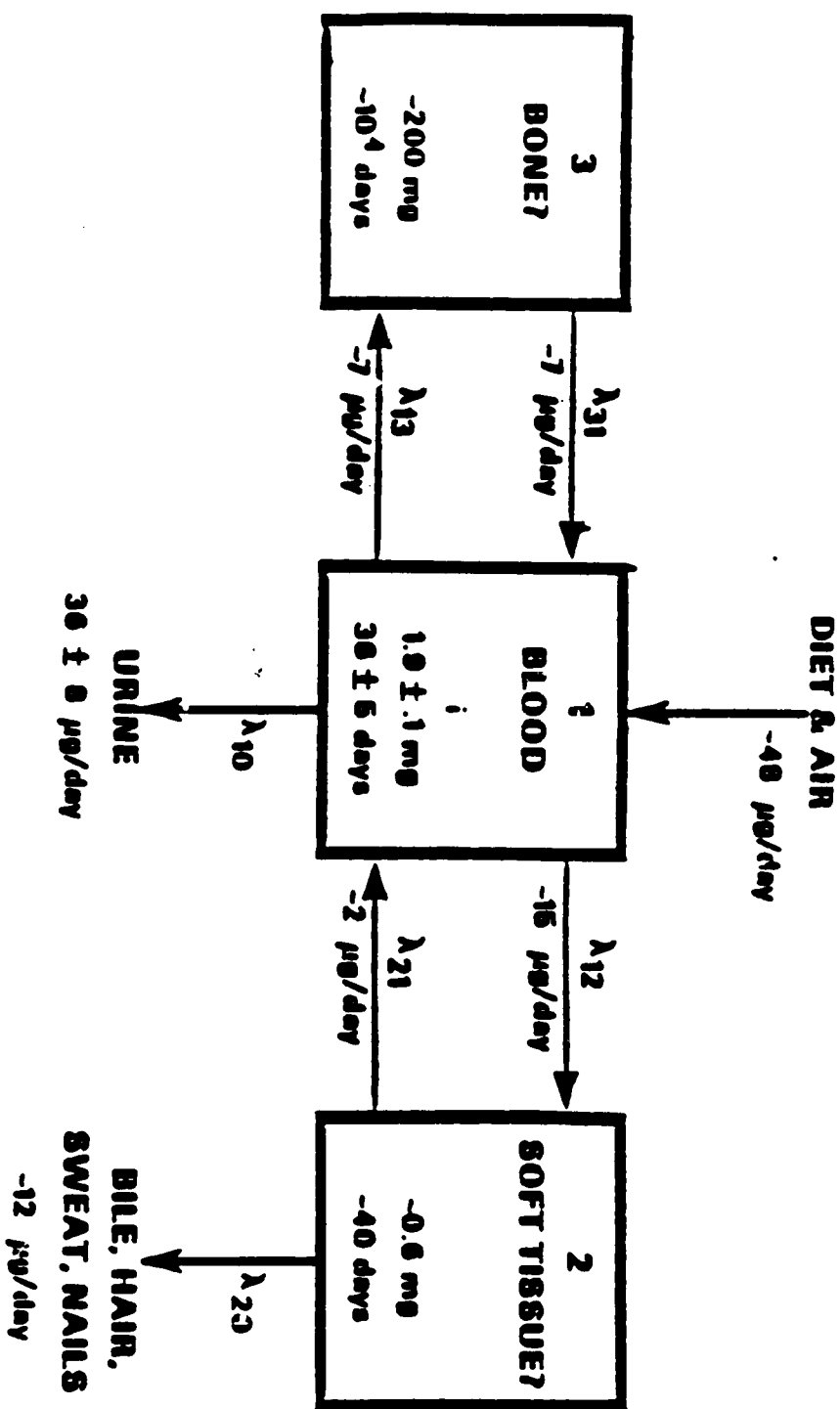


Fig. 1. A three-compartmental model of human lead metabolism

This model is derived from tracer and balance data from five healthy men. The lead content and mean life of each pool and the rates of lead movement between pools (λ) are shown. Numerical values represent the mean values ($\pm$ SD) for all subjects for whom data were available. Loss of lead from the body via pool two (120) is from integumentary structure (hair, nails, sweat) and alimentary tract losses, such as salivary, biliary, gastric, and pancreatic secretions.

Source: Rabinowitz, Wetherill and Kopple (1976).

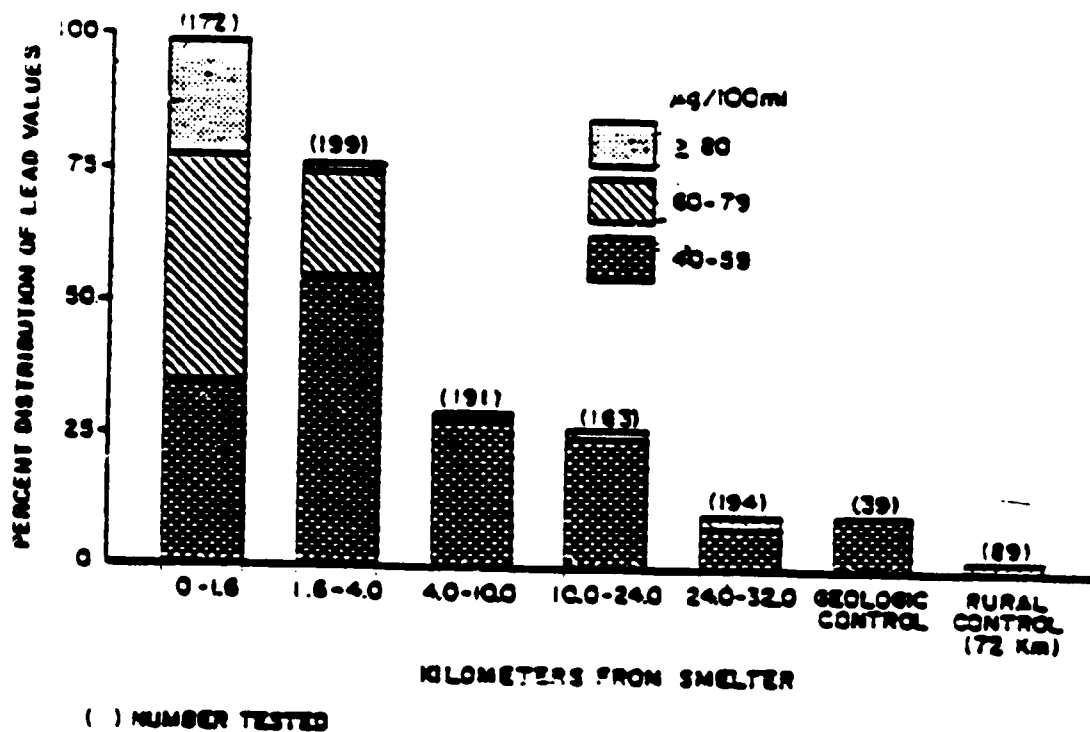


Fig. 2. Distribution of blood lead levels in one-to nine-year-old children, near north Idaho smelter, August 1974.

Source: Landrigan, Baker, Feldman, et al. (1976).

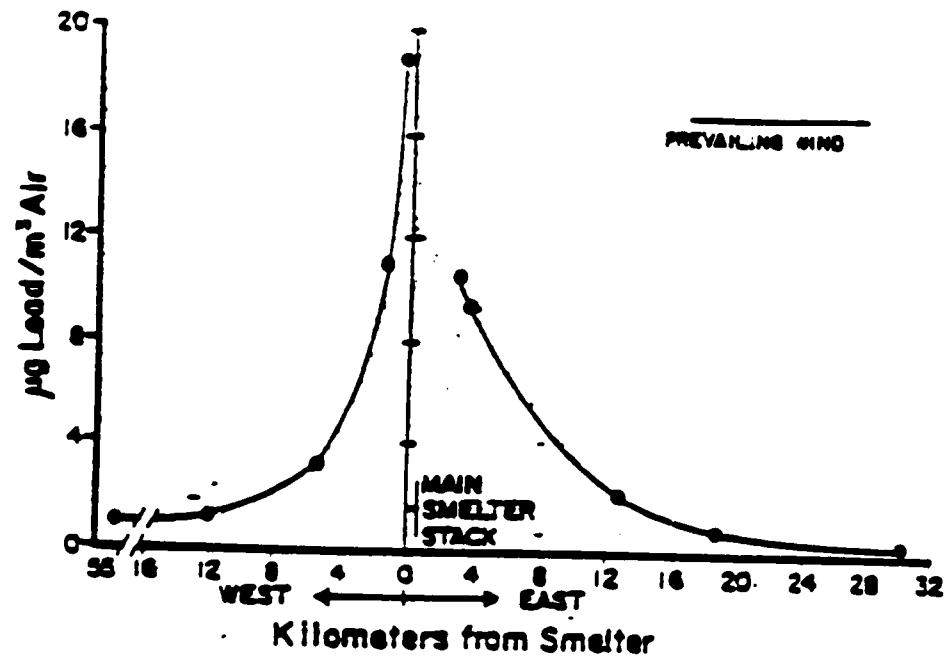


Fig. 3. Airborne lead concentrations (suspended particulates) by distance and direction from smelter, Shoshone County, Idaho, July 1974 (monthly average).

Source: Landrigan, Baker, Feldman, et al. (1976).

Chapter 14

Relation of Surrogate Measures
to Measures of Exposure

Walter J. Rogan

RELATION OF SURROGATE MEASURES TO ACTUAL MEASURES OF EXPOSURE

Walter J. Rogan

Measurements of exposure on individuals can range from anamnestic information gathered by questionnaire to detailed biochemical measurement at the toxic agent's site of action. Leaving issues of individual susceptibility aside, the clearest understanding of an agent's action emerges when we know its concentration at the site where the damage is done, such as DNA in a cell in the bronchial mucosa, limb buds in the developing fetus, or the sperm or ovum in the gonad. On its way to its site, the agent may have been metabolized to a more or less toxic form. The biologically active dose can thus differ from the administered dose, for reasons of absorption, metabolism, or excretion, among others, and recalled dose may differ from administered dose. In a strict sense, any exposure information other than the biologically active dose is a surrogate. What is usually available for studies of the effects of toxic agents is exposure information that bears some relationship, often unknown, to the biologically active dose. Sometimes several different kinds of exposure information are available.

Epidemiologists deal with these problems under the general rubric of misclassification, since it is desirable to classify the exposed according to their true dose; however, the information for doing so is not available. As a rule, the assumption is that the estimate of exposure bears some relationship to the dose at the target site, and that any bias introduced into the measurement of relative risk is a bias toward the null if the misclassification is independent of outcome. The better the measure, the less the bias. There is also a loss of power to declare a risk different from the null. This loss

of power is roughly proportional to the correspondence between the estimate of exposure and the actual dose. The general statistical behavior of these ideas in the framework of regression analysis has recently been discussed by Kupper (1984).

While the misclassification issue is straightforward in theory, in practice it can be unexpectedly complex. The examples will be presented here in the framework of surrogate information leading to misclassification, but each case differs somewhat in how those ideas apply. Several examples will be considered in which two or more measures of exposure are or might be available, one a more meaningful estimate of exposure, and one a surrogate. In the first example, polybrominated biphenyls, the "true" measure of exposure is the concentration of the chemical in fat tissue, which is thought to represent a good estimate of integrated exposure and body burden; the surrogate measure is the level of the same chemical analyzed in blood. Other examples are aflatoxin in liver versus aflatoxin in blood or urine and lead in calcified tissue. The use of urinary mutagens or other biochemical measures will be discussed in the context of a complex exposure, that of pharmaceutical workers. The specific case of surrogate respondents to interviews will not be discussed here; that situation has a literature of its own.

PBB In Michigan

In 1973, there was a food contamination in Michigan which was eventually shown to involve the entire lower peninsula. The contaminant chemical was a mixture of polybrominated biphenyls (PBBs), which are fat soluble, fire retardant chemicals that had been marketed under the name of Firemaster. The Firemaster had been mixed up in shipment with Nutrimaster, a MgO₂ supplement for dairy cows. Livestock on affected farms sickened, but the cause of their

illness was not discovered until many had been slaughtered and eaten. PBB was not metabolized substantially by the food animals nor by the people who consumed them. Although there was not acute toxicity among the exposed, there was suspicion among health officials that delayed toxicity might occur, since the chemical was toxic to nerve, lymphocytes, and liver in the laboratory, and was also a carcinogen. Thus plans were made for follow up surveillance of those with what were thought to be the highest exposures; namely the farmers themselves, those who had bought and consumed contaminated meat and dairy products, chemical company workers, and those who lived closest to the contaminated farms.

There are three ways one might estimate people's exposures to the chemical: ask them, measure it in something readily available, like blood or urine, or measure it in something hard to get, like fat tissue or liver. The pharmacology of the chemical would predict that the majority of the ingested chemical would be stored in fat, and that the level found there would be a good measure of cumulative dose, and thus of risk of disease. The amount of chemical in fat should be in rough equilibrium with the fat in serum, in rough proportion to the fat concentration. Adipose tissue is about 65 percent fat, and blood about 1 percent. Thus the blood concentration should be about 1.5 percent of the fat concentration. As it happens, it is much easier to measure PBB precisely at concentrations in the 1-2 part per million range at which they are found in fat than it is to measure them at the 1-12 part per billion range at which they are found in serum. PBBs do not appear in urine and have a complex appearance curve in liver, so fat and blood were really the only candidates for analysis for subjects other than nursing mothers, who are able to give conveniently what amounts to a fat sample.

Thus analysis of fat tissue obtained by fat biopsy would be the best approximation of total dose, and blood analysis could be viewed as a surrogate for fat analysis. The question would then be, what price in terms of bias and power had to be paid for the use of the more convenient blood analyses?

Data are available to look at this question (Gladen and Rogan, 1979). The Michigan Health Department collected 198 pairs of blood and fat samples and analyzed them for PBB. One way to look at these data was to divide the sample into three exposure categories on the basis of what the blood analysis showed, construct three ordered categories of fat values which corresponded as closely as possible to the blood values, assign some plausible relative risks to the fat categories, and look at what would be observed in a study in which blood values were used to assign exposure status. For small relative risks, the bias and power loss were modest, but the misclassification had greater effects as the relative risks got larger (Table 1).

As it developed, it was not feasible to do fat sampling on large numbers of people in the context of the follow-up study, although a large field survey done by Selikoff and Anderson (1979) did include them. As yet, significant long term effects such as changes in cancer mortality have not emerged in the PBB studies, so the validity of this line of reasoning has not been checked. There has been one study that looked at the relationship between fat level of PBB and performance on a psychological test of school readiness; it showed modest decrements in performance among the children with above average PBB concentrations (Seagull 1983) (Figure 1).

Finally, an interesting post hoc question that was not examined is: given that one has information on the relationship among blood, fat, and some historical marker of exposure, such as living on or consuming goods from a

quarantined farm, might the history be as good as the lab test at assigning people into exposure categories? A similar line of reasoning was followed by Levine, Blunden, DalCorso et al. (1983), who judged reproductive histories to be more sensitive than sperm analyses in detecting chemically induced male infertility.

Aflatoxin in liver and blood

There had been suspicion at least since the early seventies that Reye Syndrome, an often fatal disorder of the brain and liver in children, might have an environmental etiology. A leading candidate among potential agents was aflatoxin, a naturally occurring hepatotoxin and carcinogen produced by the mold *Aspergillus flavus*. Outbreaks of acute aflatoxicosis, an illness similar to Reye Syndrome, had been reported from Thailand and Czechoslovakia among children who had consumed heavily contaminated rice or corn. There was clearly some contamination of food in the US, particularly in the South. The capability to analyze for aflatoxin existed at most universities and state health departments, yet until 1978, there was only one case reported in the US of aflatoxin in a child with Reye Syndrome (Chaves, Carballo, Ellefson and Gomez, 1976) (Table 2).

In 1979, Ryan, Hogan, Hayes et al from Mississippi reported aflatoxin B1 in the livers of six out of seven children dying of Reye Syndrome compared with one of twelve controls (Table 3). This was an extraordinarily strong association (estimated odds ratio = 66). In another study done at the same time but reported later, the Centers for Disease Control had found about a 21 percent prevalence of detectable aflatoxin in urine from affected children, their families, and controls (Table 4). Thus they had evidence of substantial population exposure, but no evidence of association with Reye Syndrome (Nelson, Kimbrough, Landrigan et al, 1980).

quarantined farm, might the history be as good as the lab test at assigning people into exposure categories? A similar line of reasoning was followed by Levine, Blunden, DalCorso et al. (1983), who judged reproductive histories to be more sensitive than sperm analyses in detecting chemically induced male infertility.

Aflatoxin in liver and blood

There had been suspicion at least since the early seventies that Reye Syndrome, an often fatal disorder of the brain and liver in children, might have an environmental etiology. A leading candidate among potential agents was aflatoxin, a naturally occurring hepatotoxin and carcinogen produced by the mold *Aspergillus flavus*. Outbreaks of acute aflatoxicosis, an illness similar to Reye Syndrome, had been reported from Thailand and Czechoslovakia among children who had consumed heavily contaminated rice or corn. There was clearly some contamination of food in the US, particularly in the South. The capability to analyze for aflatoxin existed at most universities and state health departments, yet until 1978, there was only one case reported in the US of aflatoxin in a child with Reye Syndrome (Chaves, Carballo, Ellefson and Gomez, 1976) (Table 2).

In 1979, Ryan, Hogan, Hayes et al from Mississippi reported aflatoxin B1 in the livers of six out of seven children dying of Reye Syndrome compared with one of twelve controls (Table 3). This was an extraordinarily strong association (estimated odds ratio = 66). In another study done at the same time but reported later, the Centers for Disease Control had found about a 21 percent prevalence of detectable aflatoxin in urine from affected children, their families, and controls (Table 4). Thus they had evidence of substantial population exposure, but no evidence of association with Reye Syndrome (Nelson, Kimbrough, Lanrigan et al, 1980).

Initially, it was felt that this was a clear example of the phenomenon just noted, that is, that the measurement that was closest to a measure of the biologically active dose was the liver analysis, and study of urine would be biased toward finding no association. Because of the large and unexpected association found in the liver study, it appeared that a small confirmatory study was indicated. The study implemented was a simple one. A letter was sent to the pathology departments and coroners' offices in the Southeast in which they were asked to submit autopsy specimens of liver tissue from children who had died of Reye Syndrome. They were also asked to submit a control liver sample from an age and sex matched child who had died suddenly. There was concern that the children who were hospitalized for long periods would be eating differently from the way they would be if they were at home. Fatal Reye Syndrome is often a catastrophic illness with cases dying quickly; unless controls were also recently admitted before they died, any aflatoxin residues might have had a chance to be metabolized, therefore finding a difference might represent differential opportunities for exposure rather than causality.

Twelve cases were collected and only one control, none had detectable aflatoxin B1 (Rogan, Yang and Kimbrough, 1985) (Table 5). One case did have a trace of the M1 metabolite. The analytical methods were at least as sensitive as those used in the positive study. Although not a great deal was known about aflatoxin pharmacokinetics in humans, the predicted residue would not be the B1 form.

This is somewhat of an interpretive dilemma. There have been no confirmatory studies published, but there is no obvious fault in the original paper to which the apparently aberrant result can be attributed. One further study that could be done would be to look for adducted aflatoxin. Although the

antibody in this study had some cross reactivity with the plausible aflatoxin adducts, an exhaustive search was not done.

Lead in tooth and blood

Lead is a potent neurotoxin for mammals, and lead encephalopathy in childhood has been associated since the 1920s with major nervous system sequelae. For almost as long, there has been suspicion that doses lower than those needed to produce encephalopathy might produce disorders of intellect or behavior (Needleman and Landrigan, 1981). The studies that have been done produced mixed results, which were at least partly attributable to the difficulty of assigning an appropriate exposure value to study subjects. The epidemiology of lead exposure is a combination of the distribution of lead sources (classically, peeling paint in deteriorated inner city housing), and childhood habits, with two year olds exhibiting the most hand-to-mouth behavior and thus showing the highest blood lead levels. Lowered IQ and disorders of behavior become much more apparent among school age children but assessing their exposure with their current blood lead might be misleading, since blood leads are most affected by recent exposure. The crucial peak exposure might have taken place years before.

Lead is a di- or tri-valent cation that is metabolized similarly to calcium. Most of the lead that cannot be excreted is stored in calcium rich sites, such as tooth and bone. Even though the lead in such sites is sequestered away from its site of toxic action, the amount stored represents an integration of long term exposure, and might be a better indicator of the total active dose.

Needleman, Gunnoe, Leviton et al (1979) exploited this in a study of the relationship between lead exposure, as measured by tooth lead, and IQ among school age children who had not had encephalopathy. They collected teeth and

analyzed the dentine layers for lead from 2335 first and second grade children. They then recontacted the children in the highest 10th percentile and lowest 10th percentile of lead levels, and found that children who had higher dentine leads scored on the average 4 IQ points below children who had low dentine leads (Figure 2) (Needleman, Leviton and Bellinger, 1982). Although not reported, such a relationship would presumably not be present if blood leads had been used.

Although dentine lead might be the best indicator of integrated long term exposure, it should be clear that it is not the best screening method to detect children at high risk of lead poisoning. Shed teeth are not available from two year olds. Of interest is the fact that most screening programs do not use blood leads, but rather screen with erythrocyte protoporphyrin, a heme precursor that accumulates as a consequence of the inhibition by lead of the enzyme ferrochelatase. Although EP is neither the proximate toxin nor a real measure of integrated dose, it is less susceptible to short term variation and to contamination than is blood lead. Thus the choice of an appropriate surrogate measure depends on the use to which it will be put.

Mutagens in urine

The Ames test uses strains of mutant bacteria with a relatively unstable mutation that makes them require histidine for growth. When these bacteria are plated on histidine free medium with a test substance, visible colonies of the bacteria will grow only if the test substance causes back mutation to histidine independence. Since Ames, Yamasaki and Ames (1977) (Figure 3), reported almost a decade ago that smokers have mutagenic urine, hundreds of substances have been tested directly in the assay, and at least a dozen human studies have been reported of mutagenic urine among various occupational groups, cancer

chemotherapy patients, passive smokers, and patients with cirrhosis or bilharzial bladder cancer (Everson, Ratcliffe, Flack et al, 1985).

The idea presented is a simple but elegant one, particularly applicable to complex exposure situations. The presence of a mutagen in urine implies exposure to and absorption of, some active dose of a toxin of general interest, without having to specify exactly what that toxin is. In theory, among a group of subjects with similar exposures, those who have mutagenic urine should be at higher risk for mutation related disease, such as cancer.

A partial example can be drawn from work done with a study of chemotherapy workers (Everson, Ratcliffe, Flack et al, 1985). Many cancer chemotherapeutic agents are themselves mutagens and carcinogens. There had been several reports that nurses administering or pharmacists preparing such drugs had mutagenic urine. The exposures were complex, with workers commonly exposed to eight or more drugs. The advantages were the simplicity of the assay and the fact that an analytic technique would have to be able to give data on eight or more different substances. As it turned out, the workers studied were already employing adequate protective measures so that no mutagenic substances were detectable in their urine.

Mutagens in urine are not strictly an exposure measure that is closer to the target organ of interest; in a sense, they are both an outcome of exposure and a kind of biological dosimeter. They serve to identify those study subjects who have in some sense a greater exposure, and thus could, under the right circumstances, increase power and reduce bias. The price to be paid for this power is that a positive Ames test in urine has very low specificity; in particular, low levels of urinary mutagens might well result from relatively common exposures other than the one under study, and thus the possibility of confounding is quite high.

There are now available several other laboratory techniques that might also serve as dosimeters. The most familiar are cytogenetic changes, which are known to occur as a consequence of radiation exposure, and which more recently have been shown to be induced by chemical carcinogens. Sister chromatid exchanges have also been used in the same way. Certain drugs and environmental chemicals, like polychlorinated biphenyls and the polycyclic aromatics in tobacco smoke, are inducers of mixed function oxidase enzymes, whose activity can be measured in tissue. A dramatic example of the latter was seen in some work done with enzyme activity in placentae of smokers and non-smokers, in which a 400g birthweight decrement was seen in the offspring of those smokers whose enzymes were induced; smokers with normal enzyme activity had normal birthweight children (Pelkonen, Karki, Koivisto et al, 1979). Although measurement of enzyme induction is still a research procedure, there are non invasive screening methods under development (Kotake, Schoeller, Lambert et al, 1982). This brings us to adducts, which are right at the line between outcome and exposure, and are discussed elsewhere in the literature.

Discussion

The effects on power and estimation of surrogate dose or exposure information are fairly familiar, and it appears that most epidemiologists will find nothing that strains credulity in the examples used in this chapter. The use of the laboratory in environmental epidemiology is increasing, and many of the techniques that are becoming available pack a great deal of biological firepower. Breaks in chromosomes or mutagens in urine appear to be undesirable conditions or diseases themselves, rather than indicators of exposure. Because of the plausibility of their association with diseases like cancer or inherited disorders, the temptation to skip a step and regard the laboratory measure as an illness in itself, or as identical with the occurrence of a cancer or a

genetic disease is quite strong. It should be kept in mind that actual validation studies, in which persons with and without the laboratory abnormality have been followed to see if they indeed develop disease more frequently, are all but unavailable. A nice example of this is the fact that although persons exposed to the atomic bomb are known both to develop cytogenetic abnormalities and leukemia, it has not been shown that those with chromosome breaks are more likely to develop leukemia, conditional on dose. A great deal of work is necessary before these most powerful of surrogate measures can be substituted for diseases in etiologic studies.

REFERENCES

- Chaves-Carballo, E., Ellefson, R.D. and Gomez, M.R. (1976) An aflatoxin in the liver of a patient with Reye-Johnson Syndrome. *Mayo Clin. Proc.* 51:48-50.
- Everson, R.B., Ratcliffe, J.M., Flack, P.M. and Watanabe, A.S. (1985) Detection of low levels of mutagen excretion not associated with exposure to cytotoxic drugs among chemotherapy workers at two hospitals. *Cancer Res.* 45:6487-97.
- Gladen, B.C. and Rogan, W.J. (1979) Misclassification and the design of environmental studies. *Amer. J. Epidemiol.* 109:607-16.
- Kotake, A.N., Schoeller, D.A., Lambert, G.H., Baker, A.L., Schaffer, D.D. and Josephs H (1982) The caffeine CO₂ breath test: Dose response and route of N-demethylation in smokers and nonsmokers. *Clin. Pharm. Therap.* 32:261-69.
- Kupper, L.L. (1984) Effects of the use of unreliable surrogate variables on the validity of epidemiologic research studies. *Amer. J. Epidemiol.* 120:643-48.
- Levine, R.J., Blunden, P.B., DalCorso, R.D., Starr, T.B. and Ross, C.E. (1983) Superiority of reproductive histories to sperm counts in detecting infertility at a dibromochloropropane plant. *J. Occup. Med.* 25:591-97.
- Needleman, H.L., Gunnoe, C., Leviton, A., Reed, R., Peresie, H., Maher, C. and Barrett, P. (1979). Deficits in psychological and classroom performance of children with elevated dentine lead levels. *N. Engl. J. Med.* 300:689-95.
- Needleman, H.L. and Landrigan, P.J. (1981) The health effects of low level exposure to lead. *Ann. Rev. Public Health* 2:277-98.

- Needleman, H.L., Leviton, A. and Bellinger, D. (1982) Lead-associated intellectual deficit. N. Engl. J. Med. 306:367.
- Nelson, N.J., Kimbrough, R., Landrigan, P., Hayes, A.W., Yan, G.C. and Benanides, J. (1980) Aflatoxin and Reye's syndrome: A case control study. Pediatrics 66:865-69.
- Pelkonen, O., Karki, N.T., Koivisto, M., Tuimala, R. and Kauppila, A. (1979) Maternal cigarette smoking, placental aryl hydrocarbon hydroxylase and neonatal size. Toxicol Lett 3:331-35.
- Rogan, W.J., Yang, G.C. and Kimbrough, R.D. (1985) Reye syndrome and aflatoxin: A study of livers from deceased cases. Arch. Environ. Hlth. 40:91-5.
- Ryan NJ, Hogan GR, Hayes AW, Unger PD and Siraj MY: (1979) Aflatoxin B₁: Its role in the etiology of Reye syndrome. Pediatrics 64:71-5.
- Seagull, E.A.W. (1983) Developmental abilities of children exposed to polybrominated biphenyls. Am. J. Public Health 73:281-85.
- Selikoff, I.J. and Anderson, H.A. (1979) A survey of the general population of Michigan for health effects of polybrominatedbiphenyl exposure. Report to the Michigan Department of Public Health. September 30. Environmental Sciences Laboratory, Mount Sinai School of Medicine of the City University of New York.
- Yamasaki, E. and Ames, B.N. (1977) Concentration of mutagens from urine by absorption with the non-polar resin XAD-2: Smokers have mutagenic urine. Proc. Natl. Acad. Sci. 7:3555-59.

TABLE 1: Blood versus fat levels of PBB (polybrominated biphenyls) Michigan, U.S.A. 1973.

Fat levels (ppb) ^a	Blood levels (ppb) ^a			Total
	Low (nd) ^b	Medium (<3)	High (>4)	
Low (0-300)	73	18		91
Medium (301-1000)	20	32	6	58
High (1000+)	3	3	43	49
Total	96	53	49	198

Source: Gladen and Rogan (1979)

^appb = parts per billion

^bnd = not detectable

TABLE 2: Reports through 1974 of liver aflatoxins in patients with Reye Syndrome

Authors	Year	Country	Age	Sex	Aflatoxin B ₁ ug/kg liver
Shank et al	1971	Thailand	2 yr. 13 yr.	Male	93
Becroft and Webster	1972	New Zealand	22 mo. 8 mo.	Male Female	5-50 5-50
Dvorackova et al.	1972	Czechoslovakia	1 yr. 22 mo.	Female Female	ND ^a ND
Chaves-Carballo et al	1974	United States	15 yr.	Female	22.5

Source: Chaves-Carballo, Ellefson and Gomez (1976)

^a ND = no data

TABLE 3: Relevant laboratory tests for Reye Syndrome cases, Mississippi USA
1975-1978

Laboratory tests	Cases							
	<u>R.W.</u>	<u>M.F.</u>	<u>J.T.</u>	<u>J.P.</u>	<u>N.W.</u>	<u>D.W.</u>	<u>K.M.</u>	<u>J.T.</u>
SGOT (IU/I)	370	336	153	585	580	396	>1000	169
Blood ammonia (ug/dl)	228	707	ND ^a	229	1600	2000	133	471
Blood urea nitrogen (mg/dl)	46	18	12	76	27	40	47	26
Blood glucose (mg/dl)	82	24	211	141	28	36	10	26
Liver aflatoxin B ¹ (ng/gm)	12.64	17.33	3.36	ND	-	16.13	3.41	2.2

Source: Ryan, Hagan, Hayes et al (1979)

^aND = no data

TABLE 4: Reye Syndrome patients and controls with detectable aflatoxin B₁ levels in serum and or urine.

Subject categories	No.	Percent of subjects with detectable aflatoxins
Rye Syndrome patients	17	11.8%
Patients' family members	43	27.9%
Neighborhood controls	26	15.4%
Hospital controls	22	22.7%
Total	108	21.3%

Source: Nelson, Kimbrough, Landrigan et al (1980)

TABLE 5 Relevant laboratory tests for Reye Syndrome cases and one control
Southeastern USA 1983

Patient information and laboratory tests	Cases and Controls											
	A	B	C	D	E	F	G	H	I	J	K	L
Age	3 yrs	7 yrs	4 yrs	1 yr 2 mos	6 yrs	9 yrs	10 yrs	3 yrs	6 yrs	4 yrs	4 yrs 6 mos	12 yrs
Sex	F	F	F	F	F	F	M	F	F	F	F	M
Serum ALT (IU/l)	1700	1630	-	liver function	3080	459	1090	432	-	1400	-	4600
Serum AST (IU/l)	1557	1480	-	tests elevated	-	-	-	128	-	188	-	2190
Serum ammonia (ug/dl)	686	816	-	normal	36	172	122	820	-	150	110	616
Blond urea nitrogen (mg/dl)	221	-	-	41	-	-	-	-	-	52	-	-
Serum glucose (mg/dl)	-	-	-	-	-	-	-	24	-	-	39	-
Liver aflatoxin B ₁ or M ₁ (ppb)	ND ^a	ND	ND	0.04 ^b	ND	ND	ND	ND	ND	ND	ND	ND

Source: Rogan, Yang and Kimbrough (1985)

^a ND non detected

^b detected, not confirmed

Mean standardized scores on McCarthy scales of children's abilities
by PBB exposure level
Michigan U.S. 1980

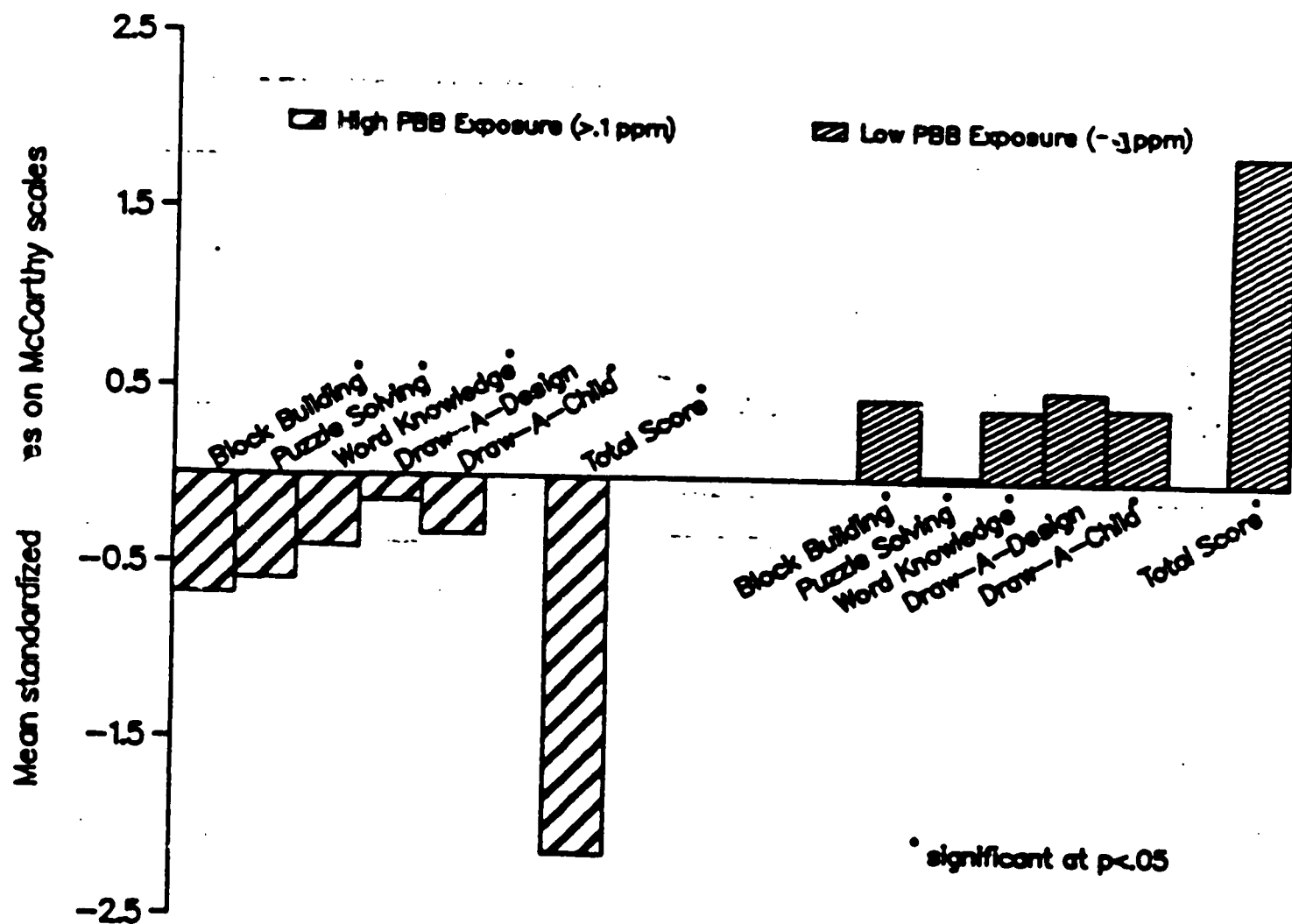


Fig 1.

Source: Seagull (1983)

Rogan

✓

Cumulative frequency distribution of verbal IQ
scores in subjects with low or high levels of lead

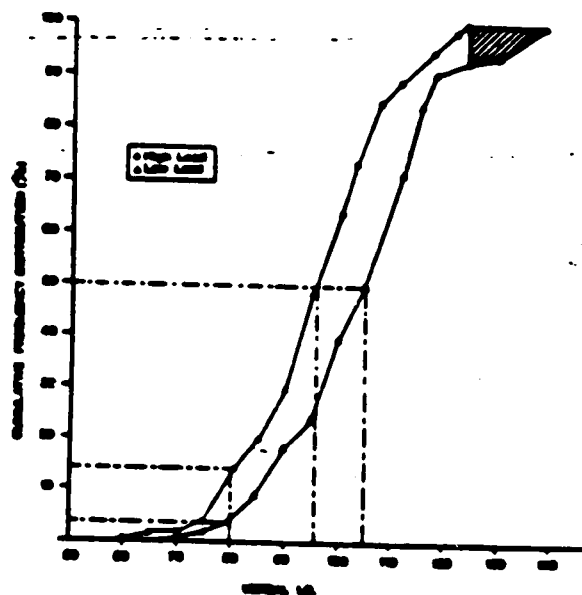


Fig 2

Source: Needleman, Leviton and Bellinger (1982)

Rogan

-20-

Chapter 15

Molecular Cancer Epidemiology: The Use of New Laboratory Methods in Studies on Human Cancer Causation

I. Bernard Weinstein

MOLECULAR CANCER EPIDEMIOLOGY: THE USE
OF NEW LABORATORY METHODS IN STUDIES
ON HUMAN CANCER CAUSATION

I. Bernard Weinstein

About one hundred years ago Robert Koch identified, through a combination of ingenious clinical, pathologic and laboratory studies, a bacterial agent as the cause of tuberculosis (1). This discovery launched the field of infectious disease epidemiology. It is rather paradoxical that epidemiologic studies on cancer have, in general, followed quite a different methodology. I believe that it is time that cancer epidemiologists adopt a more interdisciplinary approach, similar to that used in the field of the infectious diseases. For example, until recently it was rare to find a cancer epidemiologist who would go out into the field and bring along his laboratory colleagues, whereas in infectious disease epidemiology this team approach is commonplace. A few years ago my colleagues and I began to emphasize a new approach, which we called "molecular cancer epidemiology", in part because at the present time research in biology is not considered exciting unless it is prefaced by the word molecular. Actually, we were sincerely impressed with the possibility of developing a number of biochemical and molecular tools that epidemiologists could use to better define the etiology of specific human cancers. A variety of highly sensitive and specific laboratory procedures are now available for this purpose and others can be developed. They can be grouped into four categories in terms of serving as biomarkers of: (1) genetic or acquired host susceptibility; (2) metabolism and tissue levels of carcinogens; (3) levels of covalent adducts formed between carcinogens

and cellular macromolecules; and (4) markers of early cellular responses to carcinogen exposure; for example, sister chromatid exchange, DNA repair, altered gene expression, and mutation. The subject of molecular epidemiology has been reviewed in detail elsewhere (2,3). This paper will largely emphasize markers related to category (3) and their potential use in epidemiologic studies.

It is important to clarify certain terms (2, 3). "Exposure" deals with ambient levels of potential toxins in various environmental sources, air, water, food, etc., whereas "internal dose" is the amount of the parent compound that actually enters the body. What is particularly important in the field of chemical carcinogenesis is to distinguish the "internal dose" from the "biologically effective dose", i.e., the dose of the active form of the toxic agent at the level of the target cells. It is also important to identify "early markers of response" to the toxic agent, i.e., markers that can be detected long in advance of disease activity itself. At the present time, it is difficult to quantitate "background" levels of hazardous chemicals in our environment, or to know the background levels of such chemicals within the human body. In the field of chemical carcinogenesis one cannot simply measure background with a Geiger counter. Thus, we must develop new methods to assess at least three parameters related to chemical carcinogenesis: biologically effective doses, early markers of response and background levels. Indeed, many of the inconclusive epidemiologic studies in this field and the vexing problems of risk extrapolation reflect the lack of precise information on these parameters.

A fundamental biochemical principle related to many chemical carcinogens is that these chemicals are often not biologically active until they undergo metabolism in the body. (For a detailed review on chemical carcinogenesis see ref. 4). Benzo(a)pyrene (BP), for example is not a highly reactive chemical. It must first undergo oxidative metabolism to form a dilepoxide. Normally, the latter compound will be reduced, and often conjugated to yield an innocuous derivative that is excreted. However, if the dilepoxide derivative of BP accumulates it can react with guanine residues in DNA and thus distort the structure and function of the genetic material. This principle applies to a number of other chemical carcinogens, including aflatoxin, vinyl chloride, and nitrosamines. Thus, within the framework of the above-described terminology we can distinguish: 1) BP exposure level, the level of BP in the air; 2) BP dose, the amount of the substance that enters the body; 3) BP detoxified metabolites, the level of inactive derivatives in the blood and urine; and 4) the biologically effective dose of BP, i.e., the actual level of BP dilepoxide that becomes bound to cellular DNA.

Most of the remainder of this paper will emphasize methods which exploit the principle of carcinogen-DNA adducts, since it is possible to develop highly sensitive methods to measure the extent to which DNA is modified in human tissues by potential environmental carcinogens. In essence, DNA can be used as an endogenous trapping agent, an internal dosimetry badge, for assessing types and levels of exposure to carcinogens and mutagens. The extent to which DNA is modified by such chemicals in individuals in the general population can be considered the "background" level. Levels of adducts that are significantly elevated in

subpopulations might then serve as a quantitative marker of increased disease risk in that subpopulation. There now exist highly specific monoclonal antibodies and highly sensitive immunoassays for quantitating the levels of carcinogen adducts on cellular DNA. A variety of highly sensitive analytical techniques are also available to analyze levels of carcinogens and their metabolites in serum, urine, feces and various tissues obtained from human subjects (For review see 2-5).

It should be emphasized that active derivatives of several chemical carcinogens, for example, the diepoxide derivative of BP, often form stable chemical bonds with not only DNA but also protein. For several carcinogens there is a rather good correlation between extent of binding to DNA and to protein. Since it is often easier to obtain protein samples from subjects, one can use protein-binding as a surrogate for assessing DNA binding. For example, hemoglobin samples are easily obtained from human subjects. A drop of blood will provide more hemoglobin than an equivalent amount of DNA obtained from 50ml of blood. Thus, investigators are now using this approach in molecular epidemiology studies (2, 6, 7).

Examples of biomarkers of early responses to DNA damage include unscheduled DNA synthesis, chromosome aberrations, sister chromatid exchanges and the micronucleus test. These markers can be scored long in advance of preneoplastic lesions that are detected by histologic methods. There is recent evidence that cancer is often associated with abnormal activation of a limited set of cellular genes, termed proto-oncogenes. Oncogene activation can occur by a variety of mechanisms including point mutation, gene amplification and gene translocation (for review of this

subject see 8). It is possible, therefore, that highly sensitive methods might be developed to detect changes in cellular oncogenes at early stages of the carcinogenic process in humans. This approach is relevant to not only cancer epidemiology and risk assessment but also to new approaches to early detection and prevention.

Table 1 lists potential applications of assays for levels of carcinogen-DNA adducts in humans. One use of assays of carcinogen-DNA adducts in humans is to improve the methodology of quantitative risk extrapolation, particularly if we are dealing with a compound to which humans are exposed. Thus, if we know that in experimental animals a certain level of DNA adducts is associated with a certain cancer risk, then, based on levels of DNA adducts detected in a human subpopulation, we might estimate their cancer risk. There are of course, likely to be interspecies variations, but, with experience and an expanding data base, this approach could provide a valuable ancillary method to the difficult field of risk extrapolation. This approach might be particularly valuable in assessing the hazards of low level exposures to certain carcinogens in humans, because DNA adducts can be measured with great sensitivity and the number of cancer cases occurring in the exposed population may be too small to be statistically significant.

A limitation of the DNA adduct approach is that it probably is an indicator of recent exposure and not of exposure in the distant past, because of the process of DNA repair. Actually, there is very little data on how rapidly DNA adducts are excised from human tissues in the intact individual. There is the possibility that because of the long life span

of certain lymphocytes, and the fact that resting lymphocytes have very little DNA repair, that these cells might provide a marker of exposure which occurred in the distant past. It seems likely, however, based on studies in rodents and cell culture systems that most human tissues will excise DNA adducts within weeks or months following exposure to the substance of interest.

Many toxic exposures occur in a complex environment where there is exposure to a mixture of substances rather than to a single toxic agent. Thus, it is often difficult, using conventional epidemiologic methods, to identify the specific chemical agent in a mixture which is most likely to be carcinogenic. If we assume that the latter compound is the one which becomes most extensively bound to the cellular DNA of the exposed individual, then, by using the DNA as a trapping agent, we can identify the specific agent by detecting its presence on cellular DNA. One might, for example, determine which of the many chemicals present in cigarette smoke is most carcinogenic in cigarette smokers by determining the types and abundance of specific DNA adducts in cells obtained from cigarette smokers. Analogous approaches might be used for detecting potential carcinogens in specific work places and in specific diets. There is an extensive data base indicating that cellular levels of DNA adducts often correlate with carcinogen potency and dose (2, 3, 9), although there may exist discrepancies in this correlation because of differential rates of DNA repair and the multistage nature of the carcinogenic process (2-4).

Table 2 lists some of the current methods that can be used to detect carcinogen-DNA adducts. These methods are discussed in detail elsewhere (2-4). At the present time, immunoassays for specific carcinogen-DNA

adducts appear to be the most promising for epidemiologic studies because of their high sensitivity and relative simplicity. One can currently detect about one residue of a carcinogen per about 10^7 nucleotides of DNA. Since human cells contain about 3×10^9 nucleotides of DNA per cell, this is equivalent to about 300 carcinogen residues per cell. If the sensitivity of the current method can be increased 300 fold, which seems possible, then a negative result would indicate that there is less than one carcinogen residue per cell. In animal experiments, when carcinogens produce tumors in about 50% of the exposed animals, the levels of DNA adduct formation are in the range of about 1 per 10^5 to 1 per 10^6 nucleotides of DNA. Thus, the current methodology and that now being developed may be in the appropriate range of sensitivity for epidemiologic studies. A number of polyclonal and monoclonal antibodies are now available to detect specific carcinogen DNA adducts, and others are being developed in various laboratories. Published studies indicate that it is possible to use immunoassays to detect BP-DNA adducts in human tissues of cigarette smokers and coke oven workers (10, 11). The p^{32} post-labeling method of Randerath (12) also appears very promising for epidemiologic studies, since it is also highly sensitive and, in contrast to the immunoassays, does not require the preparation of an antibody specific to each compound that is in question. This method has been used to detect DNA adducts in the placental DNA of pregnant cigarette smokers (13).

A recent study on esophageal cancer in the Peoples Republic of China (14) provides an exciting example of the use of the immunoassay approach

to identify DNA adducts in a high risk human population. Antibodies have also been developed to DNA modified by the cancer chemotherapeutic agent cis-diamminodichloroplatinum (15). Since patients receive a precisely known dose of this agent, and since the treatment is associated with increased risk of developing leukemia, these subjects provide an ideal population for pilot studies on DNA adducts and risk extrapolation in humans. Indeed, recent studies (15) indicate that there is a very high level of drug-DNA adducts in patients treated with cis-diamminodichloroplatinum i.e., levels of adducts comparable to those seen in rodents administered high doses of potent carcinogens.

It is also possible to detect carcinogen-DNA adducts in single cells using immunofluorescence methods (2, 3). This should permit the use of cell cytology methods in epidemiology studies. Thus far, however, these methods are less sensitive than assays on DNA extracted from tissue samples or blood lymphocytes. New image intensification microscopy methods and other approaches should increase the sensitivity of the cytologic methods.

Carcinogenesis is obviously a multistep process and, in contrast to initiating carcinogens, tumor promoters do not appear to act by binding to cellular DNA (4, 16). Thus, it will be necessary to develop other laboratory methods that can be used in epidemiologic studies to detect the exposure of humans to tumor promoters. Recent studies indicate that the phorbol ester tumor promoters act by altering the phosphorylation of specific cellular proteins (17). If this proves to be a general principle, then highly sensitive and specific methods might be developed to assess the profiles of phosphorylated proteins in biopsy specimens of

human tissues, as an approach to monitoring high risk groups for possible exposure to tumor promoters. The role of tumor promoters in human cancer causation and problems of risk extrapolation for tumor promoters are discussed in detail elsewhere (16).

Finally, I want to emphasize that it is very likely that several major human cancers are due to combined exposure to chemical carcinogens and viral agents(4). Therefore, as epidemiologists go out into the field, they should bring along not only their colleagues in genetic toxicology but also those with expertise in virology. Recent studies suggest that human liver cancer in Central Africa and in Asia is due to combined exposure to Hepatitis-B virus and aflatoxin (18). Thus, it is important to be able to score for human exposure to viruses as well as chemical carcinogens in future studies on human cancer causation. This principle might be particularly important in future studies on human breast and colon cancer.

In summary, there are three fields which have, until recently, generally worked in isolation of each other. Cancer epidemiology (which did not follow the historical course of infectious disease epidemiology, in terms of methodology), experimental studies on carcinogenesis, and clinical studies. I believe that the concept of molecular epidemiology provides an exciting opportunity to bridge these three fields ^{together} through the use of powerful new biochemical and molecular methods. Hopefully, these combined approaches will make it possible to more precisely identify the environmental causes of the major forms of human cancer and thus lead to more effective approaches to cancer prevention.

TABLE I

ASSAYS FOR CARCINOGEN-DNA ADDUCTS: POTENTIAL APPLICATIONS
IN HUMAN STUDIES

1. Demonstrate and quantitate recent exposure to potential environmental chemical carcinogens.
2. Quantitate the "biologically effective dose" of carcinogens to which humans are exposed.
3. Use as one of a series of parameters in quantitative risk assessment (the parallelogram approach).
4. Determine background levels of chemical carcinogens.
5. Determine the most hazardous carcinogen under conditions of exposure to complex mixtures of chemicals.
6. Use in studies on mechanisms of carcinogenesis - multistage aspects, etc.

TABLE II

METHODS FOR DETECTING CARCINOGEN - DNA ADDUCTS

1. Use of radioactively - labeled carcinogen (useful in rodents or cell culture systems but not applicable in humans).
2. Fluorescent^{cc} detection methods (if the carcinogen is fluorescent).
3. Immunoassays using polyclonal or monoclonal antibodies and radioactive immunoassays (RIA) or enzyme-limited immunoassays (ELISA).
4. Immunofluorescence assays on single cells.
5. Gas chromatography and mass spectrometry.

REFERENCES

1. Koch, R. Ueber bakteriologische forschung. In: Vern. X. International Medical Congress, Berlin, 1890, p. 35, 1892.
2. Perera, F. P., Weinstein, I. B. Molecular epidemiology and carcinogen-DNA adduct detection: New approaches to studies of human cancer causation. J. Chronic. Dis. 35:581-600, 1982.
3. Perera, F. P., Santella, R., Poirier, M. C. Potential methods to monitor human populations exposed to carcinogens: Carcinogen-DNA binding as an example. In: Hoel, D., Merrill, R., and Perera, F.P., (Eds.), Risk Quantitation and Regulatory Policy. Banbury Report 19: Cold Spring Harbor, NY: Cold Spring Harbor Labs., pp. 211-223, 1985.
4. Weinstein, I. B. Basic aspects of environmental carcinogenesis. In: Medical Oncology, Calabresi, P., Schein, P., and Rosenberg, S. A. (Eds.), MacMillan Publishing Co., New York, pp. 63-81, 1975.
5. Weinstein, I. B. The monitoring of DNA adducts as an approach to carcinogen detection. Annual Reviews of Public Health, Nelson, N. (Ed.), 4, pp. 409-413, 1983.
6. Calleman C. J., Ehrenberg, L., Jansson, E. et al. Monitoring and risk assessment by means of alkyl groups in hemoglobin in persons occupationally exposed to ethylene oxide. J. Environ. Path. & Tox. 2:427-442, 1978.

7. Santella, R., unpublished studies.
8. Hunter, T. The Proteins of Oncogenes. Scientific American. 50: 70-79, 1984.
9. Lutz, W. K. In vivo covalent binding of organic chemicals to DNA as a quantitative indicator in the process of chemical carcinogenesis. Mutat. Res. 65:289-356, 1979.
10. Perera, R. P., Poirier, M. C., Yuspa, S. H., et al. A pilot project in molecular cancer epidemiology: Determination of benzo(a)pyrene-DNA adducts in animal and human tissues by immunoassay. Carcinogenesis 3:1405-1410, 1982.
11. Harris, C. C., Vahakangas, K., Newman, M. J., et al. Detection of benzo(a)pyrene diol epoxide-DNA adducts in peripheral blood lymphocytes and antibodies to the adducts in serum from coke oven workers. Proc. Natl. Acad. Sci. USA. 82:6672-6676, 1985.
12. Reddy, M. V., Gupta, R. C., Rayderath, E. and Rayderath, K. 32P-Postlabeling test for covalent DNA binding of chemicals in vivo: Application to a variety of aromatic carcinogens and methylating agents. Carcinogenesis 5:231-243, 1984.

13. Everson, R. B., Randerath, E., Santella R. M., Cefalo R. C., Avitts, T. A., and Randerath, K. Detection of smoking related covalent DNA adducts in human placenta. *Science* 231:54-57, 1986.
14. Umbenhauer, D., Wild, C. P., Montesano, R., et al. O⁶methyldeoxy-guanosine in oesophageal DNA among individuals at high risk of oesophageal cancer. *Int. J. Ca.* 33:661-665, 1985.
15. Reed. E., Yuspa, S. H., Zwelling, L. A., Osols, R. F., Poirer, M. P. Quantitation of cis-diamminedichloroplatinum II (cis platin)-DNA-in trastrand adducts in testicular and ovarian cancer patients receiving cisplatin chemotherapy. *J. Clinical Investigations* 77: 545-550, 1986.
16. Weinstein, I. B. The Relevance of Tumor Promotion and Multistage Carcinogenesis to Risk Assessment. In: Banbury Report 19: Risk Quantitation and Regulatory Policy, D. G. Hoel, R. A. Merrilland F. P. Perera, Editors, Cold Spring Harbor Labs., New York, pp. 231-240, 1985. Also: "Concluding Remarks on Page 343.
17. Nishizuka, Y. Studies and Perspectives of Protein Kinase C. *Science* 233:305-312, 1986.
18. Harris, C. C., and Sun, T-t. Commentary: Multifactorial etiology of human liver cancer. *Carcinogenesis* 5:697-701, 1984.

Part IV

Quantitative Expression of Risk

Chapter 16

Approaches to Quantitative Expression
of Dose Responses

Ethel S. Gilbert

APPROACHES TO QUANTITATIVE EXPRESSION OF DOSE RESPONSE

Ethel S. Gilbert

INTRODUCTION

An important element in the risk assessment process is deciding how to utilize results from several different epidemiological studies to determine the risk estimates needed for establishing exposure standards. In order to compare results from different studies effectively, it is necessary to have quantitative estimates of risks based on the various studies. It is particularly important that the uncertainties in the estimates be quantified so that the relative validity of various studies can be evaluated.

This chapter will focus on quantifying risks experienced by populations that have been exposed to relatively low levels of harmful agents or substances. Often these are the levels that could reasonably be expected in future use of the agent or substance. Frequently these levels have been achieved by deliberately limiting exposures because of demonstrated risks of exposure at high levels. Currently there are a number of epidemiological populations exposed at the low levels achieved under various protection standards. The objective of these studies is to provide a direct assessment of risks associated with exposures at the actual levels of interest.

This situation is well illustrated by studies of health risks due to radiation exposure. There are several epidemiological studies of human populations exposed at high levels including the Japanese atomic bomb survivors

Work supported by the U.S. Department of Energy under Contract DE-AC06-76RLO1830.

and several populations that have been exposed for therapeutic reasons. These data demonstrate quite clearly that radiation exposure at high levels can induce cancer, and have been used to quantify risks from low level exposures such as might be received through occupational or diagnostic medical use. The available evidence on radiation risks as well as the models used to develop risk estimates for low level exposure are described in a report of the National Academy of Sciences known as the BEIR III report (1980).

Currently there are several epidemiological studies of populations that have received occupational exposure to radiation including workers at the Hanford nuclear facility (Gilbert and Marks 1979), naval shipyard workers (Rinsky, Zumwalde, Waxweiler et al. 1981), workers at several Oak Ridge facilities (Lushbaugh, Fry, Shy et al. 1983), and plutonium workers (Voelz, Wilkinson, Acquavella et al. 1981). There are additional studies being conducted in Canada and United Kingdom. Exposures to these workers have been deliberately limited so that if the estimates provided by BEIR III, or by other reports that have attempted risk estimation are correct, the number of radiation induced cancers in these populations is likely to be too small to result in statistically significant correlations.

An important objective of studies of populations subject to limited exposures is to determine if the limitation of exposures has achieved its goal of keeping risks to minimal levels. Although the scientists responsible for the risk estimation models may be confident that the approaches that have been used to extrapolate from high to low level exposures or from animal to man involve a high level of conservatism, that is, have tended to overestimate rather than underestimate risks, this confidence is not necessarily shared by the interested non-scientist or the general public. For some, the direct

Gilbert

-2-

assessment of risks based on populations exposed at the actual levels of interest may have more credibility than the more indirect approach.

A question of interest to the risk assessor is the role that such studies should play in the general process of determining risk estimates and in establishing exposure standards. The appropriate role can be determined only if the uncertainty in the findings is quantified so that it can be compared with the uncertainty in other studies, particularly those based on high level exposures. Even if a study has little potential for use in modifying risk estimates, the direct assessment of risks based on populations exposed at the actual levels of interest provides a valuable check on estimates obtained by less direct means. However, for this direct assessment to be properly interpreted, the uncertainties in the data set need to be clearly indicated.

It is noted that although the focus of this chapter is on populations exposed at low levels, many of the methods that are described are applicable to populations exposed at high levels. In the latter case, however, there would probably be issues of interest that cannot be meaningfully considered using data on low level exposures.

Methods for quantifying risks

Let us now consider methods of analyzing data on populations exposed at low levels. The usual approach to the analysis of such data is to conduct statistical tests investigating the association of the exposure of interest and death or incidence of several health effects. If no significant associations are identified, it is concluded that there is no evidence that risks exceed those upon which the standards have been based. Frequently, this is as far as the analysis goes.

Is there anything further that can be done to quantify the risks that may have been experienced by the exposed population? Or, from the standpoint of those involved in risk evaluation and standard setting, is there information in these studies beyond the statement of "no significant effects" that is useful?

Gilbert

-3-

It is the opinion of this author that there are additional steps that can be taken to provide a better assessment of what can and cannot be concluded based on the population under study.

A current study of the effects on mortality of occupational exposure to radiation at the Hanford plant will be used to illustrate methods of analysis that are useful. The Hanford plant, located in southeastern Washington State, has employed many workers in jobs involving some exposure to radiation. The initial purpose of this plant was the manufacture, chemical separation, and purification of plutonium. In addition, research of a diverse nature and, more recently, power generation have been conducted at the facility. Employment records, which include data on age, sex, race, dates of employment and termination, and job classification, have been kept on 28,683 male workers and 11,190 female workers employed between 1944 and 1979. Vital status has been established primarily by periodic searches of Social Security Administration files, but has recently been supplemented by the use of other sources.

External exposure to radiation is measured at most nuclear installations, including Hanford, through the use of dosimeters worn by employees. Such exposure has been measured in rem, a normalized sum of several radiation types. The Hanford data have been analyzed by several investigators including Gilbert and Marks (1979), Tolley, Marks, Buchanan et al. (1983), Mancuso, Stewart and Kneale (1977), Kneale, Mancuso and Stewart (1981, 1984), Hutchinson, MacMahon, Jablon et al. (1979) and Darby and Reissland (1981).

Until recently, most of our own analyses of these data have focused on testing the null hypothesis of no association of radiation exposure and mortality. In these analyses, death rates from several causes are related directly to exposure without introducing an external control. In a population

such as Hanford, for which relatively good exposure measurements are available, and which exhibits considerable exposure variability, such internal comparisons can be derived from the proportional hazards model of Cox (1972); their most recent application to the Hanford data is described in Gilbert (1985) and Gilbert and Peterson (1985). As a supplement to these analyses, comparisons with U.S. vital statistics have been made by calculating standardized mortality ratios (Gilbert and Marks 1979).

The analyses noted provide no evidence of an association of overall cancer mortality and radiation dose, and no evidence of an association of leukemia mortality and dose. Leukemia is the type of cancer that has been most strongly linked with radiation exposure in studies of populations exposed at high levels. Of seventeen cancer types tested for a positive association with radiation dose, only multiple myeloma shows a significant correlation ($p < 0.01$); this finding results from three deaths with relatively high exposures. Given the number of tests conducted, the possibility that this may be a false positive finding cannot be ruled out, and the appropriate interpretation of this correlation remains uncertain.

Estimates of radiation risks based on the Hanford data are too unreliable to be of much interest. However, confidence limits provide a useful way of quantifying the uncertainty in this data set. For example, it is possible for the Hanford data to show no evidence of positive correlations of radiation and various cancer types, yet be consistent with effects several times currently accepted estimates; large upper confidence limits will demonstrate this. On the other hand, claims are sometimes made that effects are magnitudes higher than standard estimates would indicate. Upper confidence limits that are less than such extreme claims serve to demonstrate that the Hanford data are inconsistent with such claims.

Gilbert

-5-

There are a number of difficulties in obtaining confidence limits. The size of the data set (over 13,000 longer term workers each with several years exposure data) may make it impractical to carry out the computations for certain procedures, while the highly skewed exposure distribution, and the small number of deaths from some causes of interest means that the asymptotic approximations required for many statistical procedures may not be entirely adequate. A description and justification for the approach that we have used in the presence of these difficulties is given in Gilbert (1985).

In order to obtain estimates and confidence limits, it is necessary to specify a model. The model we have used in most analyses is one in which the radiation risks are assumed to be proportional to spontaneous or non-radiation related risks. This model is the so-called relative risk or proportional hazards model, and has the major advantage that parameters that quantify the radiation risks can be estimated independently of the spontaneous risks as indicated by Cox (1972). It has also been assumed that the risk due to radiation is a linear function of the cumulative dose of the worker. These two assumptions in combination mean that the hazard or risk for a worker alive with cumulative dose z in the age-calendar year category x is given by

$$S(x)[1 + Bz]$$

where $S(x)$ is the spontaneous risk for this category, and B is the proportional increase in the risk per unit of radiation exposure (rem), and which must be estimated. A latent period of 2-years has been incorporated in analyzing leukemias, and of ten years in analyzing other cancers by calculating the dose z only up to two (or ten) years prior to the time of interest.

Once the parameter B and its upper confidence limits are estimated, the number of radiation induced cancers that would be expected based on these

parameter values can be determined. For a death with cumulative exposure z , the a posterior probability that the death was caused by radiation is given by $z/(1 + Bz)$. This formula is derived by applying Bayes Theorem. The estimated number of radiation induced deaths in the study population or any segment of the population can be determined by summing such expressions over all deaths. The estimated number of spontaneous deaths can also be determined as the sum of the quantities $1/(1 + Bz)$.

The use of the above model for estimating cancer risks other than leukemia can be justified based on data on populations exposed at high levels and reviewed in BEIR III (1980). Note that since there is no evidence of radiation effects based on the Hanford data alone, it is obvious that the shape of the dose-response function or whether the relative risk model is appropriate cannot be determined. Although one can check whether the model is consistent with various models developed from data on high level exposures, there is very little power for detecting deviations from these models.

For leukemia, Darby (1983) has demonstrated that neither the relative risk model nor the absolute risk model, in which risks are assumed to be independent of spontaneous risks, provide an adequate fit to data on populations exposed at high levels. BEIR III estimates, however, are based on the absolute risk model. Estimates for leukemia risks based on both models have been obtained. In applying the absolute risk model, the spontaneous rates $S(x)$ do not "cancel out" as in the relative risk model. Since the data are not strong enough to individually estimate the spontaneous rates for each age and calendar year category, to obtain estimates based on the absolute risk model, it is assumed that the spontaneous rates are proportional to age-calendar year-specific rates

for the U.S. population. The proportionality constant is estimated from the data. To summarize, the hazard or risk for the absolute risk model is defined as

$$A S(x) + C z$$

where the $S(x)$ are taken from vital statistics and A and C are estimated. With this model, the number of radiation induced deaths is estimated by multiplying the estimated coefficient C (or its upper confidence limit) times the number of person-year-rem in the study population.

There are problems in comparing estimates and confidence limits based on the Hanford data with estimates presented in BEIR III and other documents. One difficulty is that Hanford estimates are based on film badge readings while BEIR III estimates are based on estimated dose to the organ of interest. An investigation of the relationship of film badge readings and doses to various organs is currently underway, but is not yet complete. The appropriate adjustment will probably vary by calendar year, by the type of work, and work location. For the time being a crude adjustment can be obtained by assuming that half of the amount read by the dosimeter reaches the internal organs of interest. This would have the effect of doubling Hanford estimates that are expressed per unit of exposure.

Data on the Japanese atomic bomb survivors have played a major role in determining the risk estimates presented in BEIR III. Recently, the dose estimates used in the Japanese studies have been seriously challenged, and studies are now in progress to determine new dose estimates. Jablon (1984) has noted that the likely effect of the dose revision will be to increase estimates based on the earlier dosimetry by a factor in the neighborhood of two. Since based on information available at the present time, the most plausible

Gilbert

adjustment (multiplying by a factor of two) seems to be the same for both the Hanford estimates and the BEIR III estimates, the ratio will be unaffected, and no adjustment to estimates from either source has been made. It is hoped that additional dosimetry information from both populations will soon allow a more precise treatment of this issue.

An additional difficulty in comparing Hanford upper confidence limits with BEIR III estimates, especially for cancers other than leukemia, is that the BEIR III report does not directly indicate a single coefficient appropriate for a population with the Hanford age structure and expressed in the manner of our upper confidence limits. To handle this problem, we have applied the BEIR III models to the Hanford data, taking age at exposure into account, and determined the number of radiation-induced cancer deaths that would have been expected in this population based on these models. For comparison, we have also calculated the number of radiation-induced cancers that would be expected based on estimates and upper confidence limits obtained directly from the Hanford data.

Application to the Hanford data

Estimates and confidence limits for leukemia risks based on both the relative and absolute risk model are presented in Table 1. These estimates are based on deaths of white males employed at least two years at Hanford occurring between January, 1947 and January 1, 1979. Because an association of chronic lymphatic leukemia and radiation exposure has never been demonstrated, this type of leukemia (ICD code 204) has been excluded from these analyses. In addition to the estimated parameter (B or C), the number of radiation induced cancers in the study population, based on the estimates and on their 95 percent confidence limits, are presented.

Gilbert

-9-

The estimates for leukemia, based on either the relative and absolute risk model, are negative, but the upper 95 percent confidence limits are both positive. Based on the relative risk model, the upper 95 percent confidence limit for the number of radiation induced deaths from leukemia in Hanford workers is 1.1. The upper confidence limit (obtained from the absolute risk model) is 4.2 leukemia deaths per million person-year-rem. This would result in an estimated 3.3 radiation induced leukemia deaths.

The discrepancy in the results obtained from the relative and absolute risk models results at least in part from differences in assumptions regarding spontaneous risks. The relative risk model requires no assumptions about these risks, while the absolute risk model requires the assumption that age-calendar year specific rates for the Hanford population are proportional to those obtained from U.S. vital statistics. There is some indication in our data that this assumption is not appropriate since the ratios of observed deaths and expected deaths (based on U.S. rates) show an increasing trend with calendar year.

The BEIR III estimates for leukemia are based on an absolute risk model using estimates obtained from the Japanese atomic bomb survivors. Estimates based on linear-quadratic, judged most realistic, and linear extrapolation are given. A straight-forward application of the BEIR III models yields the higher numbers, that is 2 radiation-induced leukemias based on the linear model and 0.9 radiation-induced leukemias based on the linear-quadratic model. The lower numbers represent an attempt to take into account that BEIR III estimates are based on leukemia incidence, which is no longer identical to leukemia mortality (National Cancer Institute 1981), and to account for a possible healthy worker effect. To the extent that the reduced mortality often seen in employed

populations results from incomplete ascertainment of deaths or lower fatality rates due to early treatment and cure, this effect can be expected to affect adiation-induced cancers as well as cancers due to other causes.

With the relative risk model, the maximum expected number of leukemia deaths compatible with the Hanford data (1.1) is similar to the number expected based on the BEIR III linear model. However, the maximum expected number of leukemia deaths calculated with the absolute risk model for the Hanford data (3.3) is about three times the number obtained by applying the BEIR III linear model. These results can be interpreted as meaning that the Hanford data provide direct evidence that leukemia risks due to low levels of radiation cannot be much larger than those obtained through linear extrapolation from estimates based on populations exposed at high levels.

However, these results should be interpreted with caution. Because there are only twelve deaths from non-lymphatic leukemia in the Hanford population, it is possible that the normal approximations used to obtain the confidence limits are not entirely adequate. The uncertainty in comparing estimates based on Hanford dosimetry with those based on Japanese dosimetry has already been noted.

Estimates and confidence limits for cancers other than leukemia are also presented in Table 1. These estimates are based on deaths occurring between December 31, 1954 and January 1, 1979. Recall that a ten year latent period was accounted for by lagging exposure by ten years. For cancers other than leukemia, there is fairly good evidence based on data from other sources that the relative risk model provides a better fit than the absolute model, so that absolute risk model was not considered. Applying the BEIR III linear-quadratic and linear relative risk models to the Hanford data, it can be estimated that

approximately 0.5 and 1.2 radiation-induced deaths, respectively, would have occurred.

The estimated risks based on Hanford data are negative indicating that these data provide no direct evidence of an association of radiation exposure and cancer mortality. However, the upper 95 percent confidence limit obtained from the Hanford data is nearly thirty times estimated risks based on the BEIR III linear-quadratic model, and more than ten times those based on the BEIR III linear model. Although these results could be used to counter extreme claims that effects might be 100 or more times those provided by BEIR III, it is clear that the Hanford data are relatively uninformative with respect to the magnitude of risks due to cancers other than leukemia resulting from low level exposure to radiation.

CONCLUSIONS

What do the results presented tell the risk assessor about the role of the Hanford data in providing risk estimates or setting standards for exposure? First of all, since no risks have been directly demonstrated using data from this population, the data set provides no reason for thinking that current risk estimates are too high or that standards are inadequate. For leukemia, the Hanford data appear to demonstrate directly that risk estimates obtained by linear extrapolation from data on populations exposed at high levels are unlikely to underestimate risks by a factor greater than three. It will be interesting to see if this finding is substantiated when a more appropriate method for calculating dose to the bone marrow from film badge readings is developed. For cancers other than leukemia, although the data provide no evidence of a positive correlation of radiation exposure and cancer mortality,

Gilbert

-12-

it is clear that we must continue to rely on data from population exposed at high levels. At the present time, such data allow us to place tighter limits on risks than does a direct assessment based on the Hanford data.

However, studies of populations exposed to low levels of radiation do have a role to play. There is more to be learned as follow-up on the Hanford workers is extended, and there are several other similar populations in the United States, Canada, and Great Britain that are currently being studied. If data from these populations could be pooled, these studies could potentially play a larger role in the task of assessing risks. As a minimum, they serve as a valuable check on estimates obtained from populations exposed at high levels.

It should be noted that in order to fully determine the appropriate role of studies of populations exposed at low levels, one needs to compare the uncertainty resulting from the direct approach (using low level data) with the uncertainty in estimates obtained from populations exposed at high levels. Adequately quantifying uncertainties in the latter estimates is extremely difficult, since many of the uncertainties relate to model assumptions and cannot be readily quantified using standard statistical procedures.

TABLE 1: Estimates and upper 95 percent confidence limits for risk of death from cancer in white males employed at least two years at Hanford plant

	Based on Hanford Data					
	<u>Relative Risk Model</u>		<u>Absolute Risk Model</u>		<u>Based on BEIR III</u>	
	Estimate	Upper 95% Confidence Limit	Estimate	Upper 95% Confidence Limit	Linear Quadratic Model	Linear Model
Leukemia						
Percent increase per rem	-3.4%	5.6%	-	-	-	-
Excess deaths per million person-year-rem	-	-	-7.7	4.2	1.0	2.24
Expected number of radiation-induced deaths	0	1.1	0	3.3	0.5-0.9	1.2
Cancer other than leukemia						
Percent increase per rem	-0.3%	1.9	-	-	-	-
Expected number of radiation-induced deaths	0	14.6	-	-	0.5	1.2

REFERENCES

- BEIR Report. (1980) The Effects on Populations of Exposure to Low Levels of Ionizing Radiation: 1980. Report of the Advisory Committee on the Biological Effects of Ionizing Radiation, Division of Medical Sciences, National Academy of Sciences, National Research Council, Washington, DC.
- Cox, D.R. (1972) Regression models and life tables. J. Royal Stat. Soc., Series B 34:187-220.
- Darby, S.C. (1983). Modelling age- and time-dependent changes in the rates of radiation-induced cancers. In, Atomic Bomb Survivor data: Utilization and Analysis, (ed. R. L. Prentice and D.J. Thompson). Society for Industrial and Applied Mathematics, Philadelphia, pp. 67-80.
- Darby, S. and Reissland, J. (1981) Low levels of ionizing radiation and cancer — Are we underestimating the risk? J. Royal Stat. Soc., Series A 144:298-331.
- Gilbert, E.S. (1985) How much can be learned from populations exposed to low levels of radiation? The Statistician 34:133-44.
- Gilbert, E.S. and Marks, S. (1979) An analysis of the mortality of workers in a nuclear facility. Radiat. Res. 79:122-48.
- Gilbert, E.S. and Peterson, G.R. (1985) A note on 'job related mortality risks of Hanford workers and their relation to cancer effects of measured doses of external radiation'. Br. J. Ind. Med. 42:137-39.
- Hutchison, G.B., MacMahon B., Jablon S., and Land C.E. (1979) Review of a report by Mancuso, Stewart and Kneale on radiation exposure of Hanford workers. Health Physics 37:207-20.
- Jablon, S. (1984) Characteristics of Current and Expected Dosimetry. In, Atomic Bomb Survivor Data: Utilization and Analysis, (ed. R.L. Prentice and D.J. Thompson). Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.
- Gilbert

- Kneale, G.W., Mancuso T.F. and Stewart A.M. (1984) Job related mortality risks of Hanford workers and their relation to cancer effects of measured doses of external radiation. Br. J. Ind. Med. 41:9-14.
- Lushbaugh, C.C., Fry, S.A., Shy, C.M. and Frome E.L. (1983) DOE Health and Mortality Study at Oak Ridge. In, Health Physics Society: Epidemiology Applied to Health Physics, pp. 436-43. Proceedings of the Sixteenth Midyear Topical Symposium, U.S. Department of Energy Report CONF-830101, Albuquerque, New Mexico.
- Mancuso, T.F., Stewart A. and Kneale G. (1977) Radiation exposures of Hanford workers dying from cancer and other causes. Health Physics 33:369-85.
- National Cancer Institute. (1981) Surveillance, Epidemiology, and End Results. Incidence and Mortality Data: 1973-77. NCI Monograph 57, U.S. Government Printing Office, Washington, DC.
- Rinsky, R.R., Zumwalde R.D., Waxweiler R.J., Murray, W.E. Jr., Landrigan, D.J., Bierbaum P.J., Terpilak M. and Cox C. (1981) Cancer mortality at a naval shipyard. Lancet 1:231-35.
- Tolley, H.D., Marks, S., Buchanan, U.S. and Gilbert, E.S. (1983) A further update of the analysis of workers in a nuclear facility. Radiation Res. 95:211-13.
- Voelz, G.L., Wilkinson G.S., Acquavella J.F., Tietjen G.L., Brackbill R.M., Reyes M. and Wiggs L.D. (1981) An Update of Epidemiological Studies of Plutonium Workers. LA-UR-82-123, Los Alamos National Laboratory, Los Alamos, New Mexico

Chapter 17
Statistical Modeling of Dose-Response Relationships
Ron Brookmeyer

Introduction

Time is an important factor to consider in the quantitative assessment of risk. Time may be measured from a number of different origins including time since first exposure, time since last exposure, chronological age and calendar year. This chapter considers some statistical methods and pitfalls associated with time factors in risk assessment. The focus is on the epidemiologic cohort study.

The importance of time as a determinant of risk is predicted by the multistage model of carcinogenesis (Armitage and Doll, 1961). The model hypothesizes that the transition of a normal cell to a malignant one involves K successive stages, and carcinogens act to alter at least one of the transition rates between stages. According to the model, the effects of age at first exposure, duration of exposure and time since last exposure depend on which stage the exposure acts. (Whittemore, 1977; Day and Brown, 1980; Day, 1984³). The excess risk from a single exposure which acts on an early stage, will take a long time to manifest but the risk from an exposure which acts on a late stage will emerge more quickly. The multistage model is a convenient framework for understanding the role of time in carcinogenesis.

This chapter focuses on simple statistical methods for analyzing the role of time in the epidemiologic cohort study. The statistical methods for the standardized mortality ratio will be reviewed, then some common pitfalls are considered and finally design and monitoring issues in relation to time factors will be discussed.

Statistical methods

One of the oldest and most useful epidemiologic indices of risk is the standardized mortality ratio (SMR). It has been reviewed and critiqued by Gaffey (1975) and its theoretical framework developed by Breslow (1977). In this section the focus will be on the evolution of risk over a single time scale such as follow-up time or time since exposure. The theoretical SMR θ is defined as the ratio of the hazard (risk) function for an exposed individual to the hazard function for a comparable unexposed individual. We have

$$\theta = h_i(t)/h_1^0(t) \quad (1)$$

where $h_i(t)$ is the hazard function for the i th individual in the exposed cohort t time units after exposure and $h_1^0(t)$ is the hazard for an unexposed but otherwise comparable individual; comparable usually refers to an individual with the same age, year of birth, sex and race. The hazard function $h_1^0(t)$ is typically obtained from standard population rates. Formal estimation of the SMR can proceed by consideration of the likelihood function for θ . The likelihood function $L(\theta)$ is given by

$$L(\theta) = \prod_{i=1}^N [\theta h_1^0(t_i)]^{\delta_i} [\mu_i(t)]^{1-\delta_i} \exp\left[-\int_0^{t_i} \theta h_1^0(t) + \mu(t) dt\right]$$

where t_i is the total follow-up time for the i th individual; δ_i indicates if the individual is a case at t_i ($\delta_i = 1$ if case, $\delta_i = 0$ if not); and $\mu_i(t)$ is a risk function for leaving the cohort for all reasons other than becoming a case (for example: death from other causes, loss to follow-up, withdrawn alive). If we assume that the risk function $\mu_i(t)$ is functionally independent of θ , the likelihood function $L(\theta)$ is proportional to

$$L(\theta) = \prod_{j=1}^N [\theta h_1^0(t_1)]^{\delta_1} \exp[-\theta \int_0^{t_1} h_1^0(t) dt] \quad (2)$$

The first term on the left of expression (2) is the risk of becoming a case at t_1 , while the second term is the probability of surviving up to t_1 . Setting the derivative of the log-likelihood function to 0, the maximum likelihood

estimate is $\hat{\theta} = \frac{N}{\sum_{i=1}^N \delta_i} / \frac{N}{\sum_{i=1}^N \int_0^{t_1} h_1^0(t) dt} = O/E$, where O is the observed number of cases and E (the sum of cumulative hazard functions) is the "expected" number of cases if the cohort was subject to standard population rates. In practice, the expected number is computed by classifying each person-year of observation by age and calendar year (usually by 5 year groupings), totaling the person-years in each age-calendar year cell, multiplying by the standard rates of the cell and summing over all cells. However the "expected" number E is not a true expectation in the usual sense (in fact it may be larger than the total sample size), rather it is a random variable which depends on the observed follow-up; the expectation of E is the true expected number, that is, $E(\int_0^{t_1} h_1^0(t) dt) = P(\delta_1 = 1)$ (Berry, 1983). In order to assess whether the SMR is greater than 1, the well known statistic $(O-E)^2/E$ is referred to a χ^2 distribution with one degree of freedom.

Statistical pitfalls and issues

1) Does the risk evolve over time ?

Implicit in equation (1) is the proportional hazards assumption. The relative risk for exposure is assumed constant over follow-up time. Model (1) assumes the relative risk does not evolve with time, for example the relative risk one day after exposure is the same five years after exposure.

In many situations this is unrealistic and thus the reporting of a single SMR is misleading. In fact, a single SMR computed over the whole follow-up period is not the average relative risk over time but rather a weighted average where the weights are the expected number of cases at each point in time. Since expected numbers depend on the number of person years of follow-up (as well as the standard population rates) and often there are more person years in early follow-periods, a single SMR may be weighted more toward the early risk than the later risk. For example, if the relative risk immediately following exposure was $\hat{\theta}_1 = O_1/E_1$ and later on it was $\hat{\theta}_2 = O_2/E_2$, the single SMR estimated overall follow-up time is

$$\hat{\theta} = \frac{E_1}{E_1 + E_2} \hat{\theta}_1 + \frac{E_2}{E_1 + E_2} \hat{\theta}_2$$

The interpretation of a cohort study should not depend on a single observed to expected ratio but on the way this ratio evolves with time since exposure. Consider for example the the International Radiation Study of Cervical Cancer (Day and Buice, 1984). The objective of this study was to quantify the excess risk over time of second primary cancers following radiotherapy for an initial cancer of the cervix. The study involved fifteen population-based cancer registries, many of which began operation in the early 1950's and over 82,000 women with invasive cancer treated with radiotherapy. There was a much smaller group of women treated without radiotherapy (14,175), although comparisons will be made with rates from standard populations. Follow-up ended for most registries in 1977. Standardized mortality ratios in the radiation study were calculated for

the time periods 1-4, 5-9, 10-14, 15-19, 20-24 and 25+ years after the diagnosis of the cervical cancer. For sites close to or at intermediate distance from the cervix, the results have been combined and are shown in Figure 1A for those treated by radiotherapy, and in Figure 1B for cases not treated by radiotherapy. These figures show the evolution of risk with time since diagnosis. For patients treated with radiotherapy there does not appear to be excess risk within 10 years from first exposure, however there is an increasing trend of the relative risk with time beyond 10 years. The single SMR estimated overall time is $\theta = 1.08$ for the radiotherapy group, and $\theta = 1.0$ for the no radiotherapy group and are shown in Figure 1A. The pattern of risk over time would go unnoticed with an analysis based on a single SMR.

2) Allocation of Person-Years to Time Periods

A source of confusion associated with risk estimation in multiple time periods is the allocation of person-years. Individuals contribute person-years experience to a time dependent category only during the years they actually spend in that category. For example, an individual followed 12 years in the International Radiation Study contributes 4, 5 and 3 person-years respectively to the time periods 1-4, 5-9 and 10-14 years since diagnosis. A common error has been the allocation of all 12 years to the 10-14 category. This procedure overestimates the expected number in later time periods and therefore underestimates the relative risk in later time periods. This fallacy was pointed out by Wagoner, Infante and Saracci (1976) in a cohort study of workers exposed to vinyl chloride and cancer mortality. Enterline (1976) discussed the fallacy in the context of dose-

response estimation in a cohort study of asbestos workers and respiratory cancer when estimating the SMR by cumulative years of employment. Workers should not be classified by their final cumulative years, rather their person-year contributions should be allocated dynamically as they are followed over time.

The correct allocation of person-years is clarified by formal consideration of the likelihood function. The likelihood for the SMR in a given time period T_1 to T_2 , requires slight modification to expression (2). The only individuals who contribute factors to the likelihood are those who were followed beyond T_1 . The term δ_i indicates whether the i th individual became a case in the period T_1 to T_2 and the survival probability in (2) is conditional on surviving up to T_1 :

$$L(\theta) \propto \prod [\theta h_i^0(t_i)]^{\delta_i} \exp[-\theta \int_{T_1}^{t_i^*} h_i^0(t) dt] \quad (3)$$

where $t_i^* = \min(t_i, T_2)$ that is we integrate up to the last follow-date of individual or the end of the time period T_2 , which ever comes first.

Then the maximum likelihood estimate is $\hat{\theta} = \frac{\sum_{i=1}^N \delta_i}{\sum_{i=1}^N \int_{T_1}^{t_i^*} h_i^0(t) dt}$, again the hazard is integrated only over the time the i th individual was followed in the period T_1 to T_2 . In other words, the SMR is the ratio of the observed

number of cases in the interval T_1 to T_2 to the "expected" number, where the expected number is based on the number of years each individual spent in the interval.

3) Allocation of person-years to prevalent cohorts

A similar fallacy may occur with the allocation of person-years from a prevalent cohort. An occupational study may include all workers already employed at the start of the study as well as new hires during the follow-up period. For example, suppose a study began in 1960, and continuously followed all workers. An individual who was employed since 1951 could contribute person-years only to the time period 10 or more years after first employment. It is an error to allow that individual to contribute person-years to the 1-9 year category and such an error would produce an underestimate of risk for the 1-9 year period. The intuition for this is clear: An individual hired in 1951 who died before 1960 would not be in the cohort and therefore could not contribute to the observed cases, only the survivors would be included in the cohort. When individuals must survive simply to be included in the study, it is referred to as left truncation.

Again, the correct allocation of person-years could also be obtained by careful consideration of the likelihood. An individual who started employment E_1 years prior to the start of the study could only enter the cohort if he survived E_1 years; that is entry into the cohort is conditional on surviving E_1 years. Based on a likelihood function

analogous to (3) we obtain $\hat{\theta} = \frac{\sum_{i=1}^N \delta_i}{\sum_{i=1}^N \int_{E_1}^{t_i} h_1^0(t) dt}$. The bottom line is

that individuals only contribute to the expected numbers during their period of continuous follow-up.

4) Noncontinuous follow-up

Occasionally there is noncontinuous follow-up. Suppose only currently employed or retired individuals can be followed and an individual leaves employment 10 years after first employment and returns 5 years later. The individual does not contribute person-years for the period 10-15 years from first employment, for if the individual became a case during that period we would have no records of it. This too can be made formal by careful consideration of the likelihood.

5) Continuous and discrete time

Our focus has been risk estimation for several time periods. A number of questions arise concerning how to categorize or discretize the time axis. If the time intervals are taken too wide, important trends in time may be overlooked. An estimate of the relative risk as a continuous function of time would be more ideal. Breslow (1985) proposed a technique for estimating the SMR as a continuous function of time. A computationally simpler approach, is to consider a moving window of the time axis; a smoothed estimate of the SMR over time can be obtained by considering overlapping time intervals for example 1-4, 2-5, 3-6, 4-7, 5-8 years from the first exposure. These calculations can be performed with standard person-year programs by varying the time period. However, a number of theoretical questions arise including optimum choice of the window width. If the relative risk changes rapidly with time one would prefer a smaller window

width, although at the expense of less precision because of fewer person-years of observation in narrower windows: the familiar tradeoff between bias and variance.

6) Latency and the average time to disease

Latency has been defined as the time from initial exposure to clinical appearance of disease (Land and Norman, 1978). The calculation of the average or median time from first exposure to disease among those who develop the disease can be very misleading for two reasons: First, individuals may develop the disease subsequent to the last follow-up date. Ignoring these individuals will in general produce a smaller average than if individuals were followed indefinitely. A natural suggestion is to treat these individuals as censored observations, and compute the Kaplan-Meier survival curve. This raises a second more serious objection. Implicit in survival curve estimation is the assumption of independence between the censoring and survival mechanisms. If the focus is on only one cause of death, it must be assumed that competing causes of death act independently. A more appropriate way to formulate the problem is in terms of cause-specific hazards (Prentice, Kalbfleisch, Peterson et al., 1978). A graphical display of the ratio of the cause specific hazard function of the exposed to standard population (the relative risk or SMR) as a function of time is informative and free of the methodologic problems associated with reporting an average or median latency time. The minimum time to any excess risk (the time until the SMR first becomes greater than 1) could be determined from such a display and may be a more useful epidemiologic index than an average latency time.

Multiple time scales

There are often multiple time scales, and the data analyst is faced with the task of sorting out their separate effects. Time scales of major interest in risk assessment include time since first exposure, chronological age, cumulative duration of exposure, time since last exposure and calendar year. These variables are called time dependent since they change with follow-up. The other class of variables, called fixed covariates, include race, sex and age at entry into the cohort.

It is difficult to sort out the separate effects of time dependent variables because they are highly correlated. For example, if exposure in an industry is removed in a given year, time since last exposure is completely confounded with calendar year.

Two approaches for sorting out the separate effects of time dependent variables have been described by Breslow (1983). The first involves discretizing or categorizing the time scales, thus creating a multi-dimensional contingency table and allocating person-years to each cell along the lines described previously. For example, two time scales of primary interest in a study of respiratory cancer deaths among Montana smelter workers were cumulative years of heavy arsenic exposure and cumulative years of moderate arsenic exposure. Table 1 is calculated by allocating person-years according to the number of years an individual spent in each cell. The expected number of deaths is computed by multiplying each person-year by the standard rate of the age and calendar year at the time the individual contributed the person-year. In order to

detect systematic patterns in the relative risk over the cells, the observed number of cases in the j th cell O_j may be modelled as a Poisson variable with mean $E_j \exp(\beta X)$ where E_j is the expected number of cases, X is the vector of covariates, and β are the regression coefficients to be estimated from the data. Poisson regression has been well studied (Frome and Checkoway, 1985, Nelder, 1984) and thus multiple time scales do not create serious problems provided person-years are allocated to each cell correctly.

Method 2 called the internal analysis, is based on the partial likelihood formulation of a proportional hazards model. It does not use external rates from a standard population, rather comparisons are made among individuals within the cohort. Consider the regression model

$$h(t_1, t_2) = h^0(t_1, t_2) e^{\beta X}$$

where t_1 is chronological age, t_2 is calendar year, and X is a vector of covariates possibly time dependent (such as cumulative years of moderate and heavy arsenic exposure). The likelihood consists of a product of factors, one for each case of respiratory cancer. The factor for the i th case is found by matching each case of respiratory cancer to a set of controls who were born in the same year as the case, and alive at the time of the case's death. The factor for the i th matched set is

$$\frac{\exp[\beta X_{i1}]}{\sum_{j \in R_i} \exp[\beta X_{ij}]}$$

where R_i is the risk set (i.e., the i th case and his controls) and X_{ij} is the covariate vector for the j th member of the risk set (X_{i1} corresponds to the case). The regression coefficients are found by maximizing the likelihood.

Implicit with both methods is a consistency of the relative risk over age and calendar year. Expected values in Method 1 were found by standardizing over age and calendar year. The underlying assumption is that the relative risk or SMR associated with a given exposure is the same for all age-calendar year cells. An equivalent assumption is implicit in Method 2: the regression coefficient (β 's) are assumed to be the same for all risk sets. If the relative risk is not constant across age-calendar year cells, an analysis stratified by age and calendar year is required. The moral is clear: SMR's are summary relative risk measures, collapsed over age and calendar year time scales. If we expect the relative risk for 60 year olds to be different than for 40 year olds all other things being equal, standardizing over chronological age is inappropriate.

Design and monitoring considerations

As discussed, a single SMR computed over all follow-up time has limited value. The design of a cohort study should account for this and consider risk estimation in the time periods of primary interest. Assuming that the observed number of cases has a Poisson distribution with mean ($\theta \cdot E$), the standard error of the SMR is $\sqrt{\theta/E}$; the precision depends on both the true relative risk and the expected number of cases. The critical elements in assessing the precision of a cohort study include a prediction of the expected number of cases in the time period of interest together with the hypothesized relative risk in that time period. A prediction of the expected number of cases depends on the accrual pattern over time.

If we are dealing with a prevalent cohort which will be followed forward, a simple approximation to the expected number of events can be obtained by assuming all individuals remain in the cohort over the follow-up period. Of course this will be an overestimate as there will be losses from the cohort due to competing causes of death, and loss to follow-up. The advantage of this approximation is that it is easy to compute with standard person-year programs by setting the last follow-up date for each individual to the proposed duration of the study. A more refined approximation can be obtained by incorporating standard mortality figure to account for competing causes of death as in Brookmeyer, Day and Pompe-Kirn (1985). Additional accrual can be accounted for by assuming individuals arrive according to a Poisson process with an expected accrual of μ /year. If the duration of the follow-up period is L years, and we are interested in risk estimates T_1 years after entry into the cohort, then only $(L-T_1)\mu$ individuals could be expected to make person-year contributions; that is, an individual must enter at least T_1 years prior to the end of the study. By allocating the μ arrivals each year according to anticipated ages at entry, a prediction of their contribution to the expected numbers can be obtained from standard person-years programs.

Related issues concern monitoring a cohort study over time. A question of interest in the International Radiation study of Cervical Cancer was the value of extending follow-up a further five years, in particular the extra information it would provide on risk 25 years or more after exposure. A general approach for addressing this problem has been considered

in Brookmeyer, Day and Pompe-Kirn (1985). The methods involve predicting the gain in the number of expected cases in each time period. Additional follow-up has a different impact on precision of the risk estimate in each time period. For example, a five year extension of the Radiation Study would not affect the risk estimate 1-4 years after diagnosis since all those in the cohort had already passed the four year mark. On the other hand, five years of additional follow-up could produce considerable gains in the precision of the risk estimate 25 years after diagnosis. A convenient display of the impact of five years of additional follow-up on the risk estimate 25 years after exposure is a precision-effect curve (Figure 2). This display is a joint prediction of the plausible values for the SMR and its error bound (two standard errors) after additional follow-up. We expect a gain in precision of the risk estimate in the radiotherapy group, but we do not expect the SMR to fluctuate much; we anticipate that the basic conclusions will remain unchanged. A different picture emerges for the no radiotherapy group. Not only will there be important gains in precision, but the final SMR may fluctuate substantially.

In summary, the design, monitoring and analysis of the epidemiologic cohort study must carefully consider the role of time. A single relative risk for an exposure assumes that risk does not evolve with time but this is rarely the case. A display of the risk over time is informative and avoids the ambiguities associated with a single risk estimate. In general, risk will not be a function of a single time scale, but rather a complex function of multiple time scales. Multiple time scales create a number of

problems but careful consideration of the likelihood function can clarify the statistical procedures, in particular the appropriate allocation of person-years.

Similar issues concerning time factors arise in the design and analysis of case-control studies. In order to account for the evolution of risk over time, the exposure history must include both the dose and the time of exposures. The methodologic problems associated with reporting a single odds ratio are analogous to reporting a single SMR. In conclusion, the statistical design and analysis of epidemiologic studies should routinely address the role of time for a complete assessment of risk.

REFERENCES

- Berry, G. (1983). The analysis of mortality by the subject years method. *Biometrics* 39:173-84.
- Breslow, N. (1985). Nonparametric estimation of relative mortality functions. Presented at Symposium on Time-Related Factors in Cancer Epidemiology, National Institute of Health, Bethesda, Maryland.
- Breslow, N. (1977). Some Statistical Models Useful in the Study of Occupational Mortality. In: (ed. A. Whittemore). *Environmental Health: Quantitative Methods*. Society for Industrial and Applied Mathematics.
- Breslow, N., Lubin, J., Marek, P., and Langholtz, B. (1983). Multiplicative models and cohort analysis. *Journal of American Statistical Association* 78:1-12.
- Brookmeyer, R., Day, N. and Pompe-Kirn, V. (1985). Assessing the impact of additional follow-up in cohort studies. *Am J Epidemiol* 121:611-19.
- Day, N. E. and Boice, J. D. (1984). Second Cancer in Relation to Radiation Treatment for Cervical Cancer: Results of a Cancer Registry Collaboration. IARC Scientific Publication No. 52. Lyon: International Agency for Research on Cancer.
- Day, N. E. (1983). Time as a determinant of risk in cancer epidemiology: The role of multistage models. *Cancer Surv* 2:577-94.
- Day, N. E. and Brown, C. C. (1980). Multistage models and primary prevention of cancer. *J Nat'l Cancer Instit* 64:977-89.

- Enterline, P. E. (1976). Pitfalls in epidemiological research: An examination of the asbestos literature. *J Occup Med* 18:150-56.
- Gaffey, W. R. (1976). A critique of the standardized mortality ratio. *J Occup Med* 18:157-60.
- Land, C. E. and Norman, J. E. (1978). Latent Periods of Radiogenic Cancers Occuring Among Japanese A-Bomb Survivors. In Late Biological Effects of Ionizing Radiation Volume 1, International Atomic Energy Agency. Vienna.
- Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., Flournoy, N., Farewell, V. and Breslow, N. (1978). The analysis of failure times in the presence of competing risks. *Biometrics* 34:541-54.
- Wagoner, J. K., Infante, P.F. and Saracci, R. (1976). Vinyl chloride and mortality? *Lancet* 2:194-95.
- Whittemore, A. (1977). The age distribution of human cancer for carcinogenic exposures of varying intensity. *A J Epidemio* 106:418-32.

Table 1

Table 1. SMR by Cumulative Years of Heavy and Moderate Arsenic Exposure SMR (Expected Cases)

		YEARS OF HEAVY ARSENIC EXPOSURE							
		0		< 1		1-4		5+	
Years of moderate arsenic exposure	0	1.3	(20.8) ^a	1.7	(1.2)	5.4	(.6)	10	(.6)
	< 1	1.4	(4.9)	2.6	(.8)	5.0	(.2)	8.9	(.3)
	1-4	2.6	(3.1)	12.1	(.3)	10.	(.1)	9.1	(.1)
	5-14	2.5	(1.6)	0	(.1)	0	(.1)	0	(.01)
	15+	3.5	(1.1)	9.1	(.1)	0	(.1)	0	(.03)

Source: Breslow (1983):

(a) Expected number of cases

LEGENDS TO FIGURES

Figure 1 Standardized mortality ratios for sites close to and at intermediate distance from the cervix. for invasive cervical cancer patients treated (a) with radiotherapy and (b) without radiotherapy. 80% confidence intervals presented. (Dotted line is single SMR computed overall follow-up time) (FROM DAY & BOICE, 1974)

Figure 2 Precision effect curves for radiotherapy and no radiotherapy groups with 5 years of additional follow-up and 75% and 95% prediction intervals: X marks SMR and 95% error bound at close of follow-up in 1977. (J. R. BROOKMEYER, R. DAY AND P. KIRN, 1985)

Figure 1A
 RADIO THERAPY
 CLOSE AND INTERMEDIATE SITES
 Figure 1B
 NO RADIOTHERAPY

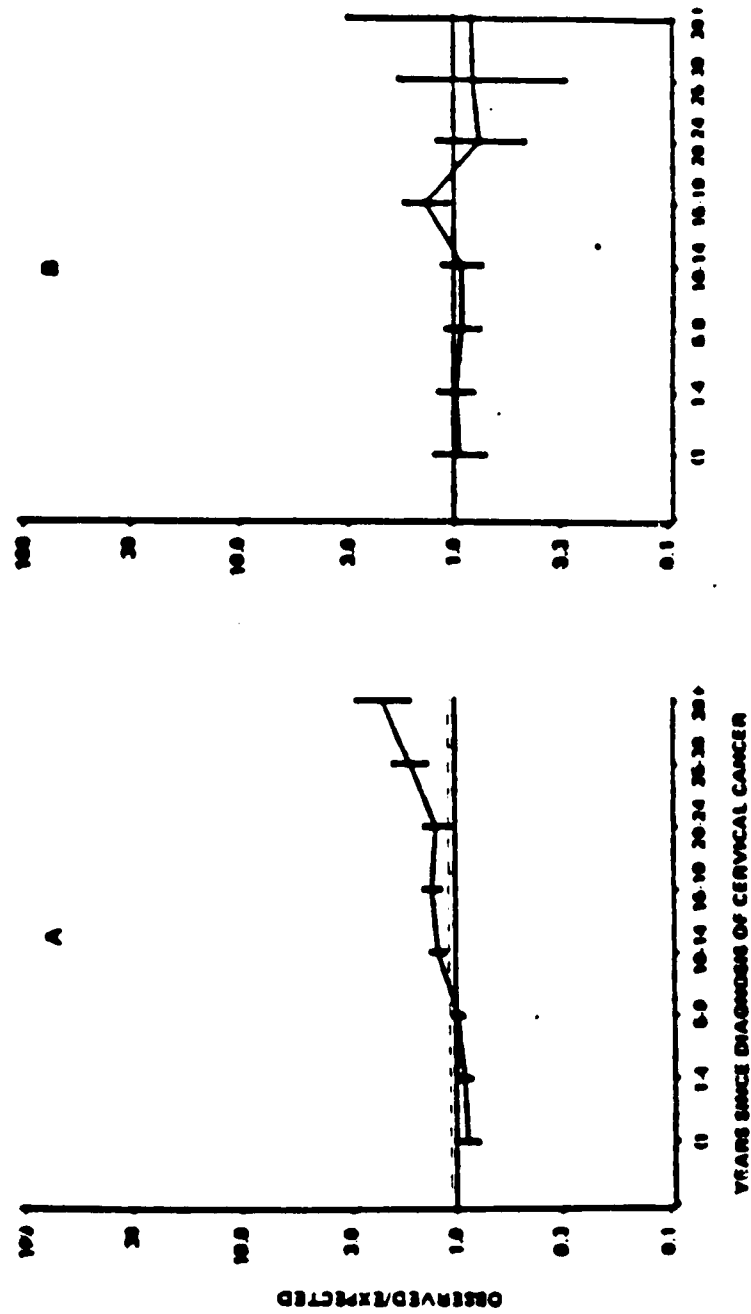


Figure 1

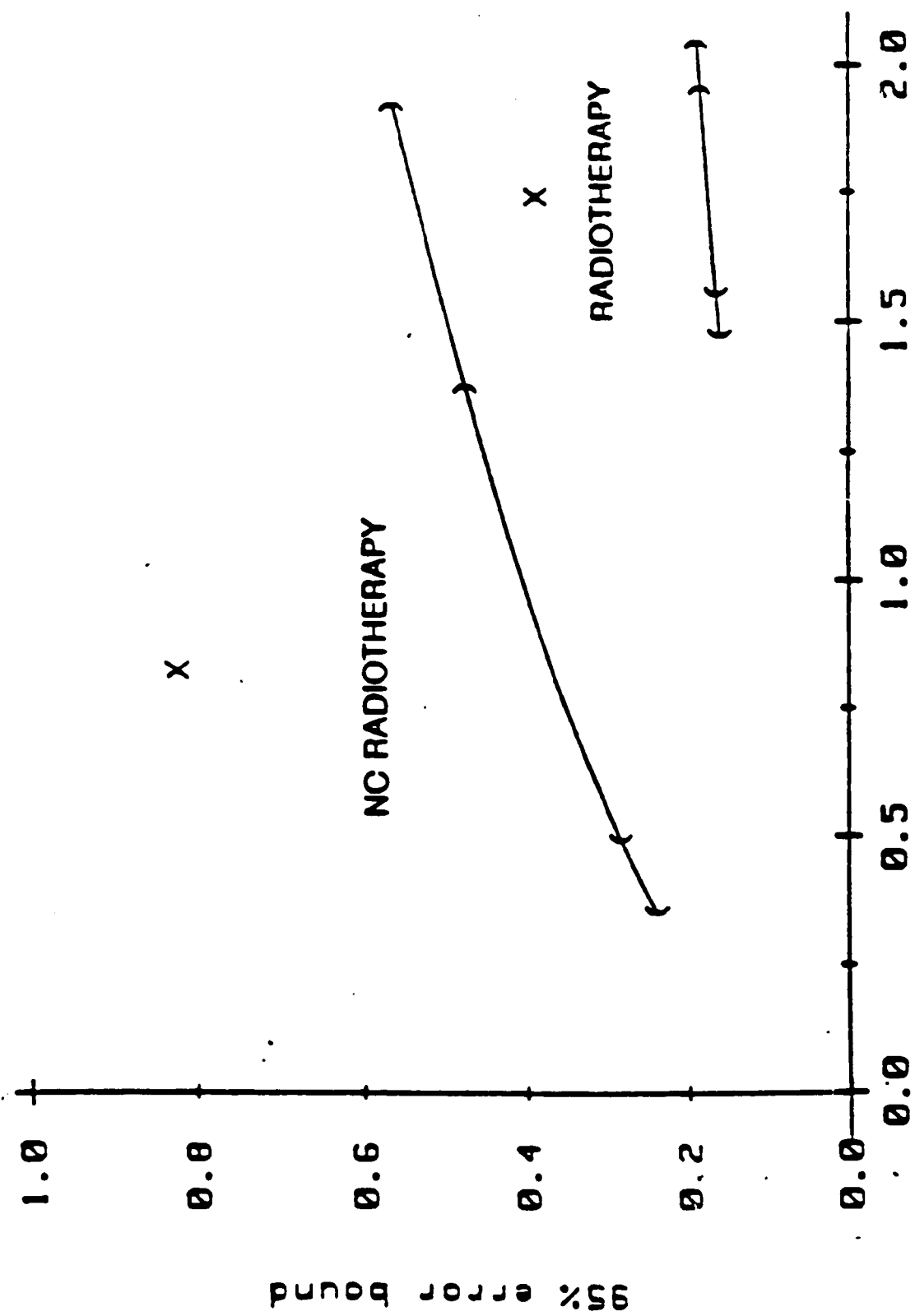


Figure 2

Observed SMR

Chapter 18

Quantification of Risk in Defined Populations

William R. Gaffey

Quantification of Risk in Defined Populations

William R. Gaffey

Introduction

In this chapter the term "risk" will be used to mean the risk of a chronic condition such as cancer associated with an exposure that will in general have occurred many years before the occurrence of the disease.

Measuring such risks in the general population, or in some subgroup such as people who live near toxic waste dumps, is extremely difficult. This is partly because migration into and out of a geographic region is difficult to keep track of, and partly because general population exposure to most substances is so low that if the exposure causes an excess risk it will probably be extremely small. Epidemiology is not completely successful in such situations (Doll 1985).

The case for studying occupationally defined populations

If one is concerned with exposure to a man-made substance, it is a truism that there exists, somewhere, a group of persons who manufactured that substance, and whose exposure in most cases was dramatically greater than that of the general population.

In addition, such employed populations are usually well defined, because a system had to exist to pay them and to make contributions to Social Security on their behalf. This means that there is a paper trail that allows us to reconstruct and identify the members of populations that were employed decades ago. Determining the subsequent fate of such a population is a task whose difficulty varies depending on the outcome variable in which one is interested (death is fairly easy to determine, for example) but it is generally possible.

As a result, it is possible to evaluate the effect of relatively heavy exposure in a well defined population, and use the results of the study to generalize, however shakily, to the lower environmental exposure of other populations.

Measuring exposure

Generally speaking, the past exposure of any individual cannot be measured directly. Investigators are usually forced to use surrogate measures such as assignment to a particular work site, or to estimate past exposures through analysis of process changes, ventilation practices and so forth. One may even find themselves settling semi-objective characterizations of exposures as "high", "medium" or "low" (Tabershaw and Gaffey, 1974).

It is worth noting that past industrial hygiene measurements, when they exist, may be suspect. Until fairly recently, industrial hygiene measurements tended to be made for compliance purposes rather than to estimate exposure levels. Measurements tended to be made at the site of leaks, for example, until values had dropped to an acceptable level.

In summary, in an occupational setting there is the opportunity to study a well defined, traceable population with generally heavy exposure that cannot easily be quantified. If the results of the study are positive, this imprecision in exposure will cause problems in extrapolating to lower exposures. At the same time, if there is a health effect from exposure, there is a better chance to uncover it.

Gaffey

-2-

Some comments on study design

The records that enable investigators to study an occupationally defined population are usually kept on the basis of the factories at which the population is employed, not by the presumptive exposures of the population. Since there are often multiple exposures, particularly in a general chemical plant, those persons with the particular exposure of concern will be a subgroup of the total work force, who can generally be identified by the fact that they were at some time or other assigned to particular departments. Unless past work histories have been computerized, identifying the subgroup of interest can be a time consuming process. It has, of course, often been done, but the task may involve sorting through thousands of records to identify a few hundred of interest, among which there may be only a handful of cases or deaths. A case-control study, unfortunately, is rarely practical because one is forced to decide in advance what particular outcome (for example, death from leukemia) is relevant.

In the absence of a computerized work history file, an attractive alternative is to use a case-control study nested within a cohort study (Marsh 1983). In such a design, a cohort study is done of all the workers at the plant, making no effort to ascertain the detailed occupational history of all the members. This is followed by a case-control study of those illnesses or causes of death that have appeared in excess, examining now the occupational histories of the small number of cases and controls. Analyses such as the Serially Additive Expected Dose model can cope with the problem of multiple exposures (Smith, Wasweiler and Tyroler 1980).

Gaffey

-3-

Interpreting studies

The International Agency for Research on Cancer (ARC) has published a set of criteria for deciding when it is justifiable to infer carcinogenicity from epidemiologic evidence (International Agency for Research on Cancer 1982). The same criteria are suitable for any other outcome variable. The criteria are, in essence, that in a study meeting certain standards of quality there should be a statistically significant excess of cancer that varies appropriately with the circumstances of exposure, and that the effect should be observed repeatedly in different independent studies.

Certainly, if two or three occupational studies met these criteria there would be little difficulty in interpretation. The problem of extrapolating to low level exposures is another matter, involving societal judgements as well as science.

However, a common occurrence is that apparently sound studies will not show a statistically significant excess that behaves appropriately, or that some such studies will show an effect and others will not. A frequent complication in interpretation is that such negative or inconclusive studies will show up in cases where animal studies show an excess of cancer. However, the relative weight to be given animal and human data is again a matter of societal judgement, and here there is concern with the treatment and evaluation of human data. The IARC criteria are of no help, since they give no advice on how to factor in the effect of negative studies.

To take a somewhat extreme example, if forty studies of the same innocuous substance were carried out and a particular cancer was observed, one would expect two of the studies to show a statistically significant dose-response relationship, at the five percent level, by chance alone. A strict

Gaffey

-4-

interpretation of the IARC criteria could lead to an inference of carcinogenicity, even though common sense dictates that we are seeing exactly what we would expect to see if there were in fact no effect. Obviously studies, if for no other reason than that they ought to enter into the estimate of the upper limit of risk that is compatible with all the data that are available.

Pooling studies

The classification of studies as "positive" and "negative" frequently creates the illusion that some studies are mutually contradictory when in fact they may not be. For example, a statistical test of the excess of a particular cancer may yield a P value of .04 in one study and .06 in a second study. At the five percent level of significance the first study is "positive" and the second is "negative", but in fact both studies provide evidence for carcinogenicity. They differ in the strength of the evidence but are obviously consistent and mutually confirmatory. To carry the reasoning a step further, if the only evidence available in this case were several "negative" studies, each with a P value of .06, intuition would reveal that these studies in toto constituted positive evidence.

The point is that, in considering the totality of studies available, each study should contribute whatever evidence it provides in terms of the actual data observed, rather than being labeled as "positive" or "negative." That labelling should be reserved to describe the totality of the evidence.

How actually to evaluate the totality of evidence is a more difficult matter. For one thing, studies of small populations that are not in themselves "positive" may never see the light of day, either because journals do not consider them worth publishing or because the authors themselves elect not to publish a study which, considered in isolation, has little significance.

If there is not a complete inventory of studies, there is however a host of possibilities for combining them, ranging from simply meditating on the array of studies to formally combining them in any of several ways (Goldsmith and Beeser 1983).

It may not be obvious to epidemiologists that they have in fact already committed themselves to one technique of pooling, which is used routinely in multi-plant cohort mortality studies. In such a study, the Standardized Mortality Ratio (SMR) for a particular cause is the total number of observed deaths in all plants divided by the total number of expected deaths in all plants. Letting O represent observed deaths and E represent expected deaths, the result for a study of two plants looks like this:

$$\text{SMR} = \frac{O_1 + O_2}{E_1 + E_2} \left(\frac{E_1}{E_1 + E_2} \right) \left(\frac{O_1}{E_1} \right) + \left(\frac{E_2}{E_1 + E_2} \right) \left(\frac{O_2}{E_2} \right)$$

In other words, in effect two studies have been pooled, one in each plant, by taking a weighted average of the individual plant SMRs, using as weights the proportion of the expected deaths that occurred in each plant.

Is this necessarily correct? Given enough assumptions, such as that the circumstances of exposure are about the same in both plants and that the relationship between age and exposure does not cause a different response in each plant, this pooling is an optimum estimate of the ratio for the force of mortality in the study population to that of the reference population, or for most practical purposes the relative risk.

If the circumstances of exposure are not so well behaved, each of the plants may represent quite different studies, whose respective statistical

Gaffey

-6-

powers should be calculated against different alternatives because of the differences in exposure at each plant. This suggests that any weighted average should in such a case take into consideration the relative powers of the two studies against their respective alternatives.

There does not appear to be an answer to this apparent dilemma, and indeed no firm convictions on how, or whether, to pool such data. It may be sufficient to recognize that the problem exists, and to deal with it on a case-by-case basis in whatever way appears to be expedient.

Summary

Epidemiologic studies of occupationally exposed populations have the advantage of dealing with relatively heavily exposed populations that are easy to follow. When such studies are positive, their interpretation is relatively straightforward, although extrapolation to low level exposures may not be. When studies are inconclusive and variable it is essential to consider the totality of available evidence. Labelling individual studies as "positive" or "negative" complicates the problem by in effect not using all of the data in each study. Although epidemiologists have in practice committed themselves to one form of data pooling in a particular situation, it is not clear that the procedure is always optimum. What is clear is that the evidence provided by so-called negative or inconclusive studies must be considered as a legitimate part of the total body of data that should be used, in one way or another, to evaluate epidemiologic associations.

Gaffey

-7-

References

- Doll, R. (1985). Relevance of epidemiology to policies for the prevention of cancer. *Human Toxicol.* 4:81-96.
- Goldsmith, J.R., and Beeser, M. (1983). Strategies for pooling data in occupational epidemiological studies. Presented at the Third Conference on Epidemiology in Occupational Health, Singapore, 28-30 September.
- International Agency for Research on Cancer (1980). Monographs On The Evaluation Of the Carcinogenic Risk of Chemicals To Tumors. Lyon: International Agency for Research on Cancer. 19:18-9.
- Marsh, G.M. (1983). Mortality among workers from a plastics producing plant: A matched case-control study nested in a retrospective cohort study. *J. Occup. Med.* 25:219-30.
- Smith, A.H., Waxweiler, R.J., and Tyroler, H.A. (1980). Epidemiologic Investigation of occupational carcinogenesis using a serially additive expected dose model. *Am. J. Epidemiol.* 112:787-97.
- Tabershaw, I.R. and Gaffey, W.R. (1974). Mortality study of workers in the manufacture of vinyl chloride and its polymers. *J. Occup. Med.* 16: 509-18.

Part V

Evaluation of Epidemiologic Information

Chapter 19

Combining Animal and Human Data: Resolving
Conflicts, Summarizing the Evidence

Linda S. Erdreich

COMBINING ANIMAL AND HUMAN DATA:
RESOLVING CONFLICTS, SUMMARIZING THE EVIDENCE

Linda S. Erdreich

Introduction

Evaluation of epidemiologic information to assess effects of chronic exposure poses problems common to evaluating any body of data in a weight of evidence approach. There may be data gaps or apparently conflicting results because of different study designs or different exposure regimens. However, the epidemiologic issues may be more difficult for those risk assessors and decision makers who are not trained in epidemiology.

Two factors contribute to the difficulties laboratory and experimental scientists have with epidemiologic data: 1) epidemiology is limited by opportunity, thus imposing limits on sample size, levels of exposure and measurement of exposures and 2) there are several types of epidemiologic data, each of which gives different information and, therefore, serves a different role. Nevertheless, it is generally recognized that a study of chronic exposure using humans is the preferable basis for risk assessment (U.S. EPA. Federal Register 45:79347-57 1980; Dourson and Stara, 1983).

The essence of risk assessment is an incomplete data base. In order to evaluate the nature and likelihood of potential risks in situations where the risks are not directly measurable, and there is little past reasoning experience, judgment and theorizing are necessary (Hattis and Smith, 1985). This reasoning and judgment may include value judgments on scientific issues. The challenge is to communicate this process to the decision makers who may then implement a decision that incorporates societal values as necessary.

This chapter is an overview of general issues in evaluating epidemiologic data for risk assessment. The objectives are to 1) illustrate issues in

integrating results of epidemiologic and experimental data including apparent conflicts among studies and "negative" results and 2) evaluate the existing framework epidemiologists use to assess causality to see where adaptations are needed to perform and communicate risk assessments.

Integrating toxicological and epidemiologic data

The data base for assessing human health risk for a given agent may contain diverse types of data: in vitro studies, pharmacokinetic studies, toxicologic and epidemiologic studies, and other human data such as clinical reports. The greater the uncertainty or paucity of human data, the greater the dependence on other sources to evaluate potential toxicity to humans. Responsible risk assessment focuses on evaluating the weight of the evidence to ascertain the likelihood that an effect observed in a sample of humans or in experimental animals will occur in the exposed population. Integration of all data for assessment of hazard is frequently complicated by apparent conflict between results of toxicological and of epidemiologic studies.

An apparent conflict occurs when an association has been observed in toxicologic but not in epidemiologic data or when an association is present in epidemiologic but not in toxicologic data. For example, because epidemiologic data often include mixed exposures, some uncertainty is introduced when epidemiologic data are evaluated for the effects of a specific agent. Some resolution may occur when laboratory data support the role of the agent in question. On the other hand, particularly for effects without cancer's long latency period, descriptive epidemiologic data, in the absence of analytic studies, will settle the uncertainty associated with even the best animal data.

When data in humans suggest an association and animal data do not, there may be some question of confounding or mixed exposures. One way to resolve this is to investigate the reason for the species difference. Inconsistencies can be clarified best by studying the mechanism of action, leading to identification of a species that does respond or by describing the metabolic reason for lack of response. For example, before the publication of the clinical case reports of the characteristic limb deformities in humans, laboratory data had been negative for thalidomide. Subsequently, a variety of congenital malformations were demonstrated by administering the drug to other species (Mattison, 1983).

The following example describes the investigation of health concerns regarding arsenic in drinking water in order to set safe arsenic levels for the U.S. population. It demonstrates this approach to integrating epidemiologic and laboratory data.

A cross-sectional study in Taiwan showed a high rate of characteristic hyperkeratotic lesions and skin cancer in individuals whose drinking water was high in arsenic (Tseng, Chu, How et al., 1968; Tseng, 1977). Potential confounding factors included the presence of ergotamine and other chemicals in the water in addition to arsenic, lack of detail on the type of arsenic, unknown water usage, low socio-economic level, different diet and other social or racial factors. These potential confounders were felt to reduce the applicability of the results to the U.S. situation. Clinical data from areas with high arsenic levels and numerous case reports demonstrated a dermatologic response to arsenic but there was a gap in relating the primary concern, skin cancer, to exposure levels (Borgono and Grecher 1972; U.S. EPA, 1980, U.S. EPA, 1984a). Epidemiologic studies undertaken in areas in the United States that had high arsenic levels in the drinking water were evaluated to see if the

Erdreich

dose-response could be further characterized in Oregon Morton, Starr, Pohl, et al. (1976) investigated the correlation between occurrence of skin cancer by census tract and arsenic contamination of the drinking water. Although levels of arsenic contamination were higher than in Taiwan, the levels in the majority of the water sources were not extremely high. Both the U.S. EPA and the authors agree these results do not contradict the reported associations (U.S. EPA, 1980, U.S. EPA, 1984a).

Two cross-sectional studies were performed in the United States but sample sizes were much smaller than in the Taiwan study (Table 1). Southwick, Western, Beck et al. (1981) looked for signs and symptoms of arsenic poisoning in 223 people recruited from these communities. Two of three communities were exposed to levels over 0.05 ppm, EPA's standard at that time, including 119 of the study subjects. Arsenic was present in hair and urine samples, but significant lesions characteristic of arsenic poisoning were absent. The ability to discern an association was reduced by small sample size, low arsenic level, and the inclusion of children in the study population. In Alaska there was the additional problem of short exposure duration because the houses served by the wells were only ten years old. Therefore, although numerous clinical reports support the link between ingested arsenic and skin cancer, attempts to determine whether effects are likely to occur in the United States population at ambient levels have not been successful.

The uncertainty regarding the causal association for a given chemical species of inorganic arsenic and the level at which it poses risk would clearly be reduced if the skin cancer could be studied in experimental animals. Despite many attempts, results are inconsistent. On the other hand, the inability to reproduce these effects in rats need not be considered an inconsistency as the rat has been found to have an anomalous mechanism for arsenic distribution.

Erdreich

-4-

Another example where there remain some doubts in the interpretation of data on humans and where experimental data could tip the scales is the health effect of asbestos in the drinking water. Humans exposed to airborne asbestos have an increased risk for some gastrointestinal (GI) tumors. Therefore, it is postulated that the health effects of asbestos in the drinking water are GI tumors, a concern because large population centers are served by water supplies containing natural or anthropogenic asbestos. Epidemiologic studies of exposed populations consist mainly of geographic correlations because it is costly to get individual data, particularly for the large population that would be required to detect effects at these low doses. Of the many sites examined in six studies summarized by Polissar, Severson, Boatman et al. (1982), only two sites, stomach and pancreas, were associated with increased risk in both males and females in more than one study. Nearly all of 13 ecologic studies in five locations had potential limitations such as absence of historic asbestos exposure data, multiple comparisons and potential uncontrolled confounding for several items (Marsh, 1983). An analysis of these studies in terms of characteristics that affect ability to detect an association including exposure duration, number exposed and exposure level (Table 2), has shown that these studies probably did not have the ability to detect the risk that would be estimated by the extrapolation model to occur at lower doses (Erdreich, 1983).

Eventually a case-control study evaluated several suspect sites and found no association with exposure to ingested asbestos (Polissar, Severson and Boatman, 1984). Exposure to asbestos was based on residence, workplace history, and individual water supply. The minimum relative risk detectable was less than 2 for digestive system, colon and all sites combined. In this study, the exposure for cases and controls was somewhat similar. The author concluded that the long latency period and the reasonable belief that risks are low

Erdreich

reduce the ability to resolve this issue by epidemiologic methods. However, in 12 toxicologic studies on two animal species, chronic high level ingestion of asbestos fibers produced no definite, reproducible organ-specific carcinogenic effects. In one study, benign epithelial tumors were observed (NTP, 1984 as reported by U.S. EPA, 1985a). Nevertheless, after considering the occupational data and some flaws in the animal studies, both the U.S. EPA and National Research Council (NRC) (1983) have judged that an association does exist and have estimated the risk from ingested asbestos by extrapolating the results of cohort studies of occupational populations exposed by inhalation, based on the assumption that inhaled fibers are swallowed and on evidence that they cross the GI wall. The presence of a plausible animal model for such tumors from ingestion of asbestos fibers in the form in which they occur in drinking water, would have given weight to the argument for an association, and could provide direction for identifying likely cancer sites for site-specific case control studies. One such study became available by 1985 (U.S.A. EPA 1984a). The body of epidemiologic evidence could reasonably have been expected to show risks if they were elevated substantially above background. Therefore, these studies should be considered evidence that risks have not been underestimated.

Resolving apparent conflicts in epidemiologic data

The previous examples have shown that conflict may appear in reported results of epidemiologic studies when there are several studies performed under different exposure scenarios for different durations and follow-up times, and examining various endpoints. This also occurs with experimental data when several stages of toxicologic studies using different dose regimens have been undertaken. Studies of acute toxicity report dose levels that cause 50% mortality (LD_{50} s) and define the general toxicity, may be followed by 30-day or 90-day subchronic studies to identify a sensitive endpoint, and finally by

Erdreich

-6-

lifetime studies to characterize a safe or effective dose. In epidemiology, the different types of study designs used impose added difficulty.

The first steps taken when results appear to conflict are to scrutinize data for systematic bias or uncontrolled confounders or design factors. In certain cases the results of a study must be viewed as less important. For example, studies showing no association which did not include a long enough interval to cover the probable latent period for carcinogenesis, cannot be construed as conflicting with studies with complete follow-up. On the other hand, studies that show an association, and are found to have a selection bias toward increased ascertainment or classification of those with both the disease and the exposure, are clearly less valid. However, as Stolley (1985) has stated, one must be careful not to confuse a proposed bias with a bias sufficiently strong to invalidate the results. The bias must be weighed against the strength of the association. The proposed bias of concern to those adapting the results of arsenic exposure in rural, impoverished Chinese in Taiwan to the general U.S. population was that the Chinese population may have been more sensitive to the agent.

When results of analytic epidemiologic studies appear to conflict, it is important to consider all the evidence. Certain types of studies that never stand alone to identify a hazard can support a casual association in the context of a larger body of data. Observational data may be useful, particularly for effects with short latency, (acute or subchronic effects of toxic substances). Many case clusters and episodes of accidental poisonings are initially investigated by an epidemiologic method similar to the infectious disease model. These observations suggest questions and directions for laboratory studies. Integration of human and experimental data can potentially provide for more efficient surveillance and treatment as well as risk assessment. For example, an epidemic of porphyria cutanea tarda occurred in Erdreich

Turkish citizens who consumed grain treated with hexachlorobenzene (HCB). Follow-up studies included clinical descriptions, biological monitoring of plasma levels, and analysis of body fat and breast milk. Subsequent studies in laboratory animals have elucidated the mechanism for the induced porphyria, which explained the continued enzyme inhibition and chronic effects occurring after elimination of HCB. An unexplained high frequency of enlarged thyroids occurred in the exposed cohort. Eventually, exposure of laboratory animals was shown to result in thyroid tumors (U.S. EPA, 1985b).

Several of these scenarios imply that repeated similar results are more valuable than a single study. Doll (1984) has stated that properly conducted, relevant, unbiased studies of small sample size can be pooled, combining both positive and negative results. Recently IARC (1982) has pooled the total number of deaths from a specific endpoint, soft-tissue sarcoma, from seven occupational cohorts exposed to herbicides and manufacturing processes to shed light on the carcinogenicity of a suspected contaminant, tetrachlorodibenzo--dioxin.

Use of studies having "negative" results

Most agents that come to the attention of Federal Agencies (other than FDA) for risk assessment have no therapeutic value, and have no observable practical use to the ordinary citizen. Therefore, evaluation of epidemiologic studies that show no association with the agent becomes problematic when the available toxicologic data demonstrate an adverse effect. It should be recognized that this is actually an issue of weighing two sets of uncertainties; the uncertainty on one hand as to the relevance of animal toxicity to the human case and on the other, the possible low power of certain epidemiologic studies. One approach to evaluating the uncertainty is to examine the confidence limits of the response that are generally presented

Erdreich

together with the epidemiologic findings. Consider the guidelines for carcinogenic risk assessment proposed by the U.S. EPA (Federal Register 49: 5294-4630) which state the following:

"negative results from a well-designed and well-conducted epidemiologic study that contains useful exposure data can serve to define upper limits of risk which are useful if animal evidence indicates that the agent is potentially carcinogenic."

Thus, when the inference is made from animal data that an agent is a carcinogen, the human data are used as a check of the risk level. One important consideration of this approach is that smaller sample sizes will give larger confidence limits and higher levels for the upper limits of risk. This can be construed as a disadvantage or as an incentive to do studies with larger sample sizes whenever possible in order to generate smaller confidence intervals.

framework for evaluating the data base

Even before the existence of federal regulations requiring assessment of potential risks, the frequency of disease and the presence of data that suggested a hazard created a need for a decision to protect the public health. The approach has always been to judge the weight of the existing evidence and that has always included the integration of toxicologic, laboratory and epidemiologic data. The criteria for this process have evolved as the focus changed from infectious to chronic disease. Koch's postulates, the first criteria for establishing causality, were modified by Yerushalmy and Palmer (1959) to be more applicable to chronic disease. Lilienfeld (1959) indicated that elucidation of the mechanism by means of which the etiological factor produces the disease strengthens the plausibility of the association and weakens the idea of a confounding factor as causative.

Erdreich

-9-

Table 3 shows a revision by Evans (1976) that includes a broader definition of causative agents and host responses. This list was designed for establishing causality and does not specifically address dose-response relationships, nor does it suggest a minimum data base, or how to combine results of different studies. When Evans presented these criteria to build upon those concepts previously summarized, he specifically stated that new technology requires change. His list of criteria includes "experimental reproduction of the disease related to exposure", thus including toxicology. Other new technology could be covered in the important but vague last postulate "the whole thing should make biologic and epidemiologic sense."

A few emerging areas in risk assessment are not specifically included in these postulates. These include short-term in vivo or in vitro assays, pharmacokinetic comparisons among species, and biochemical or molecular markers of exposure or early indicators of disease. The latter may be covered by two of Evans' postulates -- one stating that a spectrum of host responses should follow exposure and the other that a measurable host response should regularly appear or should increase in magnitude following exposure. Under current knowledge, the spectrum of host responses is more readily identified for chronic toxicity effects on the lung or liver than it is for cancer, where molecular and biochemical endpoints constitute emerging technology.

These criteria for causation are a summary of the methods and reasoning involved in evaluating data. The underlying and more general concepts proposed by Hill (1971) were used by the Surgeon General in the Report on Smoking and Health (1964). Epidemiologists know these well, but I reiterate them briefly because these general precepts work well for evaluating studies from different disciplines (Table 4). Only one of these, the lack of temporal ambiguity, is largely unnecessary for evaluating nonepidemiologic data.

Erdreich

-10-

A framework for integrating data from diverse fields

The concepts underlying the evaluation of data - strength, consistency, biologic plausibility - are common to all research fields. Therefore, the risk assessor could use these concepts to summarize the relevant data and to communicate this information to risk managers, decision-makers, the press or the public as circumstances demand. There is no doubt that these concepts are being utilized by the scientists who evaluate the data. However, it is important for risk assessors to communicate their critique and evaluation. Table 5 is a preliminary attempt to merge the concepts establishing causality with the types of data available for predicting health effects. It should serve only as a reminder of what questions to ask, particularly when faced with a data base that has more laboratory than human data.

This broad framework should stand up particularly well for the kind of risk assessment public agencies have to do, which is to evaluate many agents, perhaps hundreds in a year, with such a variety of data available that no two data bases are alike. This framework has the following characteristics: 1) it is not too binding as it does not ascribe weights or value to the data types, 2) it is not limited to a specific disease endpoint, and 3) it is flexible. No inherent value judgment in interpreting or weighing the items listed is indicated, but it is not possible to escape the values inherent in deciding what items to list under the concepts. For example, specificity is a useful concept, but less necessary for toxicity data when known confounders are under experimental control. The proposed approach could serve as a vehicle for communication between investigators in various fields and decision-makers who use this information. Revisions and additions are necessary and are freely invited.

Erdreich

-11-

Inferences

In addition to the judgments involved in evaluating data, predicting risk entails making inferences (NRC, 1983). One well-recognized inference is that toxicity in animals suggests toxicity in humans. IARC (1982) recognized this for carcinogens as does the U.S. EPA (Federal Register 49:46294). The U.S. EPA also recognizes this for noncarcinogens (Federal Register 45:79318) and for developmental toxicants (Federal Register 49:46324). It may help to examine an inference that is currently being made frequently, but unofficially and of necessity, regarding epidemiologic data. Another frequent inference is that results from studies of workers or other subgroups of the population on the nature of a hazard can be generalized from these workers to all workers or to the larger population. While epidemiologists and risk assessors recognize some uncertainty in doing this and measure some of it, generalization is nevertheless often done. However, much of the epidemiologic work on risk factors and newer work on genetic differences and sensitive subgroups demonstrates the variability and attendant uncertainties in this process of generalization which need further study.

Conclusions

The challenge in risk assessment is to critically analyze and reconcile data from diverse fields. It is crucial to recognize that inferences from experimental data in animals to humans are far from validated and that epidemiologic data on many agents are needed to characterize this relationship

Erdreich

-12-

and to validate or refute it for specific organ systems, effects or chemical classes. People are reluctant to be exposed to nontherapeutic, non-utilitarian chemicals that are toxic to animals. Therefore, the bridges must be built by evaluating short-term test procedures and bioassays, by using to their fullest all types of human data and by developing and compiling background rates for trend and time analyses. The case studies presented in this section solve specific problems in evaluating risks to humans. They also help to build these bridges.

REFERENCES

- Borgono and Grecher. (1972) Epidemiological study of arsenicism in the city of Antofagasta. In: Trace Substances in Environmental Health V, Proceeding of the 5th Annual Conference, Univ. of Minnesota, 1971, D.C. Hemphill, Ed. Univ. Missouri, Columbia, MO. pp. 13-24.
- Doll, R. (1984) Occupational cancer: Problems in interpreting human evidence. Am. Occup. Hygiene 28:291-305.
- Dourson, M., and Stara, J.F. (1983) Regulatory history and experimental support of uncertainty (safety) factors. Regul. Toxicol. Pharmacol. 3:224-38.
- Erdreich, L.S. (1983) Comparing epidemiologic studies of ingested asbestos for use in risk assessment. Environ. Health Perspect. 53:99-104.
- Evans, A.S. (1976) Causation and disease: The Henle-Koch postulates revisited. Yale J. Biol. Med. 49:175-95.
- Harrington, J.M., Middaugh, J.P., Morse, D.L. and Housewoth, J. (1978) A survey of a population exposed to high concentrations of arsenic in well water in Fairbanks, AK. Am. J. Epidemiol. 108:337-85.
- Hattis, D. and Smith, J.A. (1985) What's wrong with quantitative risk assessment. M.I.T. Center for Policy Alternatives presented at Georgia State University September 26-28, 1985. (Draft)
- Hill, A.B. (1971) Principles of Medical Statistics, 9th ed. Oxford University Press, New York.
- IARC 9 International Agency for Research on Cancer. (1982) IARC monographs on the evaluation of the carcinogenic risk of chemicals to humans. Supplement 4 IARC, Lyon, France.
- Lilienfeld, A.M. (1959) "On the methodology of investigations of etiologic factors in chronic diseases" - some comments. J. Chronic Dis. 10:41-6.
- Erdreich

- Marsh, G. (1983) Critical review of published epidemiologic studies of ingested asbestos. *Environ. Health Perspect.* 53:49-56.
- Attison, D.R., (ed.) (1983) *Reproductive Toxicology*. Alan R. Liss, Inc., New York.
- Morton, W., Starr, G., Pohl, D., Stoner, J., Wagner, S. and Weswig, P. (1976) Skin cancer and water arsenic in Lane County, OR. *Cancer*. 37:2523-32.
- National Research Council (NRC). (1983) *Risk Assessment in the Federal Government: Managing the Process*. National Academy Press, Washington, DC.
- Polissar, L., R.K. Severson, E.S. Boatman and D.B. Thomas. (1982) Cancer incidence in relation to asbestos in drinking water in the Puget Sound Region. *Am. J. Epidemiol.* 116:314-28.
- Polissar, L., Severson, R.K. and Boatman, E.S. (1984) A case control study of asbestos in drinking water and cancer risk. *Am. J. Epidemiol.* 119:456-71.
- Southwick, J.W., Western, A.E., Beck, M.M. et al. (1981) Community health associated with arsenic in drinking water in Millard, UT. *Health Effects Res. Lab., U.S. EPA, Cincinnati, OH.* EPA 600/1-81-064, NTIS PB 82-108374, 1984.
- Stolley, P. (1985) Faith, evidence and the epidemiologist. *J. Public Health Policy*, March 1985, pp. 37-42.
- Tseng, P. (1977) Effects and dose-response relationships of skin cancer and Blackfoot Disease with arsenic. *Environ. Health Perspect.* 19:109-19.
- Tseng, W.P., Chu, H.M., How, S.W. et al. (1968) Prevalence of skin cancer in an endemic area of chronic arsenicalism in Taiwan. *J. Natl. Cancer Inst.* 40:453-63.
- U.S. EPA. (1980) *Ambient Water Quality Criteria Document for Arsenic*. Environmental Criteria and Assessment Office, Cincinnati, OH. EPA 440/5-80-021.
- Erdreich

- U.S. EPA. (1980) Guidelines and methodology used in the preparation of health assessment chapters of the consent decree water quality criteria. Federal Register 45:79347-57.
- U.S. EPA. (1984a) Health Assessment Document for Inorganic Arsenic. EPA 600/8-83-021F.
- U.S. EPA. (1984b) Proposed guidelines for the health assessment of suspect developmental toxicants and request for comments. Federal Register. 49: 46324.
- U.S. EPA. (1984c) Proposed guidelines for carcinogen risk assessment; request for comments. Federal Register. 49:46294-301.
- U.S. EPA. (1985a) Drinking Water Criteria Document for Asbestos, (final draft) Environmental Criteria and Assessment Office, Cincinnati, OH. EPA 600/x-84-199-1.
- U.S. EPA. (1985b) Health Assessment Document for Chlorinated Benzenes. Environmental Criteria and Assessment Office, Cincinnati, OH. EPA 600/8-84-015F.
- Yerushalmy, Y. and Palmer, C.E. (1959) On the methodology of investigations of ecologic factors in chronic diseases. J. Chronic Dis. 10:27-40.

TABLE 1: Prevalence of health effects of ingested arsenic

Author and location	Arsenic Level	Number Examined	Response
Tseng, Chu, How et al., 1968 (Taiwan)	0.3-0.6 ppm (300-600 mcg/)	40,421 (all residents)	Keratosis - 7.0% Skin cancer - 10.
Southwick, Western, Beck et al., 1981 (Utah)	mean 0.2 ppm (min 0.15 ppm)	232 119 >0.05 ppm (lifetime residents)	No excess dermato symptoms
Harrington, Middaugh, Morse et al., 1978 (Alaska)	0.03-0.4 ppm (mean home levels)	211 (91% of exposed residents)	No increased skin abnormalities

TABLE 2: Characteristics of the epidemiologic studies that affect the power to detect an association

Study Area	Duration of Exposure (year)	Population Exposed	Probable Exposure Level
Duluth, MN	16-17	100,000	1-30 MFL
Connecticut	19	576,000	<0.1-0.7 MFL
Bay Area, CA	>40	1,000,000	4-36 MFL
Puget Sound, WA	>50	78,000	200 MFL
Florida	>25	46,000	<1-10 MFL (AC PIPE)
	>25	87,000	
Quebec	>50	140,000	80-220 MFL

Source: Erdreich, (1983)

TABLE 3: Criteria for causation: A unified concept

-
1. Prevalence of the disease should be significantly higher in those exposed to the putative cause than in case controls not so exposed.
 2. Exposure to the putative cause should be present more commonly in those with the disease than in controls without the disease when all risk factors are held constant.
 3. Incidence of the disease should be significantly higher in those exposed to the putative cause than in those not so exposed as shown in prospective studies.
 4. Temporally, the disease should follow exposure to the putative agent with a distribution of incubation periods on a bell shaped curve.
 5. A spectrum of host responses should follow exposure to the putative agent along a logical biologic gradient from mild to severe.
 6. A measurable host response following exposure to the putative cause should regularly appear in those lacking this before exposure (i.e., antibody, cancer cells) or should increase in magnitude if present before exposure; this pattern should not occur in persons so exposed.
 7. Experimental reproduction of the disease should occur in higher incidence in animals or man appropriately exposed to the putative cause than in those not so exposed; this exposure may be deliberate in volunteers, experimentally induced in the laboratory, or demonstrated in a controlled regulation of natural exposure.
 8. Elimination or modification of the putative cause or of the vector carrying it should decrease the incidence of the disease (control of polluted water or smoke or removal of the specific agent).
 9. Prevention or modification of the host's response on exposure to the putative cause should decrease or eliminate the disease (immunization, drug to lower cholesterol, specific lymphocyte transfer factor in cancer).
 10. The whole thing should make biologic and epidemiologic sense.
-

Source: Evans, (1976)

Endreich

-19-

TABLE 4: Criteria for a causal interpretation

Strength of the association

Dose-response effect

Lack of temporal ambiguity

Consistency of the findings

Biologic plausibility of the hypothesis

Specificity of the association

Coherence of the evidence

Source: Hill, (1971)

TABLE 5: Proposed approach for summarizing diverse data

Concept 1 — Strength of the Association:

The stronger the association the more confident we are that the agent causes the effect

- . Presence of low LD₅₀. low observed no-effect-level, high potency
- . Dose-response gradient evident
- . High incidence rate, large excess risk
- . High level of statistical significance in relevant studies

Concept 2 — Consistency:

The association persists upon testing the observation in various circumstances

- . Number of experimental species
- . Various routes
- . Different dose regimens
- . Descriptive epidemiologic data
- . Analytic epidemiologic studies

Concept 3 — Biological Plausibility:

The association is plausible in terms of other scientific information related to the causal mechanism

- . A gradient of responses observed
- . Short-term or in vitro tests
- . Pharmacokinetics
- . Pathology
- . Structure activity relationship
- . Preclinical indicators
- . Biological monitoring of exposure

Chapter 20
Saccharin and Bladder Cancer
Philip Cole

SACCHARIN AND BLADDER CANCER

Philip Cole

Risk assessment is basically the description of the change in the frequency of a disease due to a specified change in the frequency of an exposure. It is not a cost-benefit analysis and it is not an exercise in public policy or in regulation formation which rest upon the cost-benefit analysis as well as upon certain human issues.

This chapter will begin with some background on the modern history of the relationship between saccharin and bladder cancer. Although this issue had been studied back as long ago as the 1930's, it is fair to say that our concern now is with the information which has developed in the last seven or eight years. A chronology is presented in Table 1.

The thrust began in 1977 with an animal study from Canada which reported an association. This study had two characteristics: 1) the animals had been exposed in utero by the feeding of saccharin to their mothers; and 2) exposure continued throughout their lives. The findings were positive and statistically significant for male rats. It has often been said that the study was negative for female rats but that was not the case; the study was, in fact, positive, but was not statistically significant. At that time, Dr. Donald Kennedy, then commissioner of the FDA, proposed a restriction on saccharin. Specifically, he proposed that saccharin only be permitted for sale as a single ingredient, so that, in effect, it would only be sold as tablets and powdered saccharin. Its use in various mixes and in foods would not be permitted.

Shortly thereafter, in 1977, a positive epidemiologic study appeared from Canada. The findings were positive for men and the study was said to be analogous to the rat study which had also been positive for males. However, unlike the rat study, the findings from this first human study from Canada were, actually negative for women. Among women, saccharin users had a lower

risk of bladder cancer than did nonusers of saccharin. The study was rather severely criticized and, in retrospect, it is safe to say it has not been given high credibility by the scientific community. But, in 1977 it was not so viewed, and on the basis of that human study Kennedy proposed a ban rather than a restriction -- that is, saccharin would be totally banned from the United States market.

Meanwhile, a series of literature reviews appeared. There were assessments of the older information from toxicology and epidemiology. A moderate number of studies, about ten, had been done. True, they all suffered from limitations in their design or in the information with which they had to work. Some had more than just limitations -- they had errors as well. Nonetheless, their findings were negative and gave no support to the idea that saccharin was a human carcinogen.

By late 1977 a National Cancer Institute study of 3000 people with bladder cancer and 9000 controls throughout the United States was initiated. At that time, the ongoing NCI study provided a basis for a compromise between Congress and the FDA: There would not be a ban but there would be a required labeling of saccharin-containing products indicating that the substance had been found to be carcinogenic in laboratory animals. With this labeling concession to the FDA, Congress imposed a moratorium on the ban, permitting the continued use of saccharin.

In 1979 the NCI study was published. It appeared to be largely negative except for small positive results in subgroups, people who were otherwise at very low risk of bladder cancer and people who were at very high risk. In 1982 other studies appeared. These studies could again be described as largely

negative. While some are subject to criticism because of methodologic and other problems, by and large, the widespread impression was that either saccharin was not a human carcinogen or, if it was, it was carcinogenic at such a low level that it probably would not be possible in the foreseeable future for epidemiologic studies to resolve the issue.

The International Research and Development Corporation published a large study in 1984 supporting the first Canadian rat study to the effect that indeed saccharin was a bladder carcinogen in rats when fed at very high doses and in utero but there was no excess of bladder cancer at low dose. Was this because the sample size was too small, or was it because saccharin is not a carcinogen at low dose? The answer is not known.

Recently the fourth moratorium on the ban was passed in the United States and the ban is now in its seventh year. At the same time saccharin is used today in ninety six countries.

What are some of the issues which emerge from the saccharin-bladder cancer chronology and from which lessons of a general nature for risk assessment may be drawn? (Table 2)

- 1) Ascertainment of exposure is very difficult. Other non-nutritive sweeteners were in use at the same time and people did not keep track of what they were using. These substances were used in different dosages in different kinds of preparations, whether tablets or powders. These substances were often mixed in foods and consumed by people without their knowledge.

- 2) The null or near null state (NNN) prevails, rather than the situation of an individual negative study. The criteria for evaluating an individual negative study, while not uniformly agreed upon, are nonetheless laid out reasonably well and are used. The circumstance that now prevails with saccharin is that there are a relatively large number of studies, virtually all of which cluster around the null state. Some are weakly positive, some are weakly negative and some are right on the null.

When we turn from the issue of risk assessment itself to that of making a cost-benefit analysis in public policy issues, the question of alternatives or substitutes becomes an issue. This is followed by public concerns about the right of the individual to do as he wishes as long as he is adequately informed of the existing information. All of these issues come to the fore and all of them did indeed come to the fore in the question of saccharin and bladder cancer. In 1977 when this was a very heated topic, there was no substitute available in the United States. Cyclamate had already been banned. Now, of course, there is aspartame.

I should like to focus on the question of dealing with the null or the near-null state. Table 3 is a list of some of the possible explanations of why the null or near-null state characterizes the data sets regarding saccharin and bladder cancer. First, saccharin may be a very weak carcinogen and may hardly ever cause human cancer. Second, it may not be a complete carcinogen, that is, its effect may depend crucially on an interaction with some other substance or behaviors such as cigarette smoking. Finally, the third explanation, of all

on this list, may be the most compatible with the data. However, other kinds of biological explanations are also tenable. For example, some studies included observations which did not encompass the full length possible of the induction period. There are also some epidemiologic problems in precision due to the high frequency of exposure, since epidemiologic studies are usually most powerful for their size when the prevalence of exposure in the population approximates 50 percent. For saccharin, the exposure has been almost 70 or 80 percent making it very difficult to obtain precise results. There is also inordinate random misclassification due to the difficulties of exposure ascertainment previously mentioned and other epidemiological problems of a general nature that characterize epidemiologic studies.

Two important questions arise: 1) How is science to deal with the null and near null state? and 2) How is regulatory practice to deal with it? The following deals only with the first question, which is a legitimate question that has received inadequate attention by scientists. How do scientists deal with the circumstance of substances that are weak carcinogens or not carcinogens at all falling under severe suspicion? Four perspectives are offered on this question.

Perspective 1: (Table 4) We must realize that scientists by their own attitudes are much more interested in positive results. They are more likely to receive rewards in terms of publication, promotion and/or grants, so they pursue studies which are likely to be positive. Laboratory scientists in particular develop techniques for enhancing the production of positive results (for example, protecting animals so that they can tolerate higher doses and exposing them in utero). Some professional media favor the publication of positive results. Recently there was a personal encounter with a journal that proclaimed it simply would not publish a negative study because it was

Cole

negative. The mass media, of course, are very interested in positive results and hardly interested in negative results. Imagine calling a newspaper and telling the editors that you are about to announce a study that you have completed showing that eating three bananas a day is strongly related to brain cancer. You'd get a tremendous reception. Now call them and tell them that eating three bananas a day doesn't appear to have any effect on health at all. Supremely disinterested.

Regulators and policy makers are also more interested in positive results. Obviously regulators are in the business of making regulations and policy makers have to make policy. These activities cannot really be done with negative results. The situation is very akin to the statement that when you put a hammer in a man's hand everything starts to look like a nail to him. Despite the fact that everybody is interested in positive results, the public is not. They're tired of them. They've heard one too many jokes about the rat that had to drink 180 cans of diet soda.

Perspective 2: (Table 5) There is an actual asymmetry in the logic of science that favors the production of positive results. First, there is an infinite set of positive outcomes but there is only one null value. As a consequence, negative studies have to be very large to be persuasive. That is, the confidence limits have to be tight around the null value if meaningful alternatives to the null are going to be reasonably well excluded. However, in regard to the whole issue of statistical significance of negative results, a positive result is required before the issue of statistical significance can even be raised. That is, statistical significance deals with the question: Is this departure from the null of a statistically significant extent? Therefore, if there is no departure or only a trivial one, the issue of statistical significance cannot even be raised.

Cole

-6-

Perspective 3: Then we have this card. You will hear it so often. You can prove a positive but not a negative. But it is not true — you cannot prove a positive. (Table 7) In the philosophy of science neither situation — positive or negative — is ultimately provable. A scientist's perception of a cause-effect relationship exists on a spectrum. It is totally unacceptable near one limit and it is totally acceptable near the other but the scientist never adopts the limits of the spectrum. His mind is open and he will move along the spectrum but never adopt the limit. So, philosophically, you cannot prove a positive and you cannot prove a negative. More realistically, the fact is that one can marshal evidence in support of a non-null position.

Finally, perspective 4: Study size considerations. All of us, either as teachers or students, have dealt with the issue of study size. (Table 8) In the spirit of protecting our study against missing a positive result, we ask if the state of nature is positive, and of such a magnitude, how large a study is needed to show it? Consider the alternative, which is very rarely discussed, and only in advanced statistical courses: If the state of nature is negative, how large does the study have to be to be persuasive? Furthermore, when we are dealing with the issue of study size for a study that may very well be negative, where the state of nature is negative, what are the criteria? One suggestion is that it be the upper limit of the confidence range. That is, one states, if nature is truly null, I want to do a study so that the upper limit of the confidence range would be X and X is specific, say 1.3, 1.5 for a certain relative risk. The fact of the matter is that the scientific literature has not dealt with the question of how to design a study that has a maximum probability of coming up with a null or near null result.

These aspects of dealing with a null state are indeed very important, and have major implications for risk assessment. They are, in fact, crucial for valid estimations of risk.

Cole

-7-

TABLE 1: Recent history of saccharin and bladder cancer

1977a	Positive in utero male study
1977b	Saccharin restriction proposed
1977c	Positive epidemiologic study for men
1977d	Proposed restriction augmented to a ban
1978a	Literature reviews provide no support
1978b	Public resistance
1978c	NCI study commissioned
1978d	Labelling plus moratorium on ban
1979	NCI study negative but for subgroups
1981	West Yorkshire study positive but dubious
1982a	Harvard study negative
1982b	Danish follow-up study negative
1982c	Danish case-control study negative
1984	IRDC rat study confirms high dose effect
1985	Fourth moratorium on ban
1985	Saccharin used in 96 countries

TABLE 2: Features of special interest regarding saccharin and bladder cancer which are relevant to risk assessment

1. Severe difficulties in exposure ascertainment
2. Null or near null state (NNN) prevails
3. Absence/presence of a substitute
4. Topical, controversial and persistent

TABLE 3: Possible explanation of NNN state

1. Complete but weak carcinogen
2. Not a complete carcinogen
3. Not a carcinogen at all
4. Other biological explanation
5. Imprecision due to commonality of exposure
6. Inordinate random misclassification
7. Other epidemiological explanation

TABLE 4: Preference for positive results

1. By scientists
Attitudes, rewards, techniques
2. By professional media
3. By mass media
4. By regulators and policy makers
5. But not by a disdainful public

TABLE 5: Actual asymmetry of statistical inference

1. Infinite set of positive outcomes
2. Only one negative outcome
Negative study must be large to be
persuasive
3. Unavailability of p-value to support NN
state
4. Difficulties with after-the-fact use of
power statements

Chapter 21
Estrogens and Breast Cancer
Barbara S. Hulka

Breast Cancer and Exogenous Estrogens
Barbara S. Hulka

Introduction

To what extent do women increase their risk of developing breast cancer by taking estrogens after the menopause? This chapter will attempt to answer this question by summarizing the results from recent epidemiologic studies and by reviewing the methodologic features which could influence the results. Thereafter, the findings from epidemiologic studies will be placed into the context of data and theories derived from other disciplines, and finally, some judgement will be exercised to form a conclusion.

Summary of epidemiologic studies

Epidemiologic studies on replacement estrogens in relation to breast cancer risk first appeared in 1974 (Craig, Comstock and Geiser, 1974; Henderson, Powell, Rosario et al. 1974). During the next five years a number of case-control and follow-up studies were published. These are summarized in Tables 1 and 2, respectively. The relative risk estimates reported from the case control studies are consistently unimpressive. The one study reporting an elevated relative risk of 3.1 had insufficient numbers of subjects to achieve statistical significance (Casagrande, Gerkins, Henderson et al. 1976). The study by Sartwell, Arthes and Tonascia (1977), was designed specifically to evaluate the effect of exogenous estrogens and oral contraceptives on breast cancer risks. Exogenous estrogens incurred no excess risk for any duration of use up to five years or more. These studies published prior to 1980 have been reviewed previously (Hulka, 1980).

The follow-up study by Hoover, Gray, Cole et al. (1976), noted in Table 2, was highly publicized, being the first epidemiologic study to incriminate

exogenous estrogens as risk factors for breast cancer. Among 1891 women who had received conjugated estrogens for the menopause in a single practice in Louisville, Kentucky, forty nine breast cancer cases were diagnosed during an average follow-up period of twelve years, whereas thirty nine cases would have been expected according to the experience of the southern centers in the Second and Third National Cancer Surveys.

Findings from an additional follow-up study appeared in 1983 (Gambrell, Maier and Sanders). Five thousand five hundred sixty three postmenopausal women seen at a U.S. Air Force Medical Center between 1975 and 1981 were treated with estrogens, estrogens plus progestins, no hormonal therapy or other hormonal regimens. After 37,236 patient years of observation, fifty three breast cancers developed. The incidence rate in estrogen users was lower than that among untreated women and no higher than the expected incidence rate based on Third National Cancer Survey age specific rates. However, the source population, research design and analytic methods were either insufficiently specified or inadequate for the purposes of the study, which raise concerns about bias and the validity of the study results.

Table 3 summarizes those case control studies published since 1980 in which exogenous estrogens were a primary focus of the research and analysis. The number of subjects available for analysis is noted and, where possible, the results from postmenopausal women only appear since premenopausal women are not generally at risk of receiving replacement estrogens. It is also questionable to presume that exogenous estrogens will have a similar effect on breast tissue of postmenopausal women with low endogenous estrogen production and no endogenous progesterone as compared to premenopausal women producing ovarian estrogen and progesterone.

Three of the eight studies reported an adverse effect of estrogens on breast cancer risk. The first of these, the study by Ross, Paganini-Hill, Gerkins et al. (1980), showed no ill effects for every use of estrogens among the total study group, but with a high total milligram-dose (the equivalent of 1.25 mg of conjugated estrogens daily for three years), the relative risk estimate reached 2.5 for women who had undergone a natural menopause. The same total milligram dose produced a relative risk of 0.7 for women having undergone bilateral oophorectomy.

Brinton, Hoover, Szklo et al. (1981) found an effect of increasing risk with increasing estrogen dose but unlike Ross, Pagani-Hill, Gerkins et al. (1980), the increase was evident primarily in women who had previously undergone a surgical menopause (bilateral oophorectomy). Hoover, Glass, Finkle et al. (1981), noted a dose effect on breast cancer risk for pre and post menopausal women combined; data were not presented separately by menopausal status. In general, doses of 1.25 mg or more of conjugated estrogens were required before increased relative risk estimates were observed. Even in these "positive" studies, duration of use effects, unrelated to dose, were not observed.

The Hiatt, Bawol, Friedman et al. (1984) study used records of women who had had bilateral oophorectomy through the Northern California Kaiser Foundation Health Plan hospitals between 1953 and 1979. From among this group, the 119 women who developed breast cancer were pair matched on birth date, year of oophorectomy and date of entry into the health plan to control women from the same source. As with the Hoover, Glass, Finkle et al. (1981) study, medical record notations were the source of estrogen use data and other risk factors. No overall increased risk was observed but a possible increase appeared with five or more medical chart notations of estrogen use, a proxy measure for dose.

Hulka

-3-

Our North Carolina study (Hulka, Chambless, Deubner et al., 1982) was essentially negative with respect to oral estrogens and breast cancer occurrence, but the estimated relative risk for breast cancer associated with injectable estrogens was four-fold. This mode of administration is likely to produce high dose availability to target tissues, since the portal circulation with liver deactivation is initially avoided.

The studies by Kelsey, Fischer, Holford et al. (1981), Wingo, Layde, Webster et al. (1983) and Kaufman, Miller, Rosenberg et al. (1984) reported no breast cancer risk attributable to estrogens irrespective of dose, duration or subset of the study group. With a much larger sample size than was available to Ross, Paganini-Hill, Gerkins et al (1980), Kelsey, Fischer, Holford et al. (1981) attempted to replicate the milligram-month strategy. All relative risks were unity or smaller. The studies by Kaufman, Miller, Rosenberg et al (1984) and Wingo, Layde, Webster et al (1983) encompassed very large numbers of subjects in diverse geographic areas of the U.S. and Canada. The former had hospital controls and the latter had population-based controls which should satisfy proponents of both types of control groups. Both studies provided negative results which persisted in all categories of estrogen dose and duration, and for subgroups recognized to be at higher than average risk of breast cancer. An estimate of the potential for biases — selection, misclassification and confounding — in the case-control studies published since 1980 appears in Table 4.

The studies of Kelsey, Fischer, Holford et al. (1981), Hulka, Chambless, Deubner et al. (1982) and Kaufman, Miller, Rosenberg (1984), in which both the case and control groups were not obtained from a defined population, get less

than top marks on the potential for selection bias. The Brinton study (Brinton, Hoover, Szklo et al. 1981) based on volunteers from a breast cancer screening program is the most subject to the selection bias criticism since program participation may be stimulated both by symptomatology related to the breast and propensity to use medical services. High utilizers of medical care are known to receive more medications, including estrogens (Mack, Henderson, Gerkins et al. 1975). On the other hand, the Brinton study excels in the avoidance of diagnostic misclassification among controls, all of whom were ascertained through the same screening procedures as the cases, the former having had normal mammograms and clinical examinations (at least none had been referred for diagnostic evaluation).

The Hoover study (Hoover, Glass, Finkle et al. 1981) was beset with the problems inherent to epidemiologic studies based on medical records used for clinical care purposes. Determination of estrogen exposure was based on prescription notations in the medical record. Number of such notations and interval between first and last notation provided proxies for intensity and duration of use. The difficulty in dealing with confounding factors was also a function of this data source; missing data varied from twelve to forty-four percent for relevant confounders.

The Hiatt study (Hiatt, Bawol, Friedman et al. 1984) receives high marks for absence of selection bias in the nested case control design derived from women all of whom had had prior oophorectomy.

Discussion of findings

In all the studies reviewed estrogen use, relative to nonuse, incurred risks for breast cancer of somewhere between 0.9 and 1.5. These best estimates for the magnitude of the risk have come from several large, carefully designed

case-control studies in different parts of the United States utilizing a variety of sources for both cases and controls. The follow-up studies, although theoretically more desirable, faced some intractable problems related to the selected nature of estrogen exposed women and the difficulty of identifying comparable unexposed women. Despite this, they reported risk ratios within the same range as the case-control studies. The Gambrell, Maier and Sanders (1983) study is an exception by suggesting a protective effect from estrogens.

When evaluating the possibility that a medication is related to the development of a chronic disease after a prolonged latent period, there are certain characteristics of drug use, which if related to risk, increase one's suspicion that the association is causal. Most frequently, dose and duration effects are sought. The higher the dosage and the longer the period over which the drug is administered, the greater should be the risk. Several of the studies reviewed showed increased risk with high dose products while showing no increased risk with daily dosages of .625 mg or less of conjugated estrogens. Our North Carolina study (Hulka, Chambless, Deubner et al. 1982) found that the overall excess risk was almost totally attributable to injectable products. This finding is consistent with a high dose effect since estrogen absorption after parenteral administration allows for target organ effects prior to hepatic deactivation, and also for sustained, high blood levels which are not likely to be achieved through oral administration. Although the dose effect did not appear in the largest studies, its occasional emergence is bothersome.

An independent effect on risk of increasing duration of estrogen use was not evident. Only in one study where follow-up duration was equated with use duration did it appear (Hoover, Gray, Cole et al. 1976).

A stimulus for pursuing an exogenous estrogen effect on breast cancer risk came in 1975 when the first epidemiologic studies showing an estrogen/endometrial cancer association were reported (Smith, Prentice, Thompson et al 1975; Ziel and Finkle 1975). Subsequent studies confirmed and further clarified this association. The endometrial cancer risks from estrogen were consistently large (four to six fold) and the relationship to particular characteristics of estrogen use supported a promotional action of the hormone. Increased risks were strongly related to increased duration of use; dose effects were evident but even modest doses (0.625 mg of conjugated estrogens) had definite adverse effects; the latency interval was short, about three to five years from first use until increased risks were observed; and the interval between stopping estrogen use and the resumption of nonuser risk levels was even shorter (two years or less) (Hulka, Fowler, Kaufman et al 1980).

Since the breast is also a target organ for estrogens, which are essential for ductal growth and proliferation during puberty and pregnancy, the endometrial cancer model was quickly transported to breast cancer investigations. The results, however, have not been as the model would have predicted. The studies summarized in the context of this chapter make that point.

A number of theories have evolved concerning the role of estrogens (endogenous or exogenous) in breast cancer occurrence. Some of these concern the varying distributions of estrogenic subfractions or metabolites; another involves the timing and amount of progesterone in relation to estrogens; or the total availability of biologically active estrogens may be the pertinent issue. The ratio of estriol to estradiol plus estrone (the three major urinary metabolites of estrogens) was one of the first markers thought to be an indicator of breast cancer risk (MacMahon, Cole and Brown 1973). The higher

the ratio, the lower the risk. The rationale for this ratio was based on laboratory data suggesting that estriol impeded the target organ effects of estradiol and estrone, which were the metabolically active estrogens. Further experimentation, however, has shown that estriol itself is biologically active. More recent epidemiologic studies in populations at high and low risk of breast cancer have been equivocal with respect to differences in the estriol ratio.

Subsequently, the theory of Korenman (1980) was proposed concerning the "estrogen window hypothesis" and the inadequate corpus luteum, resulting in decreased progesterone production. It was proposed that the breast was particularly susceptible to carcinogenic agents around the time of puberty and menopause when the normal cycling of progesterone is deficient, leaving estrogens unopposed. The same phenomenon could occur among adult, reproductive age women who experience anovulation as evidenced by irregular menstrual cycles. However, epidemiologic studies have not confirmed an association between irregular menstrual periods and increased risk of breast cancer. Women with early age at menarche and late age at menopause (risk factors for breast cancer) are more likely to have regular, ovulatory menstrual cycles, especially during their early, postpubertal years. These and other observations have lessened enthusiasm for the unopposed estrogen hypothesis.

New information has come forth on a divergent metabolic pathway for estradiol (Fishman and Martucci 1980). It has long been evident that the metabolites, estrone and estriol, along with estradiol, are the compounds which account for the feminizing effects of estrogen. It is now recognized that a significant proportion of estradiol is metabolised to catechol estrogens which have central nervous system effects on the regulation of pituitary hormone release and on control of behavior. A shift in the metabolic pathway toward

catechol estrogen production would reduce the levels of feminizing estrogens which affect the breast, and possibly alter the breast cancer risk.

The difficulty in finding differences in plasma levels of estradiol or other estrogenic metabolites between breast cancer cases and controls has inhibited the estrogen hypothesis in breast cancer causation. Further analysis indicates that total estrogens which are most readily measured in laboratory assays may not be the relevant factor (Siiteri, Hammond and Nisker 1981). Only that portion which is free, or bound to albumin, but not bound to serum hormone binding globulin, is available to the target tissues. Therefore, total estrogens may not differentiate cases and controls. However, if cases have lower serum hormone binding globulin levels than controls, the amount of available estradiol would be greater. Some data are available to support this supposition. Hence, the estrogen excess hypothesis in breast cancer pathogenesis is still being pursued.

There is a need for rethinking both biochemical and epidemiologic research along the lines of multiple hormones, their interactions, their serum binding proteins and their tissue receptor sites. The estrogen effect on breast tissue will be determined not only by availability of estrogens to the estrogen receptor but the availability of other hormones. At least progesterone, androgens and prolactin, will alter the estrogen effect. For example, increased progesterone levels inhibit estrogen receptor formation and thus reduce the availability of estrogens to breast tissue. Most epidemiologic studies of combination product oral contraceptives in which progestogens predominate show no increased risk of breast cancer (Kelsey, 1979).

Lastly, the significance of gonadal hormone administration before, during and after the menopause needs to be more clearly distinguished. The difference between hormone replacement (after the menopause) and supplemental estrogens

(before or during the menopause) and the role of biofeedback mechanisms with pituitary hormones have received inadequate attention.

Conclusion

1. Epidemiologic data

The numerical data obtained from epidemiological research do not provide strong support for an association between replacement estrogens and breast cancer risk. Although the findings are not consistent in all studies, one aspect of the data deserves notice. Those studies showing adverse estrogen effects concentrated their adversity in the high dose range (1.25 mg or more of conjugated estrogens). These doses are unnecessarily high to achieve the usual clinical goals of estrogen therapy, which are relief of menopausal symptoms and prevention of osteoporosis. A second observation, partially supported by the data but strongly supported by biologic plausibility, is the possibility of an estrogen-associated increased breast cancer risk among oophorectomized women. This increased risk is relative to the reduced risks inferred by bilateral oophorectomy at an early age, so that the increase is a reflection of the replacement function, with breast cancer risk increasing to levels assumed by women who undergo a natural menopause.

2. A risk assessment model for medications

The example of endometrial cancer occurrence in relation to estrogen use was presented to illustrate the characteristics of drug use which enhance the likelihood of a causal association between a medication and cancer occurrence. Characteristics of dose, duration, latency and recency of use appeared with reasonable repeatability and predictability in numerous studies of estrogens and endometrial cancer. The particular observations were consistent with the expected behavior of a promotional agent in cancer causation. However, in the

Hulka

-10-

breast cancer studies, only the dose effect was observed, and that appeared in fewer than half of the studies.

3. The biological interpretation

The sustained interest in an estrogen hypothesis relative to breast cancer occurrence derives primarily from the research of endocrinologists working in laboratory and clinical settings. Such research has produced clear evidence of estrogen effects on breast tumor tissue, information which has been applied in the form of hormonal manipulation as part of breast cancer therapy.

The responsiveness of breast tissue to estrogen action is a justified reason for placing estrogens high on the suspect list of tumorigenic agents. However, the biological and biochemical measurements reported in human studies have not shown clear estrogen or estrogen-related markers which could distinguish women with breast carcinoma from those without the disease, or those at high risk from those at low risk. Renewed efforts to understand the interactive effects of both sex steroid and protein hormones in relation to serum binding proteins and cellular receptors may improve the data base on which to justify and modify existing theory.

References

- Boston Collaborative Drug Surveillance Program. (1974) Surgically confirmed gallbladder disease, venous thromboembolism, and breast tumors in relation to postmenopausal estrogen therapy. N. Engl. J. Med. 290:15-9.
- Brinton, L.A., Hoover, R.N., Szklo, M., et al. (1981) Menopausal estrogen use and risk of breast cancer. Cancer 47:2517-22.
- Brinton, L.A., Williams, R.R., Hoover, R.N., et al. (1979) Breast cancer risk factors among screening program participants. J. Natl. Cancer Inst. 62:37-43.
- Byrd, B.F., Burch, J.C. and Vaughn, W.K. (1977) The impact of long term estrogen support after hysterectomy. Ann. Surg. 185:574-80.
- Casagrande, J., Gerkins, V., Henderson, B.E., et al. (1976) Brief communication: exogenous estrogens and breast cancer in women with natural menopause. J. Natl. Cancer Inst. 56:839-41.
- Craig, T.J., Comstock, G.W. and Geiser, P.B. (1974) Epidemiologic comparison of breast cancer patients with early and late onset of malignancy and general population controls. J. Natl. Cancer Inst. 53:1577-81.
- Fishman, J. and Martucci, C.P. (1980) New Concepts of Estrogenic Activity: The Role of Metabolites in the Expression of Hormone Action. Chapter 5 In, The Menopause and Postmenopause, (eds. N. Pasetto, et al.), MTP Press Limited, Lancaster, England.
- Gambrell, R.D., Maier, R.C. and Sanders, B.I. (1983) Decreased incidence of breast cancer in postmenopausal estrogen-progestogen users. Obstet. Gynecol. 62:435-43.

- Hammond, C.B., Jelowsek, F.R., Iae, K.L., et al. (1979) Effects of long-term estrogen replacement therapy, II. Neoplasia. Am. J. Obstet. Gynecol. 133:537-47.
- Henderson, B.E., Powell, D., Rosario, I., et al. (1974) An epidemiologic study of breast cancer. J. Natl. Cancer Inst. 53:609-14.
- Hiatt, R.A., Bawol, R., Friedman, G.D., et al. (1984) Exogenous estrogen and breast cancer after bilateral oophorectomy. Cancer 54:139-144.
- Hoover, R., Glass, A., Finkle, W.D., et al. (1981) Conjugated estrogens and breast cancer risk in women. J. Natl. Cancer Inst. 67:815-20.
- Hoover, R., Gray, L.A., Cole, P. et al. (1976) Menopausal estrogens and breast cancer. N. Engl. J. Med. 295:401-05.
- Hulka, B.S. (1980) Effect of exogenous estrogen on postmenopausal women: The epidemiologic evidence. Obstet. Gynecol. Surv. 35:389-99.
- Hulka, B.S., Chambles, L.E., Deubner, D.C., et al. (1982) Breast cancer and estrogen replacement therapy. Am. J. Obstet. Gynecol. 143:638-44.
- Hulka, B.S., Fowler, W.C., Kaufman, D.G., et al. (1980) Estrogen and endometrial cancer: Cases and two control groups from North Carolina. Am. J. Obstet. Gynecol. 137:92-101.
- Kaufman, D.W., Miller, D.R., Rosenberg, L., et al. (1984) Noncontraceptive estrogen use and the risk of breast cancer. JAMA 252:63-7.
- Kelsey, J.L. (1979) A Review of the Epidemiology of Human Breast Cancer. In, Epidemiologic Reviews, (ed. Sartwell), The Johns Hopkins University Press, Baltimore.
- Kelsey, J.L., Fischer, D.B., Holford, T.R., et al. (1981) Exogenous estrogens and other factors in the epidemiology of breast cancer. J. Natl. Cancer Inst. 67:327-33.
- Korenman, S.G. (1980) The endocrinology of breast cancer. Cancer 46:874-78.

- Mack, T.M., Henderson, B.E., Gerkins, V.R., et al. (1975) Reserpine and breast cancer in a retirement community. N. Engl. J. Med. 292:1366-71.
- MacMahon, B., Cole, P. and Brown, J. (1973) Etiology of human breast cancer: A review. J. Natl. Cancer Inst. 50:21-42.
- Ross, R.K., Paganini-Hill, A., Gerkins, V.R., et al. (1980) A case-control study of menopausal estrogen therapy and breast cancer. JAMA 243:1635-39.
- Sartwell, P.E., Arthes, F.G. and Tonascia, J.A. (1977) Exogenous hormones, reproductive history, and breast cancer. J. Natl. Cancer Inst. 59:1589-92.
- Siiteri, P.K., Hammond, G.L. and Nisker, J.A. (1981) Increased Availability of Serum Estrogens in Breast Cancer: A New Hypothesis. In, Banbury Report 8: Hormones and Breast Cancer, Cold Spring Harbor Laboratory.
- Smith, D.C., Prentice, R., Thompson, D.J., et al. (1975) Association of exogenous estrogen and endometrial carcinoma. N. Engl. J. Med. 93:1164-67.
- Wingo, P.A., Layde, P.M., Webster, L.A., et al. (1983) Risk of breast cancer in women using postmenopausal estrogens. Presented at American College of Obstetrics and Gynecology Meeting, Atlanta, GA, May 1983.
- Wynder, E.L., MacCornack, F.A. and Stellman, D.S. (1978) The epidemiology of breast cancer in United States caucasian women. Cancer 41:2341-54.
- Ziel, H.K. and Finkle, W.D. (1975) Increased risk of endometrial carcinoma among users of conjugated estrogens. N. Engl. J. Med. 293:1167-70.

TABLE 1: Case-control studies of exogenous estrogen and breast cancer, 1974-1979

<u>Authors</u>	<u>Cases</u>		<u>Controls</u>		<u>Relative Risk^a</u>
	<u>Year of Diagnosis</u>	<u>No.</u>	<u>Source</u>	<u>No.</u>	
Wynder, MacCronack and Stellman (1978)	1969-75	785	Hospital	2231	1.1
Sartwell, Arthes and Tonascia (1977)	1969-72	284	Hospital	367	0.8
Boston Collaborative Drug Surveillance Program (1974)	1972	51	Hospital	774	1.0
Brinton, Williams, Hoover et al. (1979)	1973-77	405	Screening program	1156	1.0
Henderson, Powell, Rosario et al. (1974)	1969-72	308	Physicians' offices	308	0.8
Casagrande, Gerkins, Henderson et al. (1976)	1972-73	47	Neighborhood	31	3.1
Craig, Comstock and Geiser (1974)	1949-67	132	Population and neighborhood	260	1.0
Mack, Henderson, Gerkins et al. (1975)	1971-75	111	Community	444	1.6

^aRelative risk for "ever use" of estrogen.

Source: Hulka (1980)

TABLE 2: Practice-based follow-up studies of exogenous estrogen and breast cancer, 1976-1984

<u>Authors</u>	<u>Number Estrogen Exposed</u>	<u>Number Not Estrogen Exposed</u>	<u>Intake and Fol- low-up Period</u>	<u>Defini- tion of Estrogen Use</u>	<u>Rela- tive Risk</u>
Byrd, Burch, and Vaughan (1977)	1016(33) ^a	0	1948-76	>3 yr	1.4
Hoover, Gray, Cole et al. (1976)	1891(49)	0	1939-72	≥6 mo	1.3
Hammond, Jelovsek, Lee et al. (1979)	301 (4)	309 (4)	1940-75	≥5 yr	1.1
Gambrell, Maier and Sanders (1983)	15,606 ^b (22)	5,258(18)	1975-81	not stated	0.4

^aNumber of incident breast cancer cases in parentheses.
^bPerson-years. Number of women not stated.

Source: Hulka (1980)

TABLE 3: Case-control studies of replacement estrogens and breast cancer risk among postmenopausal women, 1980-1984

Journal and Date	Study Site and Authors	Number and Source of:		Relative Risk		Comments
		Cases	Controls	Estrogen Use	95% CI	
JAMA (1980)	Los Angeles Ross, Paganini-Hill, Gerkins et al.	Retirement community 131	- 262	overall	0.8-1.9	risk with total milligram dose. No data on duration of use.
				ovaries removed	0.5-3.5	
				0.8 intact	0.7-2.4	
Cancer (1981)	28 geographic areas Brinton, Hoover, Szklo et al.	BCDDP - screening 881	863	overall	1.0-1.5	risk with dose No duration risk.
				1.2 removed	0.9-2.8	
				1.5 intact	0.9-1.5	
J Nat'l Cancer Inst. (1981)	Connecticut Kelsey, Fischer, Holford et al.	3 hospitals 332 (age 45+, not all postmenopausal)	1353	ovaries removed		No effect milligram months of use on ri
				0.5-1.5 intact		
				0.9	0.6-1.2	
J Nat'l Cancer Inst. (1981)	Portland Oregon Hoover, Glass, Finkle et al.	Kaiser Foundation Health Plan 324 (pre and post menopausal)	549	overall		risk with number of prescriptions and daily dos
				1.0-2.0 ovaries removed		
				0.3-6.6 intact	0.8-2.1	
				1.3		

TABLE 3 (continued)

Journal and Date	Study Site and Authors	Number and Source of:		Relative Risk		95% CI	Comments
		Cases	Controls	Estrogen Use			
Am. J. Obstet. Gynecol. (1982)	North Carolina Hulka, Chambless, Deubner et al.	3 community hospitals 199	451 852	ovaries not removed 1.3 1.2 intact 1.3 1.2		risk with significant estrogens	injectable No dose or duration effects on risk.
American College of Obstet. and Gynecol. presen- tation	8 geographic areas Wingo, Layde, Webster et al.	population based 705 (age <55)	1027	overall 1.0 ovaries removed 1.1 intact 0.6	0.8-1.3 0.8-1.5 0.4-1.2		No dose, duration or sub- group. effect on risk.
JAMA (1984)	8 metro- politan US & Canada Kaufman, Miller, Rosenberg et al.	multiple hospitals 925	1127	overall 0.9	0.7-1.1		No dose, duration or sub-gro effect on
Cancer (1984)	Northern California Hiatt, Bawol, Friedman, et al.	Kaiser Foundation Health Plan 119 (all with bilateral oophorectomy)	119	ovaries removed			risk with number of prescripti duration o risk.

TABLE 4: Potential for biases in case-control studies of replacement estrogens and breast cancer risk among postmenopausal women, 1980-1984

Authors	Risk	Biases ^a		
		Selection	Misclassification	Confounding
Ross, Paganini-Hill, Gerkins et al. (1980)	natural menopause	2 0	2	1
Brinton, Hoover, Szklo et al. (1981)	surgical menopause	0	2+	2
Kelsey, Fischer, Holford et al.	—	1	2	2
Hoover, Glass, Finkle et al. (1981)	pre and post meno- pausal	2	0	0
Hulka, Chambless, Deubner et al. (1982)	—	1	2	2
Wingo, Layde, Webster et al. (1983)	—	2	2	2
Kaufman, Miller, Rosenberg et al. (1984)	—	1	2	2
Hiatt, Bawol, Friedman et al. (1984)	—	2+	1	1

^a bias

0 very likely; 1 possible; 2 not likely

Hulka

-19-

Chapter 22

Leukemia Risks in Relation to Benzene
Exposures: An Examination of the Data

Darwin R. Labarthe

LEUKEMIA RISKS IN RELATION TO BENZENE EXPOSURES:
AN EXAMINATION OF THE DATA

Darwin R. Labarthe

Introduction

The central issues in current discussions about the relation between benzene exposures and the risks of leukemia are not whether there may be a causal link but rather under what circumstances of exposure and for what types of leukemia might such a link operate. This chapter will focus on the first of these two questions. The studies available to address the quantitative aspects of risk are few in number, small in size, and subject to varying interpretations. They are reviewed here from the point of view that an attempt to construct a quantitative risk pattern on the basis of these studies may help to identify questions of importance for further resolution of this issue. Four studies have reported risks in relation to some measure of benzene exposure, as ppm in air in the workplace, the units on which most of the discussion and debate has centered (Figure 1). These studies were published between 1974 and 1983 (Aksoy, Erdem and Dincol 1974; Infante, Rinsky, Wagoner et al. 1977; Ott, Townsend, Fishbeck et al. 1978; Tsai, Wen, Weiss et al. 1983). Comprehensive discussion of each study is not possible here, but for each study the risks, exposures, and some general comments will be presented.

Aksoy and colleagues

In December, 1974, Aksoy and colleagues at the Section of Hematology of the Istanbul Medical School reported an analysis of their clinical experience with leukemia cases among shoe-workers exposed chronically to benzene (Aksoy, Erdem and Dincol 1974). Over a seven year period they had identified twenty six cases which they attributed to an underlying population of 28,500 such workers. A calculation was described the result of which the authors took as evidence of a "marked and statistically significant increase" in the incidence

of leukemia among such workers over that of the general population. They claimed to have provided "cogent evidence for the clinical impression of leukemogenic effects of benzene in man." Their conclusion was based on a ratio of 13 per 100,000 events observed to 6 per 100,000 expected, or 2.17. However, attentive reading of both their report and the source cited for the expected value indicates some problems with this calculation. The observed cases were incident cases of preleukemia and acute leukemia in Istanbul hospitals from 1967-1973. The expected number was based on a report of leukemia deaths in twenty-four countries in 1965 (Gunz 1970). No adjustment was made for age differences between the populations compared, incidence was compared with mortality, and no evidence was provided to suggest that the leukemia rates for Turkey resembled those of the twenty-four countries. These and other issues (such as those raised by Matanoski in an earlier presentation in this symposium) dictate some caution in evaluating the risk estimate presented by these investigators.

More crucial to the immediate discussion is the information on exposures to benzene related to this estimate of risk. The 1974 report just cited identified all twenty-six cases and their source population as shoe workers, exposed to benzene as a solvent, said to be present in 9 to 88 percent of the composition of the adhesives in use (Aksoy, Dincol, Akgun et al. 1971; Aksoy, Erdem and Dincol 1974). The authors stated, "The working conditions in the industry are not good; their shops are usually small, not hygienic, and poorly ventilated; the concentration of benzene was found to reach a maximum of 210-650 parts per million when adhesives containing benzene were in use." In six additional reports, variously citing these cases and others, other information is given to suggest benzene concentrations of 150-650 ppm, as

exposures occurring for eight hours or more daily, for a mean duration of 6.7 years, among these cases (Aksoy 1980; Aksoy 1977; Aksoy, Dincol, Akgun et al. 1971; Aksoy, Dincol, Erdem et al. 1972; Aksoy and Erdem 1978; Aksoy, Erdem and Dincol 1976). Benzene concentrations ranging from 15-30 ppm were reported outside working hours (Aksoy, Dincol, Akgun et al. 1971).

These reports do not indicate which locations were monitored or how many, how they were selected, when the reported measurements were obtained in relation to the detection of cases of leukemia, or precisely what the exposures were in terms of eight hour Time-Weighted Averages (TWA). They do not, with few exceptions, provide estimates of exposure specific to individuals. However, they do suggest that in a large occupational group exposed for several years at levels commonly reaching into the hundreds of ppm, an excessive rate of leukemia occurrence was observed. This excess risk of leukemia in the late 1960s and early 1970s was considered to have abated by 1975, the apparent decline being attributed to substitution of other material for benzene beginning in 1969 (Aksoy 1980).

Infante and colleagues

In July, 1977, Infante and colleagues, of the U.S. National Institute for Occupational Safety and Health, reported the results of a study of a cluster of 7 leukemia cases and a population of 748 pliofilm workers, at two localities in Ohio, all of whom were judged to have had "direct exposure to benzene" in the period 1940-1949 (Infante, Rinsky, Wagoner et al. 1977a). Two overall estimates of leukemia risk were presented, one being an SMR of 474 when pliofilm workers and fibrous-glass workers were compared, with seven deaths observed versus 1.5 expected. When white US males were the comparison group, the SMR was 506. An SMR specific for myelogenous and monocytic leukemia of

1004 was presented later in the report, challenged in subsequent published correspondence, and revised to 857 (Infante, Rinsky, Wagoner et al. 1977a; Infante, Rinsky, Wagoner et al. 1977b; Tabershaw and Lamm 1977). A still later analysis, restricted to workers with greater than five years of exposure, resulted in an SMR of 2100.

Regardless of which risk estimate one considers, the levels of benzene exposure which may have been experienced by these workers are of particular interest. The available reports give differing impressions. The authors of the original report described the process of pliofilm production during which benzene exposure could occur and cited monitoring data for one of the two localities (Infante, Rinsky, Wagoner et al. 1977a). Levels from 0 to 10 or 15 ppm had been described in a 1946 document which was cited, although the authors made the less specific claim that "employees' benzene exposure was generally below the recommended limit in effect at the time of each survey." A chronological table of these recommended limits was then presented which ranged from a maximum allowable concentration of 100 ppm in 1941 to an eight hour TWA of 10 ppm in 1971. However, the 10 ppm standard was not in effect until eight years after the death of the last of their seven cases, as indicated in a later report (Rinsky, Young and Smith 1981).

In that later report, Rinsky and others provided very extensive detail as to the available exposure information. They emphasized the conclusion, based on this study, that "benzene is a human carcinogen at a range of exposures not greatly above the current legal standard," which, at the time of that report, was an eight hour TWA of 10 ppm (Rinsky, Young and Smith 1981). However, that same report gave contrary information on individual exposure estimates for the leukemia cases, with five of seven having 35 or 40 ppm eight hour TWA

exposures, one having "probably less than 100 ppm," and one, among those employed prior to any ventilation of the work area, having no quantitation provided. In a still later report, White, Infante and Chu (1982) stated: "The average benzene exposure over the period of 1937 to 1954...was 83 ppm. Over the period 1937-1975, the average benzene exposure was 50 ppm". Since, according to Rinsky and others, the great preponderance of the potential exposure period among the seven cases occurred in the 1937-1954 interval, the figure of 83 ppm appears to be the more relevant of the two (Rinsky, Young and Smith 1981). The basis for these estimates is unclear, however, since no monitoring data were available, according to the Rinsky report, for the location where five of the seven workers were employed, until a date after all but the last case had discontinued employment there (Rinsky, Young and Smith 1981).

Some additional information can be gleaned from the reports, as follows. Limited benzene monitoring data were available for one location well within the actual exposure period of interest (Rinsky, Young and Smith 1981). The authors estimated the personal benzene exposure histories for six of the seven cases but indicated a lack of information both on the earliest years of exposure, for all but one case, and on possible benzene exposures additional to those involved in the pliofilm process, for four of the seven cases. The reported values could therefore be interpreted as possible underestimates of the actual exposures of these workers. Use of respiratory protective equipment by workers at some times and locations was mentioned, but not in reference to any of the leukemia cases themselves. For the six or seven cases, exposures apparently occurred prior to any workplace ventilation, such conditions perhaps accounting for 50 percent of the cumulative exposure period among these workers. No

attempt at characterization of personal exposures among other members of the study population was reported. Notably, also, the last of these seven deaths occurred in 1961 (Rinsky, Young and Smith 1981).

Ott and colleagues

In 1978, Ott and Colleagues, from the Dow Chemical Company, reported on mortality among workers occupationally exposed to benzene (Ott, Townsend, Fishbeck et al. 1978). Of 594 such workers, 335 had been exposed prior to 1950, the three work areas involved having been operated from 1920, 1935, and 1936, respectively. In the analysis most specific for benzene exposures, the occurrence of one death coded as leukemia was related to 0.9 cases expected on the basis of US white male mortality over the corresponding calendar period, which would yield a relative risk of 1.11. In a larger group with substantial other exposures, 53 additional subjects added 0.1 expected cases of leukemia and one observed leukemia death, which would yield a relative risk of 2. A third death was necessarily excluded from these calculations since it had not been recorded as due to leukemia as the underlying cause on the death certificate.

Benzene exposure levels, by work area, were reported as eight hour TWA values from 1944 at the earliest, in one area, to 1952-1953 in the other two areas. The eight hour TWA values over the period of study were estimated from varying numbers of samples, at specified locations. The minimum and maximum values of these estimates were 0.1 and 35.5 ppm, respectively. For the individual cases, the work histories permitted estimates of eighteen months of potential exposure below 2 ppm eight hour TWA, for one case, and two years of potential exposure at 2-9 ppm, in addition to other chemical exposures, for the other case.

The workplace monitoring data for the mid-1940s and early 1950s, when these workers were exposed, were limited, but the corresponding estimated 8-hour TWAs were in the 9-35.5 ppm range. Multiple occupations, and multiple exposures while employed by Dow, occurred for the first and second of these two cases, respectively. An incidence estimate was also given, in a single sentence in the discussion section of the Ott report. However, there is a difficulty with this calculation, in that the expected number of cases excluded lymphocytic and monocytic cell types, while the cell types were not known for all three of the cases included. For one case, death was ascribed to 'leukemia' and not further specified; a later statement gives the impression of myelocytic leukemia, type not specified. Accordingly, the expected value may have been underestimated, the ratio of observed to expected therefore overestimated, and the impression of a statistically significant excess was incorrect. The sensitivity of all risk estimates from this study to the addition or subtraction of a single case is to be noted, although the expected numbers of deaths are only slightly below those of the Infante study (Infante, Rinsky, Wagoner et al. 1977).

Tsai and colleagues

In 1983, Tsai and colleagues, of the Gulf Oil Corporation, reported on a study of workers in benzene areas of a refinery (Tsai, Wen, Weiss et al. 1983). The study cohort included all male workers ever employed in the principal petrochemical units of a single refinery from late 1952 through 1957. From mortality data for US males, 0.42 deaths from leukemia were expected among the 454 workers studied. No case of leukemia was observed, so that the SMR was 0. It was noted as well that no still-living employee had developed leukemia, as of 1979, but the methods of ascertainment were not stated.

Monitoring data were judged to be inadequate prior to 1973, but personal sampling data, based on a total of 1394 samples taken from 1973-1982, indicated benzene concentrations from below 0.1 ppm (43 percent of samples) to 25 ppm or above (1 percent of samples). These values were not adjusted to eight hour TWAs. The duration of exposure was on average 8.0 years for white and 4.5 years for non-white workers, with one-half of the cohort working in benzene areas prior to 1965 and 80% so employed for at least one year.

The size of the study population provided for only a very small expectation of leukemia deaths. The fact that exposures were reported as single sample values rather than 8-hour TWAs suggests that potential exposure levels were well below those in the Dow study, at least in the period for which data were available. Qualitative information on the nature of the benzene exposures of these workers was not provided.

Conclusions

The data on benzene concentrations in these studies are valuable, though far short of the ideal, as a guide to arraying the reported leukemia risks in relation to potential exposures. Differences in data quality, manner of reporting, and other aspects not addressed here, limit the comparability of these studies. Further, the term "environmental data" and not "exposure data" should perhaps be used, since actual individual exposures can only, at best, be inferred.

The qualitative information on the work environments represented by these studies, when available, is also valuable in further characterizing the nature of the potential exposures, as to general hygienic features, chronology of process changes, and so on. More detailed and more consistent reporting of such information, as in the Rinsky (Rinsky, Young and Smith 1981) report, would be useful.

LaBarthe

-8-

It has been implied that existing data demonstrate significant risks of leukemia for workers near the current exposure standard of 10 ppm eight hour TWA (Rinsky, Young and Smith 1981). The strongest evidence from these studies is that excessive risks are related to exposures at levels in the hundreds, eighties, or fifties of ppm, for several years; the conflicting impressions about the pliofilm workers' exposures, perhaps in the 35-40 ppm range plus other, unquantified exposures, should be resolved. In any case, these observations are not consistent with the contention that the 10 ppm level has been demonstrated to relate to significant excess risk. If valid, this contention could be of great importance for the health and safety of potentially exposed workers. Resolution of this issue would be aided most by further epidemiologic studies which addressed, among others, the following questions:

— Can the observations among the Ohio pliofilm workers be replicated elsewhere, thus possibly adding to the weight of evidence from only seven cases of leukemia concerning long-term exposures in the 25-100 ppm range?

— Can actual excess risks of leukemia at or near a 10 ppm eight hour TWA exposure level be demonstrated? If so, under what conditions of exposure? If not, can such risks reasonably be excluded at this or higher exposure levels but below those described for the affected pliofilm workers?

— What is the distribution, and what are the conditions, of benzene exposure levels in the workplace today? To what extent do the hazards shown in two of these four historical studies persist?

If there were, as claimed, some two million potentially exposed workers as recently as 1977 (Infante, Rinsky, Wagoner et al. 1977a), and presumably many more in the more remote past in the U.S. alone, studies of greater size and of

at least the quality of those reported to date should be feasible. Such studies could add importantly to the limited data presently available on the quantitative relations between benzene exposures and the risks of leukemia. Observations on specific types of leukemia would be of greatest value. The implications of these questions for public health practice, for regulatory policy, and for the extensive litigation in progress and in prospect would seem to lend some urgency to establishing a firmer basis for the answers.

REFERENCES

- Aksoy, M. (1980) Different types of malignancies due to occupational exposure to benzene: A review of recent observations in Turkey. *Environ. Res.* 23:181-90.
- Aksoy, M. (1977) Leukemia in workers due to occupational exposure to benzene. In, *New Istanbul Contribution to Clinical Science* 12:3-14.
- Aksoy, M., Dincol, K., Akgun, T., Erdem, S. and Dincol, G. (1971) Haematological effects of chronic benzene poisoning in 217 workers. *Br. J. Ind. Med.* 28:296-302.
- Aksoy, M., Dincol, K., Erdem, S. and Dincol, G. (1972) Acute leukemia due to chronic exposure to benzene. *Am. J. Med.* 52:160-66.
- Aksoy, M. and Erdem, S. (1978) Followup study on the mortality and the development of leukemia in 44 pancytopenic patients with chronic exposure to benzene. *Blood* 52:285-92.
- Aksoy, M., Erdem, S. and Dincol, G. (1974) Leukemia in shoe-workers exposed chronically to benzene. *Blood* 44:837-41.
- Aksoy, M., Erdem, S. and Dincol, G. (1976) Types of leukemia in chronic benzene poisoning. A study of thirty-four patients. *Acta. Haematol.* 55:65-72.
- Gunz, F.W. (1970) Problems in leukemia etiology. Plenary Sessions Scientific Contributions. XIII International Congress of Hematology. Munich, August 2-8, pp 48-57.
- Infante, P.F., Rinsky, R.A., Wagoner, J.K. and Young, R.T. (1977a) Leukaemia in benzene workers. *Lancet* 2:76-78
- Infante, P.F., Rinsky, R.A., Wagoner, J.K. and Young, R.T. (1977b) Benzene and leukaemia. *Lancet* 2:868-869

- Ott, M.G., Townsend, J.C., Fishbeck, W.A. and Langner, R.A. (1978) Mortality among individuals occupationally exposed to benzene. Arch. Environ. Health 33:3-10.
- Rinsky, R.A., Young, R.J. and Smith, A.B. (1981) Leukemia in benzene workers. Am. J. Ind. Med. 2:217-45.
- Tabershaw, I.R. and Lamm, S.H. (1977) Benzene and leukaemia. Lancet 2 867-68.
- Tsai, S.P., Wen, C.P., Weiss, N.S., Wong, O., McClellan, W.A. and Gibson, R.L. (1983) Retrospective mortality and medical surveillance studies of workers in benzene areas of refineries. J. Occup. Med. 25:685-92
- White, M.C., Infante, P.F. and Chu, K.C. (1982) A quantitative estimate of leukemia mortality associated with occupational exposure to benzene. Risk Analysis, 2:195-204

Relative Risk of Leukemia vs. Benzene Concentrations 4 Reports

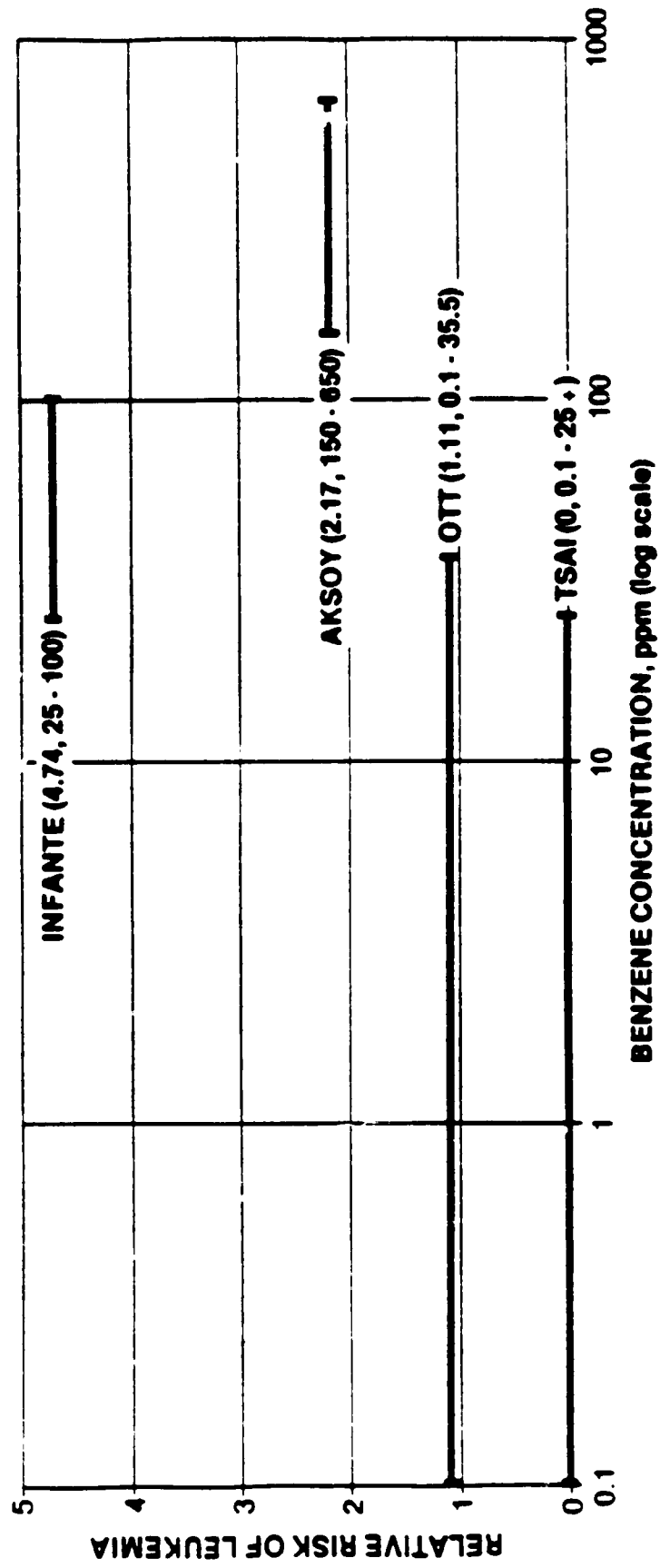


Figure 1

Chapter 23

The Influence of Theoretical and Experimental Radiobiology
on the Epidemiology of Radiation Carcinogenesis

Charles E. Land

The Influence of Theoretical and Experimental Radiobiology
on the Epidemiology of Radiation Carcinogenesis

Charles E. Land

Introduction

Ionizing radiation is a known cause of cancer, to which everyone is exposed, every day. In round numbers, yearly accumulation for the average US resident from natural background sources is about 0.1 rad to most internal organs (Committee on the Biological Effects of Ionizing Radiation 1980). The bronchial epithelium receives about 1.5 times that much from inhaled radon gas and its radioactive decay products (NCRP 1984). It is known that radiation can cause cancer because excess risks have been observed consistently in various human populations exposed to radiation doses hundreds of times higher than yearly background levels, and because cancer can be induced experimentally by irradiation of laboratory animals. These observations are so consistent that, for many cancer sites, there is no longer any serious doubt about the existence of risk at some level of exposure.

Epidemiological research tends now to be focused, therefore, upon the details of risk, and in particular its magnitude at low exposure levels, rather than on its existence. Some of the questions which arise are as follows:

- How does risk vary by various quantitative measures of radiation exposure, that is, by radiation dose and its distribution in time and space?

- How does it vary over time following exposure?
- How does it vary by properties of the irradiated individual, such as sex, age, or genetic background?
- How does it depend on other exposures and life events, either before or after irradiation?

Variation by radiation dose and its distribution in time and space

One reason for the remarkable productivity of experimental radiobiology is that radiation exposure to tissues and organs of interest can be determined with great precision. Even in epidemiological studies, one can often be quite specific about the exposure to quite small volumes of tissue. In fact, often it is possible to get far more detailed dosimetry than that which forms much of the basis of current risk estimates. Potentially, therefore, existing data may contain more information than has been gained from it thus far.

Radiation dose is the amount of energy deposited in tissue, averaged over some volume of tissue and accumulated over some period of time. It is measured in units of rad or gray ($1 \text{ Gy} = 100 \text{ rad}$). The temporal distribution of dose is described in terms of fractions, or periods during which dose accumulates separated by periods during which it does not, and by the rate at which dose accumulates during each fraction. The spatial distribution of dose can be described in terms of variation over different parts of the body, or over different parts of an organ. It can also be described microscopically, in terms of uniformity or non-uniformity of energy deposition within extremely small volumes of tissue.

The theoretical basis for dose-response modelling

The mechanism by which radiation is thought to cause biological effects is through the breaking of chemical bonds by ionization, for example, by adding or removing electrons from atoms or molecules. The number of ionizations per unit volume of tissue is proportional to the amount of absorbed energy, or dose. Ionizations occur along radiation tracks, which can be thought of as paths of photons or sub-atomic particles such as electrons, protons, neutrons, and alpha particles (helium nuclei). Radiations of different kinds can differ in their ability to penetrate materials, the amount of energy released along a single track, and the density of ionizations along each track. These properties are summarized under "quality", or "linear energy transfer" (LET).

Figure 1 is a schematic diagram of ionizations along high-LET and low-LET radiation tracks (Upton 1982a). The point of this diagram is that if a certain locus in a cell (a spherical volume around a particular point on a DNA strand) is affected by one ionization from a high-LET track, an event whose probability is proportional to dose, it is very likely that the same locus will be affected by several more ionizations from the same track. If the locus is affected by an ionization from a sparsely ionizing track, it is unlikely that it is also affected by many more from the same track. Therefore, many ionizations at that same locus are likely to occur only if two tracks intersect there. The probability of two intersecting tracks at a given locus is proportional to the square of dose. If it is assumed that damage from closely spaced ionizing events tends to be more consequential than that from the same number of separated events, these considerations lead to a linear-quadratic expression in dose, in which the linear part corresponds to single-track events, while the dose-squared part corresponds to multiple-track interactions (proportional to dose-squared and higher powers of dose, but dominated by dose-squared).

Land

-3-

A competing radiation effect, well documented in experimental studies, is cell inactivation or cell killing. A cell that cannot divide cannot develop into a cancer. A conventional modification of the dose-response model to allow for cell sterilization is to multiply the linear-quadratic function of dose, which corresponds to cancer initiation, by a negative exponential in a linear or linear-quadratic function of dose (Figure 2). The curve is intended to represent leukemia risk in an irradiated human population with uniform, whole-body exposure. Because the magnitude of the effect depends upon age at exposure, sex, and time after exposure, the ordinate scale begins at zero but is otherwise unspecified. Note that, for whole-body irradiation, doses above 500 rad or so are acutely lethal with high probability. The right-hand portion of the graph is therefore irrelevant to uniform, whole-body exposure; its usefulness lies in describing the effects of partial-body exposure.

The practical meaning of the event just discussed is to limit the universe of possible dose-response relationships to a parametric family of functions. That is, the problem is reduced to one of more manageable size and complexity, through the incorporation of hard-won information from experimental and theoretical radiobiology. The assumption that radiation carcinogenesis in man is analogous to various endpoints that are amenable to experimental study, including carcinogenesis but also including mutagenesis, clastogenesis, and cell killing, is thus both crucial and extremely powerful.

Further influences of experimental radiobiology

The incorporation of theoretical and experimental information goes somewhat further. In fact, no existing epidemiological dose-response data are strong enough to select a single member of the above family of parametric

Land

-4-

functions, without further restrictions on the range of possibilities. With the atomic bomb survivor data, it is at least plausible to ignore the cell-inactivation component on the grounds that, for whole-body, acute exposures, a substantial cell inactivation effect may be inconsistent with survival of the exposed person. Even so, it is not possible, for most cancer sites, to select a single linear-quadratic curve with any confidence. The reason is that estimation of the relative importance of the one-track and two-track components of risk depends upon the ability to estimate excess risk at low as well as at high dose levels. Vastly more observations are required to assess the effect of radiation at low doses, where the excess risk is small, than at high doses, where risk tends to be high relative to background rates (Land 1980). The 1980 BEIR Committee found it necessary, for most cancers, to further restrict the family of possible dose-response functions by assuming a fixed ratio between coefficients for dose and dose-squared, corresponding to equivalence for one- and two-track events at about 116 rad (Committee on the Biological Effects of Ionizing Radiation 1980). This equivalence was based partly on analyses of leukemia data but also on relationships observed in experimental radiobiology. Thus, presently achievable low-dose risk estimates depend not only on observed risk at higher dose levels, but also on acceptance of analogies with experimental data.

Some testable consequences

The hypothesis that cancer is more likely to result from ionizing events that are closely spaced than from events that are separated has many testable consequences (NCRP, 1980). Upward curvature of dose-response curves for low-LET radiation is easily demonstrated in experimental models like the induction of pink stamen hair mutations in *tradescantia* or killing of cells in

Land

-5-

culture; it also can be seen, although less clearly and consistently, in experimental carcinogenesis, and in the frequency of stable chromosome aberrations in cultured leukocytes obtained from A-bomb survivors. Such curvature has been difficult to demonstrate unequivocally for cancer risk in man, however, for reasons of sample size as discussed above.

Another consequence is that dose-response curves for high-LET radiation should have less upward curvature, because the likelihood of closely-spaced ionizing events is not increased by multi-track interactions, or at least not as much as for low-LET radiation. Also, low doses of high-LET radiation should, in general, be more effective at causing biological damage than low doses of low-LET radiation. Both these predictions are confirmed by a large number of experimental studies of cancer induction in animals, as well as studies using other models. Negative curvature of the dose response for cancer risk, ascribable to the competing effect of cell inactivation, is apparent at lower doses of high-LET radiation than it is for low-LET radiation.

Confirmation from epidemiological data of these experimental findings is difficult to obtain because there are far fewer populations that can be studied for the effects of exposure to high-LET radiation. It is clear, however, from studies of patient populations injected with alpha-emitting bone seekers or with thorium-containing contrast media that accumulate in the liver, and from studies of uranium miners and other miners whose bronchial epithelia were exposed to alpha irradiation from radon daughter products, that risks of cancers of the bone, liver, and lung, respectively are far higher on a dose-specific basis than has been observed in populations exposed to similar doses of low-LET radiation (Committee on the Biological Effects of Ionizing Radiation 1980).

Land

-6-

Implications of DNA repair

Two ionizing events may occur at the same locus, but be separated in time. Provided that DNA repair removes the effect, if any, of the first event by the time the second occurs, the two events cannot interact (NCRP 1980; Upton 1982b). Thus the same reasoning that would predict upward curvature of the dose-response curve for effects of low-LET radiation and for high-LET radiation a greater effectiveness per rad, and less curvature, would also predict a decreased effectiveness for low-LET radiation that is spread out over time by fractionation or protraction of exposure. The same phenomenon should not be seen for high-LET radiation, because DNA repair would be far less likely to prevent the interaction of ionizing events. The effectiveness of high-LET radiation might well increase with fractionation or protraction, especially at high dose levels, because cells killed early in the exposure sequence, and therefore incapable of carcinogenic development, might be replaced by other cells which would be targets for later irradiation. For both types of radiation, fractionation or protraction of dose should reduce curvature, both positive and negative, of the dose-response curve; that is, with increasing separation in time, the carcinogenic response should increasingly resemble the summation of effects of separate, independent exposures.

Breast cancer as a possible exception to the general rule

Experimental data tend to confirm the above predictions (NCRP 1984; Upton 1982a; United Nations Scientific Committee on the Effects of Atomic Radiation 1977). Mammary cancer appears to be somewhat of an exception, both experimentally and epidemiologically. In a reanalysis in parallel of data from the Japanese atomic bomb survivors, who each received a single, acute exposure, a group of American women who received several therapeutic x-ray exposures,

Land

-7-

separated by a few days or weeks, for acute postpartum mastitis, and another group of tuberculosis patients who typically received 100 low-dose exposures, separated by about two weeks, during the course of pneumothorax therapy, it was found that similar risks were associated with similar accumulated doses, in women irradiated at similar ages and followed over similar periods of time after exposure (Land, Boice, Shore et al. 1980). Apparently, fractionation of dose had little or no effect on risk. This finding is consistent with the model only if multi-track events are no more effective than single-track events, in which case the dose-response curve should be approximately linear. In fact, the US fluoroscopy data (which should be linear in any case) and the mastitis data are consistent with linearity, while the atomic bomb survivor data are strongly linear in the sense that only a small amount of upward curvature is consistent with the data.

The analysis of three breast cancer series just described above can be viewed as supportive of the general model, since the apparent departure from the usual case is consistent with the model (no fractionation effect is equivalent to linearity). Another breast cancer study, recently reported, is inconsistent with other data and is difficult to reconcile with the model. In this large study of breast cancer mortality among Canadian women previously given multiple chest fluoroscopies during treatment for tuberculosis, the dose response has marked upward curvature (Howe 1984). This result is inconsistent with the atomic bomb survivor data in particular, and can be reconciled with the general model only by assigning what seems to be undue importance to interactions between unrepaired damage from single ionizing events and damage from later events. There are other explanations possible, such as a

Land

-8-

discrepancy between the dosimetry used for the data from Nova Scotia, where doses were high, and the rest of the population, for which doses tended to be much lower. It is clear, however, that the context of the model gives this apparent discrepancy between different data sets far more etiological significance than it would have otherwise.

Partial-organ exposure

A consequence of the general, non-linear dose-response model is that partial-body, or partial-organ, exposures may be expected to result in dose-response relationships different from what would be predicted if the same exposure were evenly distributed over the target organ. Suppose, for example, that an exposure results in a 500 rad dose to one-fifth of the active bone marrow, and essentially none to the remaining four-fifths. Assuming that any radiation-induced leukemia will arise in the irradiated bone marrow, and using the BEIR model, the excess risk should be the average, weighted by the proportions of active marrow, of the risks computed separately for the unirradiated four-fifths and the irradiated fifth; that is, risk should be proportion to

$$(.8)(0) + (.2)(500+250,000/116) = 531.$$

Applying the model to the 100-rad dose averaged over the entire bone marrow yields a risk that is lower by a factor of 3:

$$100 + 10,000/116 = 186.$$

As another example, suppose that one-fourth of the bone marrow receives 800 rad, and the remaining three-fourths receives nothing. The average dose over the entire bone marrow is 200 rad, but the dose over the irradiated fourth is so high that the cell-killing effect of radiation is substantial to predominate. Thus the leukemia risk should be somewhat less than that

Land

-9-

calculated assuming a uniform dose of 200 rad over the entire bone marrow (Figure 3). This is probably why studies of leukemia risk in patients given radiation therapy for ankylosing spondylitis (Smith and Doll 1982) or cervical cancer (Boice, Day, Anderson et al 1984), with markedly uneven exposures to the active bone marrow, have failed to exhibit dose-response patterns consistent with a smoothly increasing risk with increasing dose. Fortunately, however, detailed partial-organ dosimetry is often feasible from exposure records for therapeutic irradiation, and it is reasonable to expect that, by explicitly taking account of the distribution of dose considerably more information can be obtained from these data.

Variation by time after exposure

The preceding discussion has been concerned with how cancer risk, averaged over some fixed period of time following exposure to radiation, depends upon various aspects of that exposure. The variation of risk over time within the period of observation, and afterward, is also of interest for both practical and theoretical reasons. Here the epidemiological approach has definite advantages over the experimental approach. Radiobiological experiments are usually designed for maximal efficiency with respect to some question that has little to do with time to tumor. Thus there may be little relevant information to be recovered from a reanalysis of old data; also, there are certain analytical artifacts that may lead to misleading results with respect to time (Guess and Hoel, 1977). The exposures of irradiated human populations, on the other hand, did not occur to suit the requirements of specific epidemiological studies, and often there is much to be gained by reworking old data. Also, population statistics contain a great deal of information on the variation over time of cancer risk in non-irradiated populations; thus there is a standard

Land

-10-

with which to compare the time dependence of risk in irradiated populations. Finally, the Japanese atomic bomb survivor data are especially informative about time to response because exposure was to a single event, in 1945, and subgroups of similar age at exposure, but with very different exposures and, therefore, very different proportions of radiation-induced to baseline cancer risk have been followed, in the same way, for nearly four decades.

Site-specific comparisons of time to tumor in heavily-irradiated and lightly-irradiated populations reveal two very different temporal patterns of risk (Land and Tokunaga 1984; Ad Hoc Working Group to Develop Radioepidemiological Tables 1985). The first, which characterizes bone sarcoma in a series of German patients injected with the short-lived isotope RA-224 and also characterizes leukemia among atomic bomb survivors, is wave-like, with an approximately lognormal distribution that is unrelated to the temporal pattern of risk in comparable, nonexposed populations (Mays and Spiess 1983). For acute leukemia the pattern varies by age at exposure, whereas for bone sarcoma and chronic granulocytic leukemia it does not. The second pattern, which characterizes female breast cancer among atomic bomb survivors and at least two medically-irradiated populations, and cancers of the lung and digestive system among atomic bomb survivors, is one of apparent congruity with respect to time of appearance for cancers occurring in heavily-irradiated and non-irradiated subpopulations of similar age. There are more cancers in the heavily irradiated populations but the excess cancers do not occur until ages at which baseline cancers occur, and their distribution subsequently is similar to that of the baseline cancers. The pattern thus corresponds, except for an initial period during which there can be no excess risk because any induced tumors require time in which to reach detectable size, to a model in which excess risk as a percentage of age-specific baseline risk, is constant over time following exposure.

Land

-11-

In its simplest form, the first pattern corresponds to a model in which the process of cell transformation is essentially complete soon after irradiation, and the lognormal temporal distribution may reflect only tumor growth kinetics (Steel 1977). The second pattern suggests a model in which cell transformation must await subsequent, age-related events which also influence the appearance time of cancers not related to radiation exposure. Since there is no difference between radiation-induced and other cancers with respect to time after exposure, there is no possibility for a dose effect on time to tumor given the second pattern. For cancers following the first pattern, radiation-induced cancers may occur earlier than other cancers, but there seems to be no biological reason why a cancer caused by 200 rad should occur earlier than one caused by 20 rad.

Dependence of risk on host factors

Generally, excess risk, as a percent of baseline risk, declined with increasing age at exposure. For those cancer sites, like the female breast for which relative risk remains approximately constant over time after exposure, and for which baseline risk increases with increasing age, the pattern with respect to exposure age implies a greater lifetime risk for exposure at younger ages. There are, however, considerable differences with respect to cancer site. Lung cancer and leukemia, for example, appear to be increased by radiation exposure after age 50, whereas there is little evidence of an excess breast cancer risk for exposures after about age 40 (Howe 1984; Tokunaga, Land, Yamamoto et al 1984). The influence of age on many radiation-induced cancers is therefore two-fold: sensitivity to radiation carcinogenesis varies according to age at exposure, and if a cancer is induced, its time of appearance (diagnosis) depends strongly upon attained age following exposure.

Land

-12-

Interestingly, there is no evidence so far for an excess of cancers of the lung and digestive organs among atomic bomb survivors exposed during childhood (Kato and Schull 1982). This is entirely consistent with the apparent temporal patterns for these sites. It is only very recently that an excess of breast cancer was observed, when the youngest cohorts reached ages at which breast cancer risk is normally increased in non-irradiated populations (Tokunaga, Land, Yamamoto et al 1982; Hildreth, Shore and Hempelmann 1983).

Except for those cancer sites, like the thyroid gland or the breast, for which baseline risk differences between the sexes seem intrinsic, and not caused by exposure factors like smoking, there is little evidence that men and women differ in their sensitivity to radiation carcinogenesis (Committee on the Biological Effects of Ionizing Radiation 1980).

Possible influences on radiation carcinogenesis by host factors other than age and sex are intriguing in their implications for etiology. For example, excess skin cancer risk among persons treated by X ray for ringworm of the scalp during childhood has tended to be greater among light-complected persons, and was not seen at all among blacks (Shore, Albert and Pasternack 1976). These findings are strongly indicative of a synergistic interaction between non-ionizing ultraviolet radiation and ionizing radiation, because the cancers observed tended to be restricted to irradiated areas not protected from sunlight by hair or clothing.

Childhood immigrants to Israel from North Africa who were treated by x-ray for ringworm of the scalp had higher excess thyroid cancer risk than native Israelis and immigrants from other areas who received the same treatment (Ron and Modan 1984). Although other, more mundane explanations have not been ruled out, one interesting possibility that has been suggested is that the

Land

13-

difference may reflect greater sensitivity of heterozygotes for the genetic disease, ataxis telangiectasia, which is known to be more prevalent among North African Jews than among other groups (Ron and Modan 1984). A finding that conceivably might be related is a disproportionate excess of thyroid cancer among Jews as compared to non-Jews given X-ray therapy during infancy for what was at the time thought to be enlarged thymus gland (Hempelman, Hall, Phillips et al 1975).

Although there is a five-fold difference in breast cancer risk between the United States and Japan, an analysis in parallel of data from irradiated Japanese and American populations found essentially no difference in absolute excess risk per rad, for women of similar ages at irradiation followed up for similar period of time after exposure (Land 1980). This finding suggests that whatever causes the difference between countries interacts with radiation exposure in an additive manner. Studies of breast cancer risk among Japanese immigrants and their descendants suggest that the factors responsible for the difference between the two countries are primarily environmental, rather than genetic (Haenszel and Kurihara 1968).

Variation with exposure to other risk factors

In general, the existence of large irradiated populations, like the Japanese atomic bomb survivors, with heavily irradiated subpopulations in which a high percent of cancers of certain sites can be assumed to be causally related to radiation exposure, and with appropriate comparison populations in which radiation can be assumed to be a minor factor, provides opportunities for case-control or other approaches in which high-dose cancer cases and non-cases are compared with low-dose cases and non-cases, with respect to host factors or exposures other than to radiation. Here the question of interest is whether

Land

-14-

the difference between high-dose cases and controls with respect to some other factor is the same as the difference between low-dose cases and controls. Clearly this approach, to be successful, requires that there be a great deal of information in the population under study concerning both radiation and the other factor of interest.

One particularly interesting question concerns the interaction of radiation and cigarette smoking in the causation of lung cancer. A recent case-control interview study of Japanese atomic bomb survivors found an additive or sub-additive relationship between smoking and radiation dose (Blot, Akiba and Kato 1984). Analyses of data on US uranium miners, on the other hand, find a multiplicative relationship (Archer, Wagoner and Lundin 1973; Whittemore and McMillan 1983), while different studies of radon exposure and smoking among Swedish miners of other ores support either additivity (Axelson and Sundell 1978; Edling 1982; Radford and Renard 1984) or multiplicativity (Damber and Larsson 1982; Larsson and St. Clair Damber 1982). It may not be necessarily contradictory for the penetrating, low-LET gamma radiation from the Hiroshima and Nagasaki bombs to interact with smoking differently than the non-penetrating, high-LET radiation from inhaled radon daughters, but it clearly would be helpful to have information on the biological plausibility of such a difference. This is a problem for which extremely thorough comparative pathological reviews might provide useful information.

Discussion

The above discussion treats radiation carcinogenesis as something that can be factored into component parts, to be dealt with separately and then combined again for risk assessment. It is assumed, for example, that the shape of the dose-response curve does not change over time after exposure, although the

Land

-15-

magnitude of risk may change. Thus one can estimate dose response in terms of average excess risk over the period, say, five to thirty five years after exposure, and separately compare time to tumor among subpopulations with little or no radiation (and therefore very little radiation-induced cancer) and with a great deal of exposure (and therefore with mostly radiation-induced cancers, depending upon site). That is, assuming the two aspects of risk to be orthogonal. Likewise, the shape of the dose-response curve is usually assumed to be the same for exposures at different ages, although the level of risk for any fixed dose may depend strongly upon exposure age.

It seems that this is a reasonable approach, at least at first. There are many more complex solutions than there are simple ones, so it is better to investigate the simple possibilities first. Perhaps one of the simple solutions will turn out to fit. Also, it is usually better to have the answer to a simple problem that only approximates the truth than to have no answer to the true problem. Working on simple possibilities may provide clues to which one of the infinitely many more complex possibilities is correct.

At the same time, however, there is evidence that argues against simplicity. There are experimental indications that the presence of a tumor promoter may, for example, alter the shape of the dose response to ionizing radiation (Fry 1985). Also, the apparent dependence upon exposure age of the time-to-response distribution of radiation-induced acute leukemias corresponds neither to the simple growth-kinetics model of bone sarcomas induced by Ra-224 or OGL induced by external gamma ray, nor to the constant relative risk model that holds approximately for radiation-induced lung cancer or breast cancer. However, plausible theoretical explanations that have testable consequences are needed.

Summary

The epidemiology of radiation carcinogenesis is a fairly well developed field, largely because of the theoretical structure provided by experimental radiobiology. The chief result of this continuing development is that there are now more interesting questions to be answered than ever before.

Land

-17-

References

- Ad Hoc Working Group to Develop Radioepidemiological Tables, (1985) Report of National Institutes of Health Ad Hoc Working Group to Develop Radioepidemiological Tables. Office of the Director, National Institutes of Health, U.S. Dept. of Health and Human Services, Washington, D.C., 1985 (NIH Publication No. 85-2748).
- Archer, V.E., Wagner, J.K. and Lundin, F.E. (1973) Uranium mining and cigarette smoking: Effects on man. *J. Occup. Med.* 15:204-11.
- Axelsson, O. and Sundell, L. (1978) Mining, lung cancer and smoking, *Scand. J. Work Environ. Health* 4:46-52.
- Blot, W.J., Akiba, S. and Kato, H. (1984) Ionizing radiation and lung cancer: A review including preliminary results from a case-control study among atomic bomb survivors. In, *Atomic Bomb Survival Data: Utilization and Analysis* (eds. R.L. Prentice and D.J. Thompson) pp 235-48. Society for Industrial and Applied Mathematics, Philadelphia.
- Boice, J.D., Day, N.E., Anderson, A. et al (1984) Cancer risk following radiotherapy of cervical cancer. In, *Radiation Carcinogenesis: Epidemiology and Biological Significance* (eds. J.D. Boice Jr. and J.F. Fraumeni Jr.) pp 161-80. Raven Press, New York.
- Committee on the Biological Effects of Ionizing Radiation (1980) *The Effects on Populations of Exposure to Low Levels of Ionizing Radiation: 1980*. National Academy Press, Washington, D.C..
- Damber, L. and Larsson, L.G. (1982) Combined effects of mining and smoking in the causation of lung carcinoma. *Acta Radiol. Oncol.* 21:305-13.
- Edling, C. (1982) Lung cancer and smoking in a group of iron ore miners, *Am. J. Ind. Med.* 3:191-99.
- Fry, R. J. M. (1985) Personal Communication.
- Guess, H. A. and Hoel, D.G. (1977) The effect of dose on cancer latency period. *J. Environ. Pathol. Toxicol.* 1:279-86.
- Haenszel, W. and Kurihara, M. (1968) Studies of Japanese migrants. I. Mortality from cancer and other diseases among Japanese in the United States. *J. Natl. Cancer Inst.* 40:43-68.
- Hempelmann, L.H., Hall, W.J., Phillips, M., Cooper, R.A. and Ames, W.R. (1975) Neoplasms in persons treated with x-rays in infancy: Fourth survey in 20 years. *J. Natl. Cancer Inst.* 55:519-30.
- Hildreth, M.G., Shore, R.E. and Hempelmann, L.H. (1983) Risk of breast cancer among women receiving radiation treatment in infancy for thymic enlargement. *Lancet* 2:273.

Land

-18-

- Howe, G.R. (1984) The epidemiology of radiogenic breast cancer. In, Radiation Carcinogenesis: Epidemiology and Biological Significance: (eds. JD Boice, Jr. and JF Fraumeni Jr.) pp 119-29. Raven Press, New York.
- Kato, H. and Schull, W.J. (1982) Studies of the mortality of atomic bomb survivors. 7. Mortality, 1950-1978. Part I. Cancer Mortality. Radiat. Res. 90:395-432.
- Land, C.E. (1980) Estimating cancer risk from low doses of ionizing radiation. Science 209:1197-203.
- Land, C.E., Boice, J.D. Jr., Shore, R.E., Norman, J.E. and Tokunaga, M. (1980) Breast cancer risk from low-dose exposures to ionizing radiations: Results of parallel analysis of three exposed populations of women. J. Natl. Cancer Inst. 65:353-76.
- Land, C.E. and Tokunaga, M. (1984) Induction period. In, Radiation Carcinogenesis: Epidemiology and Biological Significance (eds. JD Boice Jr and JF Fraumeni Jr.), pp 421-36. Raven Press, New York.
- Larsson, L.G. and Damber, L. (1982) Interaction between underground mining and smoking in the causation of lung cancer: A study of non-uranium miners in Northern Sweden, Cancer Detect. Prev. 5:385-89.
- Mays, C.W. and Spiess, H. (1983) Epidemiological studies of German patients with Ra-224. In, Epidemiology Applied to Health Physics (Proceedings of the Sixteenth Midyear Topical Meeting, Health Physics Society, Albuquerque, NM. January 9-13, 1983), pp 159-66. National Technical Information Service CONF-830101.
- NCRP Report No. 64 (1980) Influence of dose and its distribution in time on dose-response relationship for low-LET radiations. National Council on radiation protection and measurements, Washington, D.C.
- NCRP Report No. 77 (1984) Exposure from the uranium series with emphasis on radon and its daughters. National Council on Radiation Protection and Measurement, Washington, D.C., 1984.
- Radford, E.P. and St. Clair Renard, K.G. (1984) Lung cancer in Swedish iron miners exposed to low doses of radon daughters. N. Engl. J. Med. 310:1485-94.
- Ron, E. and Modan, B. (1984) Thyroid and other neoplasms following childhood scalp irradiation. In, Radiation Carcinogenesis: Epidemiology and Biological Significance (eds. J D. Boice Jr. and J.F. Fraumeni Jr.), pp 139-51. Raven Press, New York.
- Shore, R.E., Albert, R.E. and Pasternack, E.S. (1976) Follow-up study of patients treated by x-ray epilation for tinea capitis: Resurvey of post-treatment illness and mortality experience. Arch. Environ. Health 31:17-24.

Land

-19-

- Smith, P.G. and Doll, R. (1982) Mortality among patients with ankylosing spondylitis after a single treatment course with x-rays. Br. Med. J. 284:449-60.
- Steel, G.G. (1977) Growth Kinetics of Tumors. Clarendon Press, Oxford.
- Tokunaga, M., Land, C.E., Yamamoto, T., Asano, M., Tokuoka, S., Ezaki, H. and Nishimori I (1982) Breast cancer in Japanese atomic bomb survivors. Lancet 2:924.
- Tokunaga, M., Land, C.E., Yamamoto, T., Asano, M., Tokuoka, S., Ezaki, H., Nishimori, I. and Fujikura, T. (1984) Breast cancer among atomic bomb survivors. In, Radiation Carcinogenesis: Epidemiology and Biological Significance (eds. J.D. Boice Jr. and J.F. Fraumeni Jr.) pp 45-56. Raven Press, New York.
- United Nations Scientific Committee on the Effects of Atomic Radiation (1977) Sources and Effects of Ionizing Radiation. United Nations, New York, (Report E.77.IX.1).
- Upton, A.C. (1982a) The biological effects of low-level ionizing radiation. Sci. Am. 246:41-9.
- Upton, A.C. (1982b) Role of DNA damage in radiation and chemical carcinogenesis. In, Environmental Mutagens and Carcinogens. Proceedings of the 3rd International Conference on Environmental Mutagens (eds. R. Sugimura, S. Kondo and H. Takebe), pp 71-80. University of Tokyo Press, Tokyo/Alan R. Liss, Inc., New York.
- Whittemore, A.S. and McMillan, A. (1983) Lung cancer mortality among U.S. uranium miners: A reappraisal. J. Natl. Cancer. Inst 71:489-99.

Legends for Figures

Figure 1: Schematic diagram of ionizing events along high-let and low-LET radiation tracks.

Source: Upton (1982a).

Figure 2: An example of a linear-quadratic dose-response curve for radiation carcinogenesis, incorporating a negative exponential in radiation dose to represent the competing effect of cell sterilization:

$$\text{Risk} = (a_0 + a_1D + a_2D^2) \exp(-b_1D).$$

Figure 3: An illustration, using the dose-response curve of Figure 2, of risk calculation for partial-body irradiation. If 25% of the active bone marrow receives a uniform dose of 800 rad, and the remaining 75% is not exposed, the estimated leukemia risk (25% of that from a whole-body dose of 800 rad plus 75% of the baseline risk) is less than that from a uniform, whole-body dose giving the same average dose (200 rad) to active marrow.

FIGURE 1

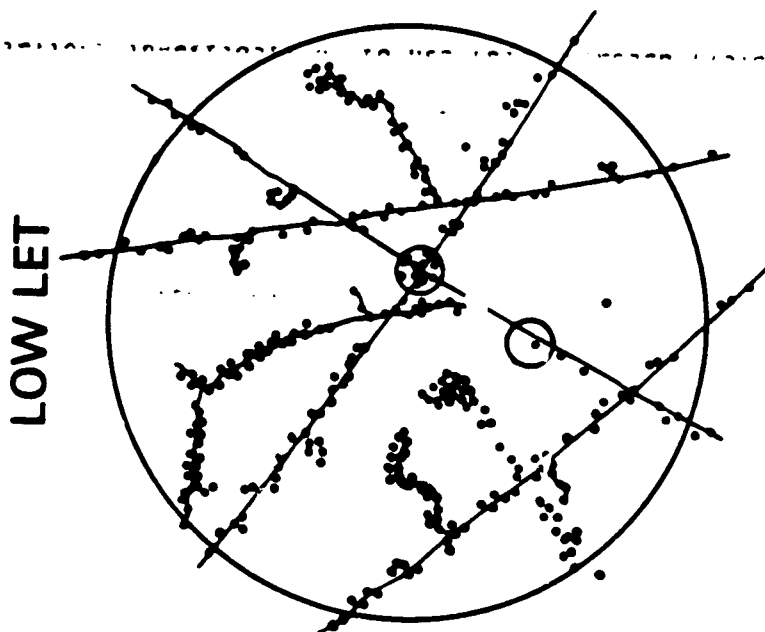
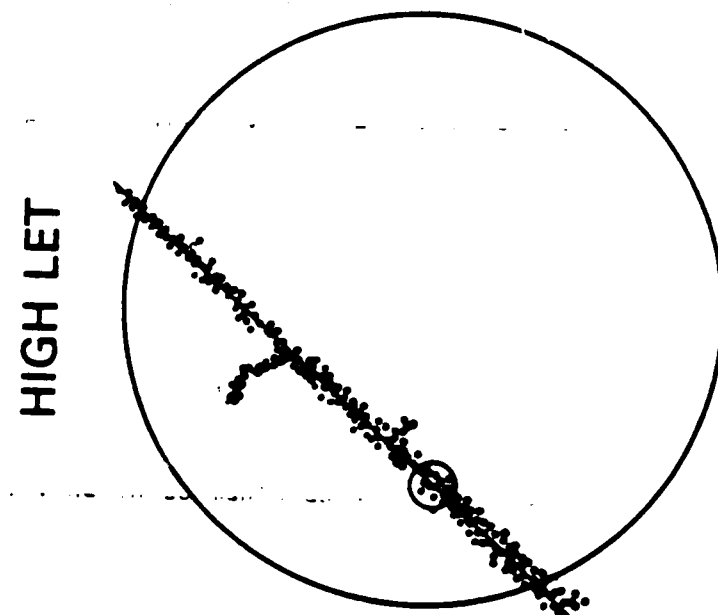


FIGURE 3

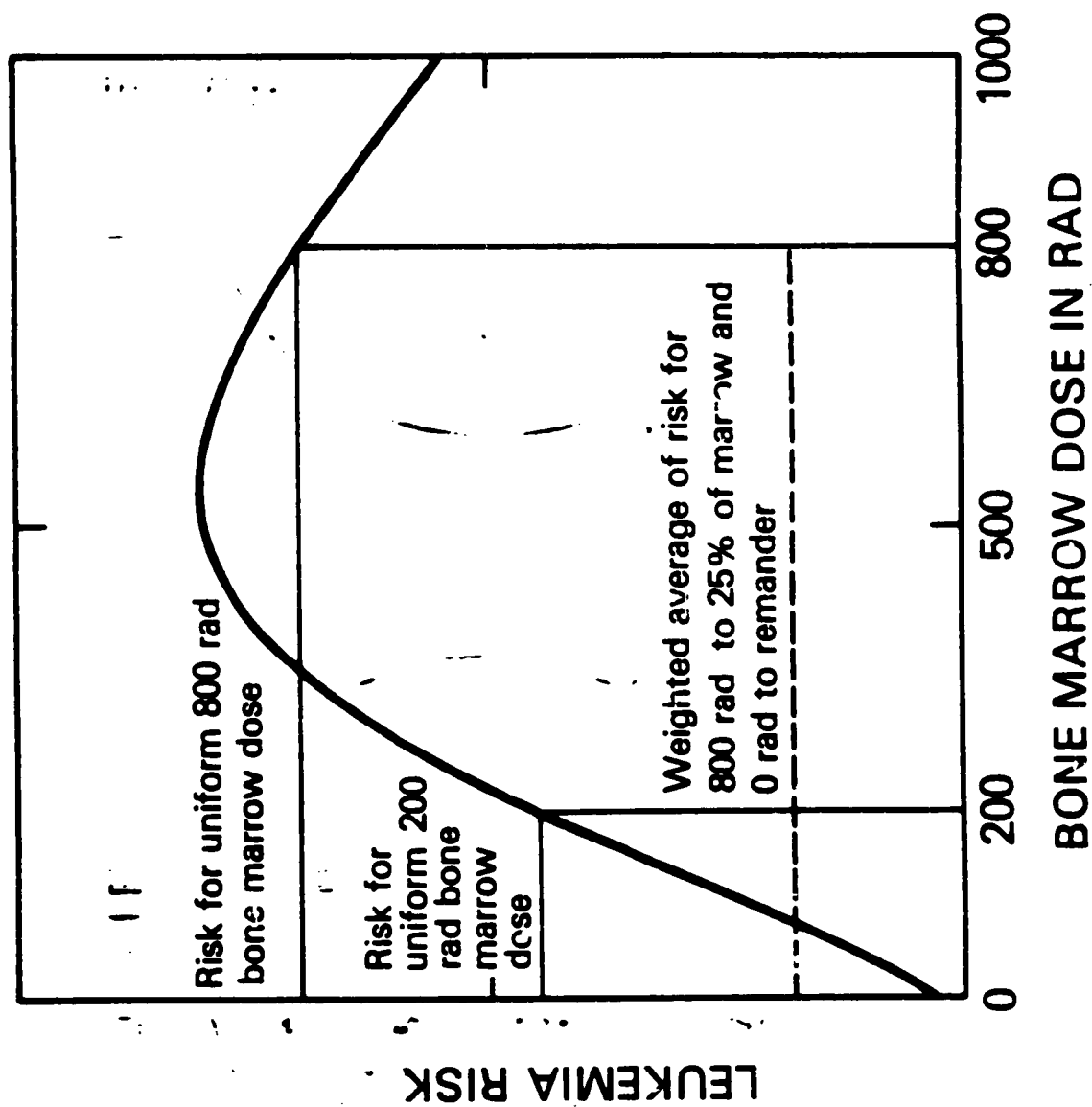
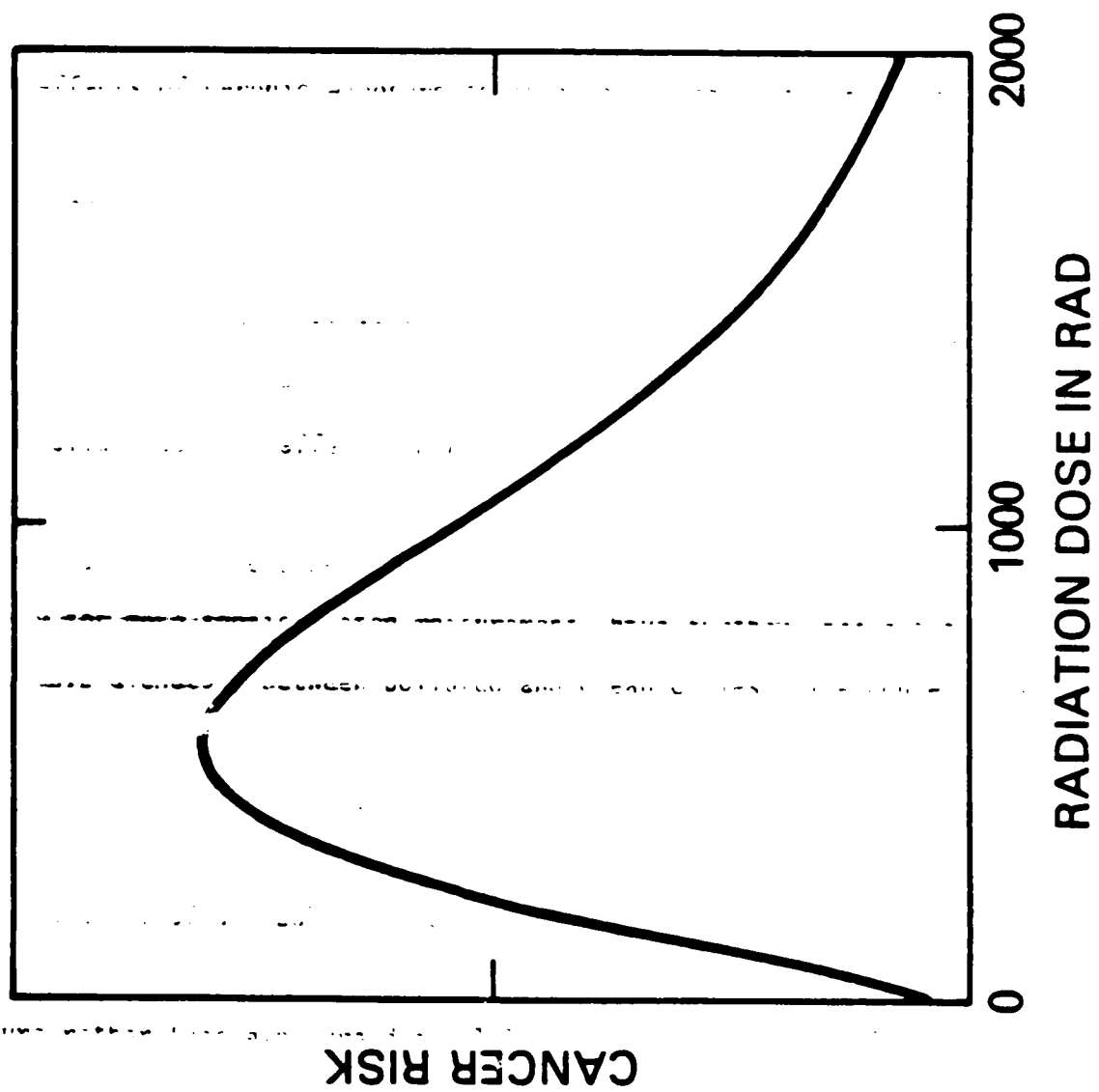


FIGURE 2



Chapter 24

SOFT WATER -- HARD ARTERIES:
An Interpretation of Ecologic Findings

George W. Constock

SOFT WATER — HARD ARTERIES:

An Interpretation of Ecologic Findings

George W. Comstock

Ecologic findings are derived from studies in which the units are populations, as opposed to epidemiologic studies based on individual persons. Ecologic studies are largely responsible for the aphorism in the title of this chapter since a large number of them have found that areas with soft water tend to have high death rates from cardiovascular diseases, especially coronary heart disease and strokes (Comstock 1979).

Interest in the possible association between composition of drinking water and cardiovascular disease started on opposite sides of the globe during the 1950's. In the United States, a pattern of high cardiovascular death rates in the South Atlantic states and low rates in the central portion of the country began to emerge (Enterline and Stewart 1956; Sauer 1980). These studies showed that the risk of coronary heart disease associated with where one lives is not trivial. Table 1 compares relative risks of dying from coronary heart disease among white males in the state and socio-economic area with the highest rates compared to the state and socio-economic area with the lowest rates (Sauer 1980). For comparison, Table 1 also shows the relative risks of a first myocardial infarction for a number of major risk factors among white male participants in several large prospective studies whose results were combined in the Pooling Project (The Pooling Project Research Group 1978). Residence clearly qualifies as a major risk factor. Among the more tenable hypotheses for these geographic differences is the possible effect of water composition on cardiovascular disease.

A more direct implication of water came from Japan, where an agricultural chemist named Kobayashi had reported in an obscure journal that death rates from stroke were highly correlated with acidity of the local river water (Kobayashi 1957). Kobayashi's work would probably have drawn little if any attention in medical circles if an American investigator, Schroeder, had not been given a grant to study cardiovascular disease in the Orient. Impressed by Kobayashi's findings, Schroeder sought to replicate this work when he returned to the United States. Because he had difficulty in obtaining information on water acidity, he looked at the association of water hardness and cardiovascular death rates in states and metropolitan areas (Schroeder 1960). He found highly significant negative correlations, -0.33 for 45-64 year old white males, and -0.24 for 45-64 year old white females, which suggested that cardiovascular death rates were related either to something harmful in soft water or something protective in hard water. Schroeder's findings, particularly those relating to coronary heart disease, were confirmed by nearly all nationwide ecologic studies, both here and abroad. The risk of dying from coronary heart disease appeared to be approximately fifteen percent higher in communities with soft water than in communities with hard water.

Not all studies agreed, however. The most recent nationwide ecologic study, summarized in Table 2, suggested that the relative risk associated with soft water is now closer to five than to fifteen percent (Sharrett, Morin, Fabsitz et al. 1984). Studies looking at variations between smaller geographic units, such as pairs of cities, have been much less likely to confirm the association. Only three epidemiologic studies of individuals have been done. One was based on a follow-up of over 30,000 adults for a period of twelve years in Washington County, MD, an area with marked variation in water hardness

(Comstock, Cauthen and Helsing 1980); a second was a case-control study among 1400 farmers in Wisconsin (Zeighami, Morris and Calle 1983); and a third was a case-control study involving approximately 12,000 persons that examined the risks associated with artificially softened water in two cities of Canada (Hewitt and Neri 1980). None of the three was able to confirm the risk associated with soft water that had been found in the nationwide ecologic studies.

These discrepant findings naturally raised questions. First, a fundamental one: Why should so many studies have been done on this particular association of exposure and disease for which there is very little biologic plausibility? Why are almost all studies on this topic ecologic in that they are based on comparisons of geographic groups rather than following the usual epidemiologic design of considering individuals as the study units? What are the problems with ecologic studies? Under what circumstances is their use appropriate?

The answer to the first question is that the search for a water factor in the etiology of coronary heart disease does have considerable appeal. If the correction of that factor could prevent even a small percentage of cases, its impact could be important. Coronary heart disease affects so many people in soft water areas that even a small reduction in risk could prevent many cases each year. Characteristics of water for drinking are routinely altered by treatment plants, so that the addition of a beneficial trace substance or the removal of a harmful one is very likely to be feasible. Finally, applying preventive measures at the community level has usually been much more effective than attempts to change individual lifestyles.

Ecologic studies also have considerable appeal. In the first place, they are inexpensive. The outcomes of interest are usually death rates from specified diseases which are readily available from routinely collected data,

Comstock

often in published form. In the case of municipalities, water hardness and many other water characteristics are also available either in published form or from the treatment plants. For most residents of most metropolitan areas, water furnished to their homes and to their places of employment will be the same. In the analysis of ecologic data, allowances can be made for the effects of other community variables such as altitude, latitude, climate, and community indices of socio-economic status such as per cent of the population with a certain level of education and per cent unemployed.

Investigators sometimes turn to ecologic studies because the alternatives, epidemiologic studies of individuals, are often not feasible, especially within a given community. If the municipal water is soft and comes from a single source, there will be no exposure gradient within the community that can be studied. If the water is hard, comparisons can be made between persons drinking hard water and those drinking water treated by home water softeners. The composition of artificially softened water, however, is not the same as that of naturally soft water. The most feasible epidemiologic design appears to be a case-control study in an area large enough to have water supplies with diverse characteristics, but this would entail considerably more expense than an ecologic study.

Unfortunately, ecologic findings are often unreliable indicators of cause-and-effect relationships. Exposures measured at the community level may not represent the exposures of individuals. For example, in studies of air pollution, it is likely that residents of the most heavily polluted areas spend more time indoors than residents of less polluted areas. A possible explanation of the diminished association of coronary heart disease deaths with soft water in 1970 compared to earlier decades could be the increasing use of water softeners in hard water areas. In Washington County, MD, twenty seven

Comstock

percent of the homes in the area with the hardest water had water softeners in 1975, a considerable increase over earlier years (Comstock, Cauthen and Helsing 1980). In many hard water areas, therefore, a large proportion of the residents may now be drinking soft water. Furthermore, it is possible that relatively little of the daily fluid intake of many persons comes from plain water from the home water tap, as suggested by a small survey conducted in Washington County, Maryland. For all these reasons, water quality data for a community may be only partially relevant to the fluid intake of its inhabitants.

In addition to the always worrisome possibility in an ecologic study that the exposure measured for the community may not apply to the individuals at risk, other risk factors must also be considered. The major known risk factors for coronary heart disease are advancing age, white race, male sex, cigarette smoking, high blood pressure, and elevated serum cholesterol. The first three factors — age, race, and sex — can be taken into account in ecologic analyses by using disease or death rates standardized for age, race, and sex. The frequency of the latter three risk factors — smoking, blood pressure, and serum cholesterol — are not normally available on an area-wide basis and hence cannot be used in ecologic studies. In Washington County, where there is marked variation in water hardness, a private health census did make it possible to know the frequency of cigarette smoking among adults in the different water hardness areas (Comstock 1980). As shown in Table 3, there was a definite tendency for cigarette smoking to be more common in the areas with softer water. In Great Britain, preliminary results from a study specifically designed to investigate the water story have shown that all three personal risk factors — smoking, blood pressure and serum cholesterol — were also

associated with soft water in the samples of adult men in the twenty four towns selected for study (Pocock, Shaper and Powell 1985). In an ordinary ecologic study, if the distribution of risk factors among individuals happens to parallel the distribution of the exposure of interest, a spurious association with the exposure can be created. The opposite effect can also occur where the individual risk factors are distributed opposite to the distribution of the factor of interest, water hardness in this case. Only rarely in an ecologic study is it possible to take any account of the confounding effects of individual risk factors.

Because of the ease with which ecologic analyses can be done, it is not difficult to find significant correlations. The World Almanac lists crime rates for a number of American cities for which it was possible to find data on the hardness of their water (Delury 1976; Feinleib, Fabsitz and Sharrett 1979). Of the six categories of crime listed, that of crimes against property showed an impressive and statistically significant association with water hardness ($p = .05$) (Figure 1). This is not a claim that hard water is like hard liquor in its effect on one's attitude toward the property of others. Figure 1 does illustrate, however, how easy it is to find a statistically significant ecologic association.

A further difficulty with most ecologic studies on water quality and cardiovascular disease is that most results have been reported in terms of correlation coefficients. As a specific example, the results from an ecologic study of trace metals in the water supplies of 99 of the 100 largest U.S. cities are available (Voors 1971). Among white males, the mortality rate from coronary heart disease showed a correlation coefficient of -0.35 with calcium and -0.43 with lithium concentrations in the municipal water. The negative

sign is meaningful, telling us that as the content of either of these metals in the water went up, death rates from coronary heart disease went down.

Does the larger coefficient for lithium signify that it has a more important association with coronary heart disease than calcium? Not necessarily. To measure the importance or strength of an association, one needs the regression coefficient. In Figure 2, two hypothetical situations are shown. Each dot represents the death rate and water hardness for an imaginary community. The lines are regression lines. Their slopes, measured by the regression coefficient, indicate the average amount of change in the death rate for each unit of change in water hardness. In both panels, the correlation and regression coefficients are positive, indicating that as water hardness increases, the hypothetical death rate also increases. In the upper panel, the correlation coefficient is $+0.60$, higher than most coefficients in studies of water quality. The regression coefficient, however, is only 0.03 , indicating that it takes 33 units of change in the water component for only one unit of change in the death rate. In the lower panel, the correlation coefficient is almost the same, $+0.59$. The regression coefficient, graphically illustrated by the slope of the line, is 0.44 , large enough to suggest that something interesting may be going on. In this instance, it takes only 2.3 units of change in water hardness for each unit of change in the death rate. In the evaluation of ecologic findings, the regression coefficient is usually more informative than the correlation coefficient because the former indicates the strength or importance of the association.

With all these potential problems, is there a place for ecologic studies? The answer is a qualified yes. If one has a hypothesis to be tested, an ecologic analysis probably should be the first step whenever appropriate data is readily available. In judging an ecologic analysis, remember that

Comstock

-7-

associations that confirm hypotheses developed before the analysis is done are much less likely to be due to chance than those that just happen to turn up in the process of looking for something else. Ask that the statistical analysis include the regression coefficients. Better yet, ask to see a correlation diagram like those in figures 1 and 2. Consider carefully whether individual characteristics that could not be measured are likely to have affected the degree of association.

Even with these caveats, ecologic studies are feeble reeds on which to lean. With respect to taking action at the present time on the specific topic in the title of this chapter, soft water/hard arteries, permit a closing quotation by Pocock, Shaper and Powell who are currently engaged in major ecologic and epidemiologic studies of water composition and cardiovascular disease in Great Britain (Pocock, Shafer and Powell, 1985).

"The statistical evidence (for a causal role of soft water in cardiovascular disease) inevitably rest(s) on 'ecologic studies' which are well known as being a weak epidemiologic method which cannot by itself provide sound evidence of a causal mechanism. Hence it would be unwise to invest in large-scale water hardening programs, the full consequences of which are poorly understood. At the same time it would seem prudent not to undertake artificial softening of drinking water supplies below certain levels." In other words, in the absence of better evidence than ecologic studies can give us, don't rock the boat.

References

- Comstock, G. (1979) Water hardness and cardiovascular diseases. *Am. J. Epidemiol* 110:375-400.
- Comstock, G.W. (1980) The epidemiologic perspective: Water hardness and cardiovascular disease. *J. Environ. Pathol. Toxicol.* 4-2, 3:9-25.
- Comstock, G.W., Cauthen, G.M., Helsing, K.J. (1980) Water hardness at home and deaths from arteriosclerotic heart disease in Washington County, MD. *Am. J. Epidemiol.* 112:209-16.
- Delury, G.E. ed. (1976) *The World Almanac and Book of Facts.* New York, Newspaper Enterprise Association, Inc. p 975.
- Enterline, P.E. and Stewart, W.H. (1956) Geographic patterns in deaths from coronary heart disease. *Public Health Rep.* 71:849-55.
- Feinleib, M., Fabsitz, R. and Sharrett, A.R. (1979) Mortality from cardiovascular and non-cardiovascular diseases for U.S. cities. DHEW Publication No. (NIH) 79-1453. Washington, D.C., U.S. Government Printing Office.
- Hewitt, D. and Neri, L.C. (1980) Development of the "Water Story": Some recent Canadian studies. *J. Environ. Pathol. Toxicol.* 4-2, 3:51-63.
- Johns Hopkins Training Center for Public Health Research: Unpublished data.
- Kobayashi, J. (1957) On geographical relationship between the chemical nature of river water and death-rate from apoplexy. *Berichte des Ohara Instituts fur Landwirtschaftliche Biologie* 11:12-21.
- Pocock, S.J., Shaper, A.G. and Powell, P. (1985) The British Regional Heart Study: Cardiovascular disease and water quality. Presented at International Symposium on Geochemistry and Health, London 16 April 1985.

- The Pooling Project Research Group: (1978) Relationship of blood pressure, serum cholesterol, smoking habit, relative weight and ECG abnormalities to incidence of major coronary events: Final report of the pooling project. J. Chronic Dis. 31:201-306.
- Sauer, H.I. (1980) Geographic patterns in the risk of dying and associated factors, ages 35-74 years, United States, 1968-72. DHHS Publication No. (PHS) 80-1402. Vital and Health Statistics, Series 3, No. 18. Washington DC: U.S. Government Printing Office.
- Schroeder, H.A. (1960) Relationship between mortality from cardiovascular disease and treated water supplies. Variations in states and 163 largest municipalities of the United States. JAMA 172:1902-08.
- Sharrett, A.R., Morin, M.M., Fabsitz, R.R. and Bailey, K.R. (1984) Water hardness and cardiovascular mortality. In, Drinking Water and Human Health, (eds. J.A. Bell and T.C. Doege), Illinois, Am. Med. Assn., pp 57-68.
- Voors, A.W. (1971) Minerals in the municipal water and arteriosclerotic heart death. Am. J. Epidemiol. 93:259-66.
- Zeighami, E., Morris, M. and Calle, E. (1983) Risk of coronary artery disease, stroke, and presence of drinking water inorganics. Am. J. Epidemiol. 118:449.

Table 1

Relative risk of coronary heart disease
by Residence, 1968-72, White Males Aged 35-74 (deaths)*

South Carolina vs.
New Mexico 1.74

Area B in Alabama vs.
Area 1 in New Mexico 2.57

Examples of Relative Risk observed for Coronary Heart Disease
in the Pooling Project, 1964-74, White Males Aged 40-64, (new
cases)**

Worst vs. best quintile

Systolic blood pressure	2.14
Diastolic blood pressure	2.11
Serum cholesterol	1.88
Relative weight	1.24
20 cigarettes a day vs. never smoked	1.66

* Sauer, 1980

**Pooling Project Research Group, 1978

Table 3

Association of Major Risk Factors
with Water Hardness

<u>Risk Factor</u>	<u>Correlation Coefficient</u>
<u>Washington County, MD (6 areas):*</u>	
Cigarette smoking	-0.85
<u>Great Britain (24 towns):**</u>	
Cigarette smoking	-0.22
Systolic blood pressure	-0.13
Serum cholesterol	-0.33

* Comstock, 1980

** Pocock, Shaper and Powell, 1985

Table 2

Unadjusted and Adjusted Relative Risks of Mortality
in Soft Water (0 ppm) vs. hard water (200 ppm) areas

	<u>R.R.</u>
Unadjusted	1.08
Adjusted for:	
Region	1.05
Region, SES, Climate	1.04

Sharrett, Morin, Fabsitz, et al. 1984

Coronary Heart Disease and Water Hardness

94 SMSA's U.S., 1959-61, Whites

<u>Sex and Age</u>	<u>Correlation Coefficient</u>
White males	
45-64	-0.33
65+	-0.34
White females	
45-64	-0.24
65+	-0.30

Legends for figures:

Figure 1. Association of property crimes and water hardness in 19 selected U.S. cities, 1973, with regression line indicated.

Figure 2. Two hypothetical associations of death rates and water hardness, with regression lines indicated.

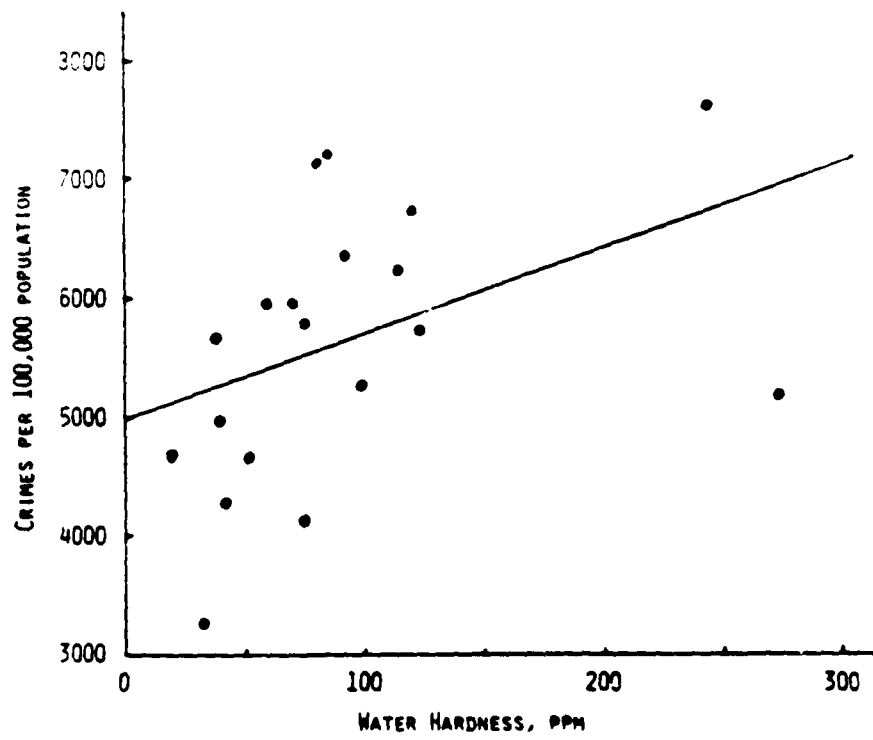


FIGURE 1

Comstock

-16-

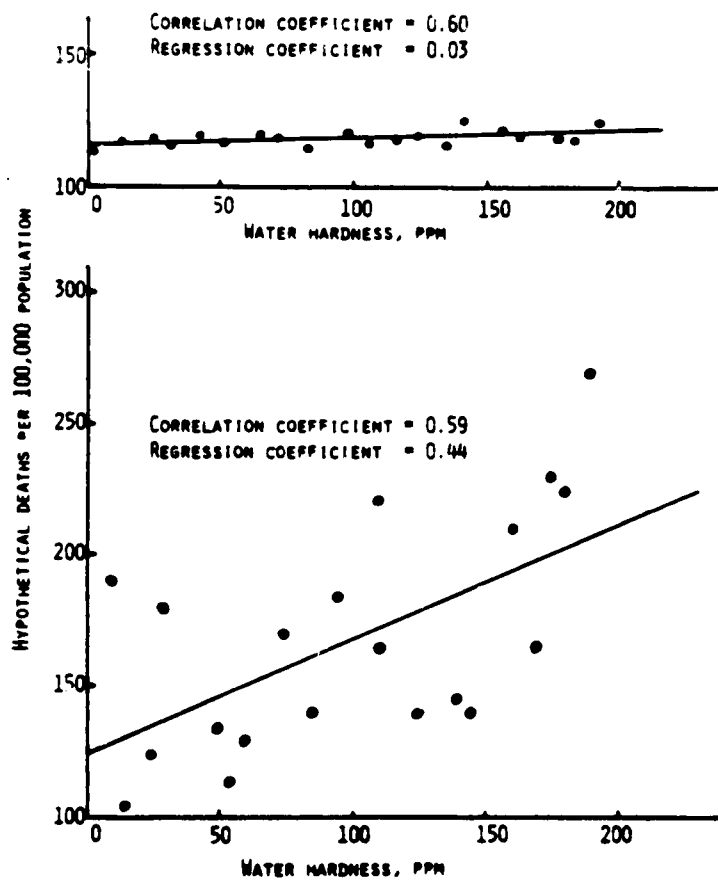


FIGURE 2

Chapter 25

Circumstances of Exposure and Reproductive Consequences

Jennie Kline and Zena Stein

Circumstances of Exposure and Reproductive Consequences

Jennie Kline and Zena Stein

Environmental exposures may have various of effects on human reproduction depending on the circumstances of exposure. The "circumstances of exposure" include not only who is exposed (the father, the mother, or the developing conceptus), but also the time, duration, level, and route of the exposure. Among possible consequences of exposure are infertility; lowered fecundability; chromosomal defect in the germ cells of either parent, of the zygote, or both; gene mutation, again either in the germ cells or zygote; spontaneous abortion of a chromosomally and genetically normal conceptus; later fetal death; premature delivery; intrauterine growth retardation; congenital malformations; developmental disability; and childhood morbidity and mortality. Even though the processes giving rise to these various outcomes are not always known, it is clear from the range of outcomes that more than one pathway must be invoked. In particular, outcomes will vary according to the timing of an exposure in relation to the time of development. In this chapter we explore the likely specificity between the circumstances of exposure and the consequences. The focus is first on exposures to the woman that precede conception, including those to which she may herself have been exposed in utero; and second, on exposures to the conceptus during gestation. Examples are selected mainly from the literature linking irradiation and cigarette smoking to reproduction.

Preconception exposure to women

We begin by considering the ways in which an environmental exposure to a woman before conception might adversely affect reproduction. At least five processes can be distinguished.

- 1) Those affecting the constitution of the woman in such a way that the conceptus will be distributed during gestation.
- 2) Those affecting the cytoplasm of the germ cell and its constituents, such as the mitochondria.
- 3) Those affecting the chromosomal material of the germ cell.
- 4) Those reducing the probability of conception through some other process, perhaps by killing the follicle.
- 5) Those affecting the probability of spontaneous abortion other than the processes listed above.

With respect to the first process, there is one well-known example. In a woman exposed to PCB the toxin is stored in the fatty tissue, to be mobilized in successive pregnancies until the stores are exhausted (Rogan, 1982). In this case, a single preconceptional exposure affects one or more pregnancies conceived subsequently.

The second process—factors that might affect the cytoplasm of the germ cell—is at this time a theoretical construct, little examined in the laboratory and not at all in humans.

Processes affecting the chromosomal material of the germ cell have received the most attention, although, as yet, no environmental exposure has been consistently associated with aneuploidy in the human conceptus (Kline and Stein 1985). In the female germ cells meiosis I, the critical process by which the pairs of chromosomes in diploid cells are separated, assorting themselves into cells that each contain a haploid complement of 23 chromosomes, begins in early fetal life. The process is arrested in each germ cell until the onset of the menstrual cycle. Thereafter, in each cycle, the follicles surrounding particular germ cells ripen, and one is selected for ovulation. Shortly before

ovulation meiosis I resumes and the first polar body is cast off. Because the arrested stage spans many years, it has been supposed that this is the most likely time for non-disjunction to arise in the oocyte. On the other hand, it may be that the oocyte is relatively invulnerable to the environment during the arrested stage, and that it is more vulnerable during the more active stages, at the completion of meiosis I or during the second meiotic division (Figure 1). In such a case, the most likely time for an exposure to produce effects would be around the time of conception. Thus in female mice, ethanol administered shortly before or after hCG-induced ovulation was associated with an increased risk of aneuploidy in fertilized embryos (Kaufman and Bain 1984). This observation supports the view that the meiotic divisions around the time of fertilization may be particularly sensitive to external exposures.

The fourth process—factors affecting the probability of conception—has been examined often in animals, less so in humans. Figures 2 and 3 summarize the origin and biologic staging of the female germ cell (Peters and McNatty, 1980). Some experimental data pertain to cell death of the small oocytes. The effects of cell death, which would deplete the pool of small follicles from which growing and ovulatory follicles develop, would be infertility. Unhappily for the human biologist, experimental work on the sensitivity of follicles shows great variation not only between species, but also between strains, by age of the animal, and by stage of germ cell development (Peters 1969, Baker 1971a; Dobson and Felton 1983). For example, in some strains of mice the small follicles are particularly sensitive to irradiation during certain times in early postnatal development. Exposure at this time results in a higher proportion of deaths of small follicles than of growing follicles. In some strains it also appears to accelerate the development of growing follicles, and at least one study suggests a temporary rise in litter size at the onset of

fecundity (Lindop, Rotblat and Vatistas, 1966). In several strains of mice there is a reduction either in litter size late in reproductive life, in the total number of litters produced, and/or in the fertility span (Peters 1969). In contrast, the guinea pig is less radiosensitive than the mouse, and in this species the growing follicles exhibit greater sensitivity than the small follicles. In humans and some monkeys too, the growing follicles are more sensitive than the small follicles. However, even among those primates that have been studied, the squirrel, rhesus, and bonnet monkeys vary in level of sensitivity (Dobson, Kwan, Brunkhoust et al. 1983). Given these differences between strains and species, it is impossible to extrapolate with any confidence to humans.

From scattered case reports on exposures in humans, it appears that the effects of high dose x-irradiation exposure on fertility depend on the age of the woman treated. Exposures to young women may produce temporary amenorrhea and temporary infertility, while in older women permanent infertility — perhaps better referred to as premature menopause — may result (Baker, 1971b). These observations are compatible with the conjecture that both the small and the growing follicles are sensitive to irradiation, perhaps no more sensitive than any other cells, and the premature menopause of older women results from the death of growing follicles in an ovary where there are few extant small follicles.

Meyer and Tonascia (1981) followed the reproductive patterns of young women (aged 23-29 years) in Baltimore whose mothers had been exposed to pelvic irradiation shortly before or at the time of their birth. The data are summarized in Table 1. The irradiated women were younger at first pregnancy; they had more often given birth to three or more children by the mid-twenties; they were more often sterilized; and had experienced menarche at an earlier

Kline

-4-

age. The data suggest precocious fertility among the exposed, a finding that Meyer and Tonascia suggest bears some resemblance to experimental studies in mice where an acceleration in the development of growing follicles appears to give rise to a temporary increase in litter size. In spite of this resemblance, the likely biology in humans is obscure. In mice many oocytes may be fertilized in one cycle; in humans usually only one oocyte is ovulated, and Meyer and Tonascia do not comment on the incidence of twinning. Another possibility is that in humans, an acceleration in the development of growing follicles is associated with a decrease in the frequency of anovulatory cycles during the teenage years. Since all of the growing follicles present at birth should have undergone atresia long before puberty, this explanation rests on the assumption that changes precipitated by an acceleration in the development of growing follicles during fetal life lead to changes in follicular development (perhaps mediated through hormonal changes) at puberty. Another possibility is that the irradiation affected the pituitary, stimulating precocious development

One last observation speaks to the possible effects of in utero exposure on fertility and fecundability in adult life. Among women exposed in utero to irradiation from the bombings in Hiroshima and Nagasaki and followed to age 27-28 years, the rates of marriage, proportions of couples childless, and the rates of births per years of marriage are similar to a comparison group not living within 1500 meters from the hypocenters (Blot, Shmiz, Kato et al 1975). Table 2 summarizes some of these data. The disparity between this observation and that of Meyer and Tonascia—if not due to chance—could indicate that the effects of in utero irradiation exposure depend either on dose (undoubtedly higher in Japan), the time of exposure (concentrated in the last few weeks of gestation in the study by Meyer and Tonascia), the type of irradiation

Kline

(including neutron irradiation and gamma rays in Japan), or various permutations of these three components. Although there is no suggestion of precocious fertility in those Japanese women exposed to low doses in the third trimester and who married before 25 years, the numbers of women are considerably smaller than the cohort studied by Meyer and Tonascia and there is room for large undetected effects.

The data from Japan also speak to the possible effects of exposure during childhood and adult life on infertility. Just as for in utero exposure, no association was observed with these later exposures, although again the numbers are small (Blot and Sawada, 1972).

Thus the finding from the Baltimore cohort stands as the only set of observations on humans to suggest an influence of in utero irradiation exposure on fertility. No study has so far considered the outcome that would seem the most likely on the basis of animal experiments—premature onset of menopause.

The possible effects of irradiation exposure during postnatal life—usually in adulthood—on aneuploidy in the conceptus have been examined on at least twelve occasions (Kline and Stein, 1985). The majority of studies focus on the relationships between diagnostic and therapeutic irradiation to trisomy 21 (Down's Syndrome) among births; one examined the effects of exposure to the bombings in Japan; two others searched for associations with aneuploidy among spontaneous abortions. The hypothesis underlying these studies is that irradiation exposure to the oocyte during the resting stage (before ovulation) might increase the likelihood of non-disjunction at the completion of maternal meiosis 1, which occurs at ovulation. The observations, taken together, are equivocal. The data from Japan speak to the possible effects of an acute and high dose exposure: the rates of trisomy at birth were similar in exposed and unexposed (relative risk = .4) (Schull and Neel, 1962). Studies dealing with

Kline

-6-

exposure through medical procedures speak primarily to the possible effects of several low dose exposures, or to cumulative exposure. There are both positive reports (Uchida and Curtis 1961; Sigler, Lilienfeld, Cohen et al 1965; Uchida, Holunga and Lowler 1968; Cohen and Lilienfeld 1970; Alberman, Polani, Fraser Roberts et al 1972a; Alberman, Polani, Fraser Roberts et al 1972b) and reports where no association was observed (Lunn 1969; Carter, Evans and Stewart 1961; Marmol, Scriggins and Vollman 1969; Stevenson, Mason and Edwards 1970; Boue, Boue and Lazar 1975; Cohen, Lilienfeld, Kramer et al 1977). These studies vary in their potential for bias, in the methods of ascertaining exposure, and in the definitions of exposure. One can only conclude, very conservatively indeed, that if the oocyte is sensitive to irradiation of the mother during her postnatal life, sensitivity is low or effects are not produced by several low dose exposures.

We now turn to the possible preconception effect of smoking. An association with fertility among women exposed to cigarette smoking in utero is a remote possibility, since several polycyclic aromatic hydrocarbons are associated with death of small follicles in mice transplacentally exposed (Dobson and Felton 1983). We have located one study that examined the effect of exposure to cigarette smoking to the daughter in utero on later reproduction; in it, among pregnant women, in utero exposure was not related to the length of time it took to conceive the pregnancy (Baird and Wilcox 1986). The evidence relating smoking during teenage and adult life to infertility and, among women who do conceive, to the time it takes to become pregnant, is sparse, derived in some instances from selected samples, and inconsistent. Five studies support an association. In one (Tokuhata 1968), among breast cancer cases and controls dying of non-cancerous conditions, women who had ever smoked were less likely to have been pregnant at least once than women who had

Kline

-7-

never smoked. In a Danish series smokers were more likely to seek care at a clinic for infertility, and among those who gave birth, a greater proportion of smokers than non-smokers took more than one year to conceive (Olsen, Rachootin, Schiodt et al 1983). This last observation is consistent with three latter series (Baird and Wilcox 1985; Howe, Westoff, Vessey et al 1985; Kline, Levin and Stein 1985) all of which suggest that among fertile women the length of time taken to conceive a pregnancy is longer among smokers than among non-smokers. For example in a New York City case-control study of spontaneous abortions, among current smokers the instantaneous relative risk of conceiving is .75 that of women who have never smoked, but for ex-smokers the time taken to conceive is significantly briefer than for never smokers (Kline, Levin and Kline 1985). These observations on current smokers contrast with a study of women delivering at the Boston Hospital for Women (Linn, Schoenbaum, Monson et al 1982) where the odds of conceiving within three months was identical in current smokers and non-smokers, and with a study of women delivering in three West-Jerusalem hospitals, where current smoking is reported to be unassociated with time to conception (Harlap and Baras 1984). Given the inconsistency of these observations, and also the perplexing difference in the New York study between ex-smokers and never smokers, it is perhaps premature to speculate on a mechanism. Nonetheless, three are offered: cigarette smoking could favor the death of growing follicles, the failure of implantation, or the loss of a conceptus between implantation and the few weeks following. In support of the first mechanism is the observation that current smokers, but not ex-smokers, experience an earlier age at menopause (Baron, 1984). In support of mechanisms relating to very early reproductive loss is the association of smoking with recognized spontaneous abortion (Kline, Levin, Stein et al 1981; Kline, Levin, Shrout et al 1983; Kline and Stein 1984).

Kline

-8-

Several observations relating smoking to trisomy speak to the possible preconception effect of maternal smoking, since trisomy arises either before or at the time of fertilization. Among births sampled in the Collaborative Perinatal Project (1955-64) and in another set assembled in Upstate New York (1980-81) smoking during pregnancy appears to be inversely associated with Down's syndrome at all maternal ages (Kline, Stein, Susser et al 1980; Hook and Cross 1985). For trisomic spontaneous abortions, the New York City data suggest that among women under thirty years a similar inverse association occurs, whereas among older women ex-smoking and current smoking are positively associated with trisomic abortion (Kline, Levin, Shrout et al 1983). The similar effects of ex- and current smoking in the spontaneous abortion data suggest a common preconception effect. The possible adverse effects of both in relation to trisomic abortions among older women may indicate that the oocytes of smokers and non-smokers have different patterns of aging. Perhaps those of smokers are slower to begin aging, but age faster towards the end of reproductive life.

Exposure after conception

The following are processes which are likely to operate during gestation.

- 1) Teratogenesis, in the narrow sense, meaning a disruption that results in malformation.
- 2) Fetal death.
- 3) Processes resulting in cell death or decreased cell growth.
- 4) Premature expulsion of the conceptus.
- 5) Processes, as yet not described, leading to childhood cancer.

Kline

-9-

Three findings relate to the possible effects of irradiation exposure during pregnancy on the outcome of pregnancy or later morbidity. First, in utero exposure to high doses of irradiation before eighteen weeks gestational age is associated with microcephaly, often accompanied by severe mental retardation, and with deficits in growth (Miller and Blot 1972). Exposure in the infant and growing child may also be associated with deficits in skeletal growth (Sutow, Conard and Griffith 1965). These findings derive primarily from Japan, although some clinical case reports are consistent with these findings (for example, Goldstein and Murphy 1929; Dekaban 1968). Second, again in Japan, irradiation exposure sufficient to cause radiation sickness in the mother appears to be associated with an increased risk of fetal death, and of neonatal and infant death (Yamazaki, Wright and Wright 1954). There are few other observations relating gestational exposure to fetal death (see Schull and Bailey 1984). Of course, radiation sickness is associated with higher levels of exposure, and thus it is unclear whether radiation acts directly on the fetus or through an association of fever or some other symptom with both radiation sickness and fetal death. Third, some studies report an association of prenatal exposure to diagnostic x-rays to the pelvic area with leukemia and other malignancies during childhood (see reviews by MacMahon 1962; Lilienfeld 1966; and Brent 1979). Taken together, these studies suggest that if there is an association of prenatal exposure to childhood mortality from leukemia, the relative risk is about 1.5. Observations for the atomic bomb cohort are not in accord with predictions that the risk of childhood leukemia may increase linearly with the dose of the prenatal exposure. In this cohort, one childhood leukemia death was observed compared to .75 expected (Jablon and Kato 1970).

For smoking, the association exposure during pregnancy with intrauterine growth retardation in the infant is well-documented. The birthweights of

Kline

-10-

infants of smokers are on average 100-300 grams lower than those of non-smokers. For this outcome, ex-smokers are similar to non-smokers. Although the nicotine content of cigarettes has changed markedly over time, the magnitude of the effect of smoking on birthweight has remained constant over the past twenty years (Kline, Levin, Stein et al 1981). These observations, coupled with the fact that intrauterine growth retardation associated with smoking resembles that associated with high altitudes in that the placenta remains relatively large in relation to birthweight, suggest a role for hypoxia in this well known association.

Lastly, maternal smoking during pregnancy has been associated with Sudden Infant Death, although an effect of postnatal exposure cannot be ruled out (Hoffman 1985).

Conclusion

Taking together the evidence from both experimental and epidemiologic studies, even in the face of the many gaps in knowledge, it seems likely that there is considerable specificity between the circumstances of exposure and the consequences. Table 3 summarizes some of the epidemiologic evidence relating irradiation exposure to reproductive outcome. The mechanisms listed are speculative and come primarily from animal experiments. Nonetheless, they serve to illustrate that depending on the time of exposure, the dose, and the outcome, different biologic phenomena are more or less plausible.

For cigarette smoking, we have attempted to list the many consequences that might be examined given various circumstances of exposure (Table 4). The possible effects of in utero exposure to the female on sexual behavior, fertility, and pregnancy outcomes, particularly aneuploidy, have not been examined in humans, although experimental evidence permits the hypothesis that

Kline

-11-

the germ cells may be affected by exposure to various constituents in cigarette smoke. The bulk of the evidence relates smoking during pregnancy to various outcomes of the pregnancy. The possible effects of having smoked prior to, but not during, pregnancy are less well studied although it might be argued that for many of the outcomes listed, such as prematurity, fetal death, and complications of pregnancy, an association would be unlikely. For outcomes where processes involving the follicles or chromosomes seem likely, it might be expected that current and ex-smoking would produce similar effects. In contrast, if the likely processes involve estrogen or other hormonal changes, hypoxia, or the immediate effects of nicotine, then ex-smokers would be expected to behave like women who have never smoked, rather than like current smokers. There are many outcomes -- such as anovular cycles, early reproductive loss, infertility -- where the biologic processes are either varied or poorly understood. In these instances, it may be that epidemiologic studies that shed light on the similarities and differences between current and ex-smokers will provide the leads as to mechanism.

Kline

-12-

TABLE 1: Adjusted rates (per cent) of reproductive outcomes for women exposed, while in utero, to x-irradiation and matched unexposed women

	Exposed (N=839)	Unexposed (N=839)
Statistically significant differences		
First pregnancy age ≤ 15 years	8.0	4.5
Left school because of pregnancy	19.9	15.2
3+ living children	17.7	12.7
Stillbirths (> 20 weeks)	3.5	1.3
Sterilized	17.7	12.7
Non-significant differences		
Menarche ≤ 10 years	11.0	8.4
Spontaneous abortions	11.2	10.3
Therapeutic abortions	15.6	14.2

Source: Meyer and Tonascia (1981)

TABLE 2: Rates of marriage and rates of childless marriages among women who were in utero at the time of the atomic bombings in Japan: follow-up at ages 27-28 years

	<u>All</u> <u>Exposed</u>	<u>100 + Rad</u>	<u>Unexposed</u>
Percent married	77.2	65.2	77.8
Percent childless			
Age < 25 at marriage	12.6	10.3	10.0
Age ≥ 25 at marriage	49.4	43.8	44.8

Source: Blot, Shimiz, Kato et al (1975)

TABLE 3: Possible relations between irradiation exposure at different times and reproductive outcomes

Time of Exposure	Effects	Possible Mechanisms
IN UTERO		
High dose, radiation sickness	Fetal death ^a	Death of embryo, toxicity to mother
Single high dose before 18 weeks gestation	Mental retardation, microcephaly, growth retardation ^b	Cell death
Third trimester low dose	Possible increase in fertility and in adolescence and early adulthood ^c	Accelerated development of growing follicles leading to an increased frequency of ovulatory cycles in teenagers
Diagnostic x-rays	Possible relation to childhood leukemia and other malignancies ^d	Mutagenesis
PRECONCEPTIONAL TO FEMALE		
Diagnostic therapeutic x-ray to abdominal area	Inconsistent evidence relating to aneuploidy in births and abortions ^e	Damage to oocyte
Therapeutic high dose during childhood or adulthood	Temporary infertility in young women; possible premature infertility in older women ^f	Death of growing follicles

^a Yamazaki, Wright and Wright 1954

^b Miller and Blot 1972

^c Meyer and Tonascia 1981

^d Reviewed by MacMahon 1962; Lilienfeld 1966; Brent 1979

^e Reviewed by Kline and Stein 1985

^f See Baker 1971b

TABLE 4: Possible relations of current and ex-smoking to various reproductive outcomes

Outcomes	Current Smoking	Ex-smoking
Sexual behavior	?	?
Teenage pregnancy	?	?
Response to oral contraception	?	?
Estrogen metabolism	?	?
Anovular cycles	?	?
Menstrual irregularity	?	?
Time to conceive	possibly lengthened	possibly shortened
Very early loss	?	?
Euploid spontaneous abortion	possibly increased	possibly no association
Trisomic spontaneous abortion		
< 30 years	possibly decreased	possibly decreased
≥ 30 years	possibly increased	possibly increased
Stillbirth	increased	?
Intrauterine growth retardation	increased	no association
Gestation at birth	possibly shortened	?
Antepartum hemorrhage	increased	?
Pre-eclamptic toxemia	decreased	?
Trisomy 21 at birth	possibly decreased	possibly decreased
Sudden Infant Death	increased	?

? = insufficient evidence

REFERENCES

- Alberman, E., Polani, P.E., Fraser Roberts, J.A., et al (1972a) Parental exposure to X-irradiation and Down's Syndrome. *Ann. Hum. Genet. Lond.* 36:195-208.
- Alberman, E., Polani, P.E., Fraser Roberts, J.A., et al (1972b) Parental X-irradiation and chromosome constitution in their spontaneously aborted fetuses. *Ann. Hum. Genet. Lond.* 36:185-94.
- Baird, D.D. and Wilcox, A.J. (1985) Cigarette smoking associated with delayed conception. *JAMA* 253:2979-83.
- Baird, D.D. and Wilcox, A.J. (1986) Future fertility after prenatal exposure to cigarette smoke. *Fertil. Steril.* (in press).
- Baker, T.G. (1971a) Comparative aspects of the effects of radiation during oogenesis. *Mutat. Res.* 11:9-22.
- Baker, T.G. (1971b) Radiosensitivity of mammalian oocytes with particular reference to the human female. *Am. J. Obstet. Gynecol.* 110:746-61.
- Baron, J.A. (1984) Smoking and estrogen-related disease. *Amer. J. Epidemiol.* 119:9-22.
- Blot, W.J. and Sawada, H. (1972) Fertility among female survivors of the atomic bombs of Hiroshima and Nagasaki. *Am. J. Hum. Genet.* 24:613-22.
- Blot, W.J., Shimiz, Y., Kato, H. and Miller, R.W. (1975) Frequency of marriage and live birth among survivors prenatally exposed to the atomic bomb. *Am. J. Epidemiol.* 120:128-36.
- Boue, J., Boue, A. and Lazar, P. (1975) Retrospective and epidemiological studies of 1500 karyotyped spontaneous human abortions. *Teratology* 12:11-26.
- Brent, R. (1979) Effects of ionizing radiation on growth and development. *Contr. Epid. Biostatist.* 1:147-83.

Kline

-17-

- Carter, C.O., Evans, K.A. and Stewart, A.M. (1961) Maternal radiation and Down's Syndrome (mongolism). *Lancet* 2:1042.
- Cohen, B.H. and Lilienfeld, A.M. (1970) The epidemiological study of mongolism in Baltimore. *Ann. N.Y. Acad. Sci.* 171:320-27.
- Cohen, B.H., Lilienfeld, A.M., Kramer, S. and Hyman, L.C. (1977) Parental Factors in Down's Syndrome — Results of the Second Baltimore Case-Control Study. In, Population Cytogenetics. (eds. E.D. Hook and I.H. Porter) New York, Academic Press, pp 301-52.
- Dekaban, A.S. (1968) Abnormalities in children exposed to X-radiation during various stages of gestation: Tentative timetable of radiation injury to the human fetus, Part 1. *J. Nucl. Med.* 9:471-77.
- Dobson, R.L. and Felton, J.S. (1983) Female germ cell loss from radiation and chemical exposure. *Am. J. Ind. Med.* 4:175-90.
- Dobson, R.L., Kwan, T.C., Brunkhous, B.F. and Straume, T. (1983) Diminished Lifetime Reproductive Capacity in the Female Following Early Radiation Exposure. Presented at the Twenty-Second Hanford Life Sciences Symposium, September 23, 1983.
- Goldstein, L. and Murphy, D.P. (1929) Etiology of the ill-health in children born after maternal pelvic irradiation. *Am. J. Roentgen.* 22:322-31.
- Harlap, S. and Baras, M. (1984) Conception-waits in fertile women after stopping oral contraceptives. *Int. J. Fertil.* 29(2):73-80.
- Huffman, H. (1985) Personal communication.
- Hook, E.B and Cross, P.K. (1985) Cigarette smoking and Down's Syndrome. *Am. J. Hum. Genet.* 37:1216-24.
- Hower, G., Westoff, C., Vessey, M. and Yeatts, D: (1985) Effects of age, cigarette smoking, and other factors on fertility: findings in a large prospective study. *Br. Med. J. (Clin. Res.)* 290 (6483):1697-700.

Kline

-18-

- Jablon, S. and Kato, H. (1970) Childhood cancer in relation to prenatal exposure to atomic bomb radiation. *Lancet* 2:1000-03.
- Kaufman, M.H. and Bain, I.M. (1984) Influence of ethanol on chromosome segregation during the first and second meiotic divisions in the mouse egg. *J. Exp. Zool.* 230:315-20.
- Kline, J., Levin, B., Shrout, P., Stein, Z., et al. (1983) Maternal smoking and trisomy among spontaneously aborted conceptions. *Am. J. Hum. Genet.* 35:421-31.
- Kline, J., Levin, B. and Stein, Z. (1985) Time to conceive in fecund women. In Preparation.
- Kline, J., Levin, B., Stein, Z. and Warburton, D. (1981) Epidemiologic detection of low dose effects on the developing fetus. *Environ. Health Perspect.* 42:119-26.
- Kline, J. and Stein Z. (1985) Environmental Causes of Aneuploidy: Why So Elusive? In, Etiology of Aneuploidy. (eds. V. Vaughan-Dellarco, P. Voytek and A. Hollaender) New York, Plenum Press, pp 149-64.
- Kline, J. and Stein, Z. (1984) The Epidemiology of Spontaneous Abortion (miscarriage). In, Perinatal Epidemiology. (ed. M. Bracken) London, Oxford University Press, pp 23-51
- Kline, J., Stein, Z., Susser, M. and Warburton, D. (1980) New Insights into the Epidemiology of Chromosomal Disorders: Their Relevance to the Prevention of Down's Syndrome. In, Frontiers of Knowledge in Mental Retardation. (ed. P. Mittler) Baltimore, University Park Press, pp 131-41.
- Lilienfeld, A.M. (1966) Epidemiological studies of the leukemogenic effects of radiation. *Yale J. Biol. Med.* 39:143-64.

- Lindop, P.J., Rotblat, J. and Vatistas, S. (1966) The Effect of Age and Hypoxia on the Long Term Response of the Ovary to Radiation. In, Radiation and Aging. (eds. P.J. Lindrop and G.A. Sacher) London, Frances and Taylor, Ltd., pp 307-324.
- Linn, S., Schoenbaum, S.C., Monson, R.R., et al (1982) Delay in conception for former "pill" users. JAMA 247:629-32.
- Linn, J.E. (1959) A survey of mongol children in Glasgow. Scott. Med. J. 4:368-72.
- MacMahon, B. (1962) Prenatal X-ray exposure and childhood cancer. J. Nat'l. Cancer Inst. 28:1173-91.
- Marmol, J.G., Scriggins, A.L. and Vollman, R.F. (1969) Mothers of mongoloid infants in the collaborative project. Am. J. Obstet. Gynecol. 104:533-43.
- Meyer, M.B. and Tonascia, J. (1981) Long-term effects of prenatal X-ray of human females. Am. J. Epidemiol. 114:304-16.
- Miller, R.W. and Blot, W.J. (1972) Small head size following in utero exposure to atomic radiation, Hiroshima and Nagasaki. Atomic Bomb Casualty Commission. Tech. Rep. 35-72:1-7.
- Olsen, J., Rachootin, P., Schiodt, A.V. and Damsco, N. (1983) Tobacco use, alcohol consumption and infertility. Int. J. Epidemiol. 12:79-84.
- Peters, H. (1969) The Effect of Radiation in Early Life on the Morphology and Reproductive Function of the Mouse Ovary. In, Advances in Reproductive Physiology. (ed. A. McLaren) London, Logos Press Ltd., 149-86.
- Peters, H. and McNatty, K.P. (1980) The Ovary. Berkeley, University of California. (In Press)
- Rogan, W.J. (1982) PCB's and cola-colored babies: Japan 1968 and Taiwan 1979. Teratology 26:259-61.

Kline

-20-

- Schull, W.J. and Bailey, J.K. (1984) Critical Assessment of Genetic Effects of Ionizing Radiation on Pre- and Postnatal Development. In, Issues and Reviews in Teratology. (ed. H. Kalter) New York, Plenum Press, 325-91.
- Schull, W.J. and Neel, J.V. (1962) Maternal radiation and mongolism. *Lancet* I:537-38.
- Sigler, A.T., Lilienfeld, A.M., Cohen, A.M. and Westlake, J.E. (1965) Radiation exposure in parents of children with mongolism (Down's Syndrome). *Bull. Johns Hopkins Hosp.* 117:374-99.
- Stevenson, A.C., Mason, R. and Edwards, K.D. (1970) Maternal diagnostic X-irradiation before conception and the frequency of mongolism in children subsequently born. *Lancet* 2:1335-37.
- Sutow, W.W., Conard, R.A. and Griffith, K.M. (1965) Growth status in children exposed to fallout radiation on Marshall Islands. *Pediatrics* 36:721-31.
- Tokuhashi, G.M. (1968) Smoking in relation to infertility and fetal loss. *Arch. Environ. Health* 17:353-59.
- Uchida, I.A., Holunza, R. and Lainier, C. (1968) Maternal radiation and chromosomal observations. *Lancet* 2:1045-49.
- Uchida, I.A. and Curtis, E.J. (1961) A possible association between maternal radiation and mongolism. *Lancet* 2:848-50.
- Yamazaki, J.N., Wright, S.W. and Wright, P.M. (1954) Outcome of pregnancy in women exposed to the atomic bomb in Nagasaki. *Am. J. Dis. Child* 84:448-62.

Timing of Meiosis in the Female Germ Cell

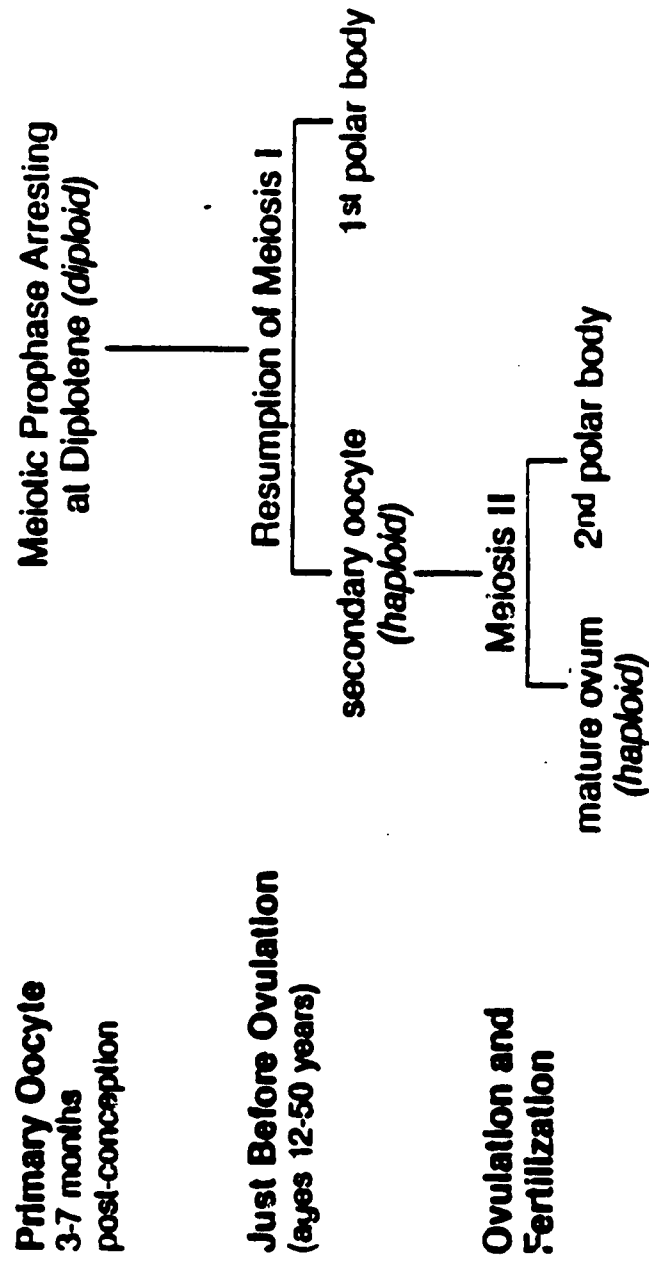


FIG. 1

Source: Peters and McNatty (1980)

OOGENESIS

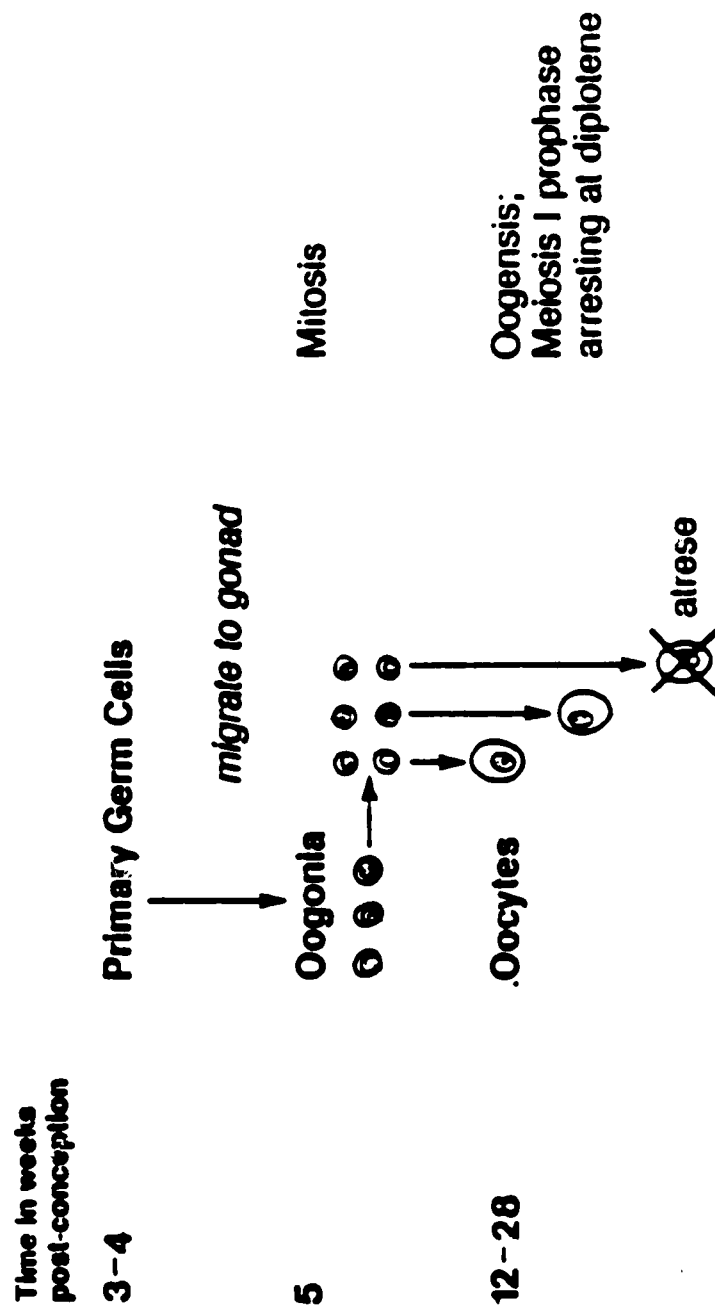


FIG. 2

Adapted from: Peters and McNatty (1980)

OÖGENESIS

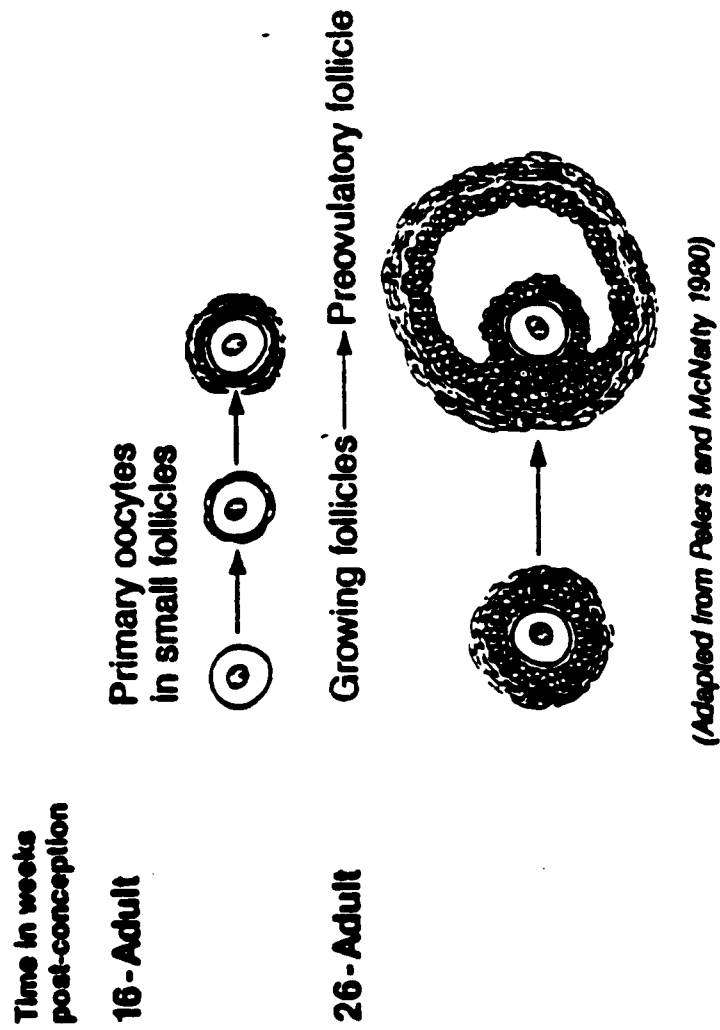


FIG 1.

Chapter 26
Evaluation of Epidemiologic Information
Moyses Szklo

EVALUATION OF EPIDEMIOLOGIC INFORMATION

Moyses Szklo

It is clear from the preceding papers that, in addition to the development of statistical techniques and models that allow us to better quantify and sort out risks associated with specific exposures, features of study design are critically important, in particular those pertaining to the validity and reliability of assessing exposure and outcome variables.

Several presentations have dealt with the need to achieve more detailed and valid measurements of exposure. For example, Comstock pointed out the problem of assessing exposure to a risk factor when the individual exposure has to be inferred from an average population value. Kline emphasized the importance of the "circumstances of exposure", including characteristics such as duration, level, route and time of exposure. The presentation by Land covered yet another problem that related to extrapolating data obtained from high level to low level risk factor exposure.

A number of speakers have also addressed problems related to assessment of outcome. Kline described different ways in which reproductive outcomes vary according to circumstances of exposure. In assessing the presence and magnitude of the leukemia risk induced by benzene exposure, Labarthe has aptly described the additional exposure factors that must be examined in detail, including dose and distinction between environmental and personal exposure data. However, as also pointed out by Labarthe, it is becoming increasingly clear that, to study the association of benzene and leukemia further specification of the outcome is also needed, since it appears that the main types of leukemia affected by exposure to organic solvents are the acute nonlymphocytic and perhaps also the chronic lymphocytic types (Linnet 1985).

Furthermore, there may be a need to be even more specific within the acute nonlymphocytic type, by examining separately different cell sub-types, such as the myelocytic, the monocytic and so on. If the association between benzene and leukemia is specific to a given cell sub-type, it obviously follows that studies which have been carried out using "all leukemias" or even broad leukemia cell types as end-points, may have underestimated the true excess risk associated with that sub-type.

Another concern in assessing outcome is that, as important as histologic classification may be, it is only one of the many criteria needed to improve the specificity of the end-point under study. As pointed out by Hulka, in some situations the ways in which a risk factor relates to the target cells will be a function of the presence of specific cell receptors. Thus, for example, further distinguishing subtypes of the same histologic type of breast cancer on the basis of presence or absence of hormone receptors, may permit more valid and reliable estimates of relative risks associated with specific risk factors, particularly those related to hormones.

I would like now to discuss an area covered in several presentations which is of great importance in evaluating epidemiologic information; interpretation of inconsistencies in results obtained from different epidemiologic studies. These inconsistencies tend to create considerable "intellectual stress", since we are often dealing with both a fairly common and preventable factor, such as estrogen therapy, and a common and serious disease, such as breast cancer, for which no modifiable risk factors have yet been identified. As a result the exciting prospect of primary prevention is raised.

Szklo

-2-

I would like to discuss briefly some of the reasons why results may vary from study to study which have been mentioned by Hulka, and to add a few others. This discussion will serve as a framework for discussing other topics relevant to evaluation epidemiologic data in general. (Table 1)

The importance of improving the specificity of the outcome has already been stressed (Slide 1). At this point it is sufficient to say that inapparent differences in types of outcome from study to study may lead to inconsistent results.

If the circumstances of exposure differ in different studies, it may also lead to differences in results. This would be the case, for example, if the chemical composition of a given drug, such as a menopausal estrogen, varied from study to study. Another example, discussed by Land, is that of differences in circumstances of exposure to ionizing radiation leading to different cellular outcomes, namely, either inactivation or breaking of chemical bonds.

Differences in study design and methods are also important in explaining differences among studies. Design features such as selection of controls and the extent to which different biases operate in different investigations may have an impact in the relative risk estimates. Also important are differences in the degree of comparability between cases and controls or between exposed and non-exposed individuals, leading to more or less biased estimates of the relative odds or relative risks. Studies may vary in the extent to which they identify and take into account confounding variables, which again may lead to differences in relative risk estimates, even though these estimates may be generally reported as "adjusted". The term, "pooling of inferences" — in contrast to "pooling of results" — has been suggested to define a special

Szklo

effort to develop a list of common confounding factors to be adjusted for and to standardize adjustment procedures in different studies, in order to assess homogeneity of results. Since the mere examination of published reports is often not sufficient to carry out the "pooling of inferences" just described, a detailed examination of methodologic features of each study and additional data analyses may be necessary.

Another factor that could explain differences among studies is of course chance resulting from sampling variability, but I would rather focus now on a less apparent reason for variability among studies, mentioned by some of the speakers, namely, the possible interaction between the factor under study and one or more characteristics of the population, the presence and magnitude of which may vary greatly from study to study and may produce effect modification. The notion that different populations or population sub-sets may differ in their susceptibility to risk factors has been extensively tested in animals, such as in Dahl's classical studies showing that salt-induced hypertension risk in rats varies as a function of genetic susceptibility or resistance (Dahl et al 1960). However, only relatively recently has interaction become one of the concerns of epidemiologists.

An example of how interaction may explain variability in study results is given in Table 2, which shows results of a study by Brinton et al. (1981), mentioned by Hulka. This study dealt with the association between estrogen therapy and breast cancer. Only bilaterally oophorectomized women were included in the analysis, because of the suggestion, pointed out by Hulka, that an estrogen-induced increase in breast cancer risk would be particularly biologically meaningful in these women. Although the results were not statistically significant, and in spite of the fact that no interaction

Szklo

-4-

hypotheses were specified by the authors in advance of the study, when relative odds for estrogen users vs. non-users were calculated separately for women with and without a maternal history of breast cancer, a large difference was seen in the magnitude of the relative odds, thus implying multiplicative interaction. Although these findings have to be confirmed in view of their lack of statistical significance, they nevertheless suggest that variation in relative odds seen in different studies may be a function of the prevalence of effect modifiers in different study populations.

Another example dealing with artificial sweeteners and bladder cancer was covered in Cole's presentation. As seen in Table 3, in the study by Hoover and Stasser (1980), for all males no association was seen between use of artificial sweeteners and bladder cancer. However, for heavy smokers, a strong and statistically significant relationship was found, with a suggestion of a dose-response pattern. Again, these results could have been due to chance, but if populations studied thus far included small proportions of heavy smokers, the relative odds in most studies would be close to 1.0, as indeed was shown by Cole.

An additional example of interaction comes from the study mentioned by Kline, showing that age seems to modify the way in which smoking relates to trisomic spontaneous abortions, in that the association is negative for younger and positive for older women. The author postulated that this interaction could be explained if oocytes of smokers and non-smokers had different patterns of aging: Those of smokers would be slower to begin aging, but would age faster towards the end of reproductive life. Thus, results of studies on smoking and trisomic abortions may be affected by the age composition of the study population.

Szklo

-5-

Interaction should be an integral part of the study hypotheses, and not merely a statistical nuisance to be considered when using multiple regression models. In addition, hypotheses involving effect modification or interaction should, whenever possible, be formulated when designing the study, so that "multiple tests" problems could be avoided.

Finally, discrepancies in results obtained from different studies may also relate to when the study was carried out. Particularly in the study of newly introduced agents, such as drugs, which have a long latent period, studies carried out at a time when minimum latency has not yet gone by, will tend to underestimate the relative risk or the relative odds. In addition, public perception of a possible increase in risk associated with a popular exposure, such as for example, coffee drinking in relationship to pancreatic cancer, may be influenced by the reports from initial study and may thus lead to reporting bias in subsequent studies.

Let me close this discussion by saying that it is quite apparent from the presentations that epidemiology is at a crossroad. Although more and more time is being devoted to refining statistical techniques for dealing with problems such as precision of relative risk estimates, it is clear that the most difficult challenges that epidemiologists have to face are related, not to the statistical aspects of risk quantification, but to the validity and reliability of the measurements of both exposure and outcome variables. As pointed out by Comstock, in the initial phases of hypothesis-testing, even ecologic studies may be appropriate and, in fact, under certain circumstances may even allow us to examine a hypothesis better (Table 4): Ecologic studies are certainly less costly, and can be particularly useful when the exposure gradient within the population is low. In addition, they allow us to estimate relative risks by

Szklo

means of regression. However, they also have the greatest potential for confounding and bias and therefore must be followed by studies that measure both exposure and outcome using individuals, not populations, as units.

However, as was extensively and carefully demonstrated by our speakers, it would be a mistake to believe that our job can be carried out by merely doing well-designed case-control, cohort or even experimental studies in which exposure and outcome variables are represented as dichotomous categories that fit nicely in a 2x2 table. Although this is still an excellent way to begin, we are slowly but surely learning that in order to improve our estimates of risk, much more detail is needed to describe and measure the ways in which a risk factor relates to the disease under study.

REFERENCES

1. Brinton L, Hoover RN, Szklo M and Fraumeni JF: Menopausal estrogen use and risk of breast cancer. *Cancer* 47:2517-2522, 1981
2. Dahl IK, Heine M and Tassinari L: High salt content of western infants' diet: Possible relationships to hypertension in the adult. *Nature* 198:1204-1205, 1963
3. Hoover RN and Stasser PH: Artificial sweeteners and human bladder cancer: Preliminary results. *Lancet* 1:837-840, 1980
4. Linet M: *The Leukemias: Epidemiologic aspects*. Oxford University Press, New York, 1985 (pp 153-157)

TABLE 1: Possible reasons for inconsistent results from different epidemiologic studies

1. Definition/type of outcome (Ex: ER+ vs. ER-BC)
2. Circumstances of exposure
3. Methods; types of bias
4. Confounding effect; degree of adjustment
5. Sampling variability
6. Effect modification
7. Timing of study

TABLE 2: Hormone-associated breast cancer relative odds among women with bilateral oophorectomy (158 cases and 177 controls)

<u>Family History of Breast Cancer in Mother</u>	<u>Relative Odds (Ever Use of Estrogens vs. No Use) 95% C.I.)</u>
No	1.28 (0.8 - 2.2)
Yes	6.67 (0.6 - 71.0)

Source: Brinton et al., Cancer 47:2517, 1981

TABLE 3: Diet drink use and bladder cancer: Synergism with smoking and race*

	Odds Ratio	
	All Males**	White Males Who Smoked More Than 40 Cigarettes/Day***
Never used artificial sweeteners	1.0	1.0
<2 servings diet drink	0.93	1.20
2-2.9 servings	1.44	3.30
>3 servings	1.01	2.62
	NS	(p = .01)

* Hoover and Stasser, 1980

** Adjusted for age, race and smoking

***Adjusted for age

TABLE 4: Comparison of characteristics of ecologic studies and case-control or cohort studies

<u>Characteristic</u>	<u>Ecologic</u>	<u>Case-Control/Cohort</u>
Unit of study	Population	Individual
Cost	Lower	Higher
Useful when:	Exposure gradient within community is low	Exposure gradient within community is high
Potential for confounding	Higher (difficult to adjust for individual risk factors)	Lower
Potential for bias in measurement of exposure	Higher (because exposure measured at community level may not represent exposure to individuals)	Lower
Measurement of strength of association	Use regression	Straightforward

Part VI
Conclusion

Chapter 27

Risk Assessment and Public Policy

Edward N. Brandt, Jr.

RISK ASSESSMENT AND PUBLIC POLICY

Edward N. Brandt, Jr.

The assessment and management of risk in our society are two of the more difficult problems that the scientific community must work to resolve. Public policy decision-making is influenced by a large number of players, often with conflicting and or contradictory data. Furthermore, the assessment of risk often becomes a very emotional issue and the public policy decision-making process tends to bring out both the best and the worst of human behavior.

Indeed, it was in the realm of risk assessment and risk management that this author was subjected to the most personal attacks and, I can assure you that it is easy to become "gun shy".

The following are several statements from the original proposal forming the basis for this symposium:

"Management does not know how to interpret negative findings, or to reconcile epidemiological studies which yield contradictory results. In addition, there is concern that preliminary or tentative epidemiological findings may be overly interpreted and lead to inappropriate anxiety or regulatory pressure."

Both of those statements are certainly true. Furthermore, the concept statement also emphasizes the lack of familiarity of many decision-makers with epidemiologic methods. One need only consider the continuing congressional hearings on the scientific aspects of cigarette smoking to understand that. In those hearings the phrase, "but that's based only on epidemiological studies" is heard repeatedly. This symposium should help in clarifying the contributions of epidemiology to risk assessment.

The following are personal views on the fundamental issues in public policy related to risk assessment and risk management. I should like to address three questions:

1. On what basis do you decide to do definitive studies to assess the risk of a specific substance?
2. How much scientific evidence is required to institute risk management strategies?
3. Which risk management strategies should be employed?

Turning to the first of these, deciding when to do definitive studies, three aspects need to be considered. The first is legal. For example, the Food, Drug and Cosmetic Act requires that there be reasonable assurance of no significant health hazard before substances can be added to food, and drugs must be demonstrated to be safe before they are permitted on the market. Furthermore, the so-called Delaney Clause says that any risk of cancer in man or animal requires that the substance not be used. Hence, in the case of food, drugs and related substances regulated by the Food and Drug Administration, it is clear that the substances must be considered safe before they can be used. That clearly says that risk assessment studies must be carried out before the substance is used. But even that does not settle the issue. What is safe? What about individuals who show hypersensitivity to a particular substance? How much risk can be tolerated? Can you assess risk at low doses?

The second consideration has to be the science. When anecdotal reports began to appear about adverse health effects from substances, how much and what kind of evidence is required to decide whether or not to engage in more definitive studies? This is a common issue in drugs and in occupational exposures in which there are some established systems for reporting adverse

Brandt

-2-

effects. In my experience, these were the most difficult decisions that had to be faced. As general guidelines, principles followed were that the more severe the adverse effect and/or the greater frequency with which it was reported, the greater the need for definitive studies.

The third aspect is public opinion. Public opinion in our society is a very potent force, as it should be. It is also easy to mobilize public opinion when dealing with the assessment of risk. It seems relatively clear that public opinion as expressed by elected officials and by the media is the most powerful voice in determining when to do definitive studies irrespective of the scientific facts. The issues of multiple risk factors and multiple exposures are not easy to explain to the public, especially when the science underlying them may not be clear. One cynical way to describe this phenomenon is to say that it is difficult to explain something you don't understand. One cannot ignore a second aspect of this decision, which is to decide what studies will be carried out. It is in this realm that I believe this symposium will yield great benefit.

In regard to the second issue, namely, when do you decide to institute risk management procedures, again, one has all three aspects to consider: legal, scientific and public opinion. In the case of the Delaney Clause for food additives or cosmetic additives, this is a non-question because the law says they must be banned. It should be added, however, that some argue that a de minimus interpretation is legally and scientifically defensible. Most situations are not very clear. Before going further, the views of the risk management options that are open to decisions will be presented.

The first of these is no action. Clearly, when there is no risk there should be no management. The second risk management action is education. Very simply education should be a component of all risk management strategies. It

Brandt

-3-

seems quite evident that only if everyone understands why and what you are doing can you then gain the necessary compliance. The best example of that is drunk driving. It has been against the law in this country in most jurisdictions for some forty years to drive while intoxicated. Yet, until just a couple of years ago there was a continuing steady increase in mortality and morbidity from drunk driving. Institution of educational programs by such organizations as the government, the media, Mothers Against Drunk Driving, Students Against Drunk Driving, and similar groups have resulted in a decline in the incidence of this preventable cause of death.

However, one form of risk management is education only. The classic example of that is, of course, labeling of such things as household chemicals.

The third risk management technique is control; that is, attempting to define safe conditions of use, or ways to protect people. Use of inhalation devices by miners, protective clothing for workers, and limiting drugs to prescriptions are all examples of this form of risk management.

Finally, there is outright ban. This is clearly a very drastic step and should be limited to unusual situations. EDB, DDT and other chemicals, of course, fall into this category.

The dilemma of deciding when to implement risk management strategies can perhaps best be examined by consideration of the Reye Syndrome and aspirin controversy. Reye Syndrome is a rare childhood disease that is frequently fatal. It occurs in children and teenagers following chicken pox and influenza. Four retrospective epidemiological studies showed increased risk for Reye Syndrome associated with the use of salicylate-containing substances during viral infections. These studies had been carried out by different investigators in three different states and were quite consistent in their results. There were, however, difficulties with the studies. First, there was the issue of recall by the parents as to what medication they had given their

Brandt

-4-

child. Second was the issue of determining whether that medication contained salicylate since, as you know, certain of the standard "flu remedies" come in two forms, with and without aspirin. Finally, there was the question of whether Reye Syndrome had been diagnosed properly, especially in its early stage. As these data were examined, it seemed clear that the evidence was strong enough to warrant a full definitive prospective study. At the same time, however, one had a disease that, although rare, had a very high fatality rate. Furthermore, since very few children with chicken pox or the flu require aspirin even for fever reduction, taking no action did not seem appropriate.

During this same period of time, the Committee on Infectious Diseases of the American Academy of Pediatrics released a report saying that aspirin should not be used in cases of flu or chicken pox because of the risk of Reye Syndrome, and an NIH Consensus Development Conference came to the same conclusion. Therefore, it was determined to institute labeling of aspirin to indicate that aspirin should not be given to children with flu and chicken pox without first consulting a physician. Let me emphasize that it was not recommended that parents be told that they should never use aspirin in cases of chicken pox or flu or that an attempt be made to remove children's aspirin from the market.

Well, many know what followed. A group called "The Committee for the Care of the Child" was formed and it began to lobby intensively to block any efforts at risk management. The Aspirin Foundation also entered this battle and produced a number of critiques of the epidemiologic studies. I met with each of these groups on several occasions to hear their arguments against labeling. However, it was felt that although the evidence was incomplete, it was adequate to label, especially given that aspirin wasn't indicated in either of those diseases to begin with. Then the Board of the American Academy of Pediatrics

Brandt

-5-

refused to endorse the Committee on Infectious Disease's recommendations which, by the way, had already been published in the Academy's journal. That action resulted in the resignation of the Chairman of that committee. Therefore, a situation existed in which there was no scientific group that was willing to support conclusions that risk management in the form of labeling was indicated. It was therefore decided to drop back one step and institute an educational campaign without labeling the product.

Those who followed this issue will then remember that I was investigated by a Congressional Committee. Certain national news services maintained that I had "caved" to the aspirin industry. At the same time, the aspirin industry was claiming that I was determined to ruin them. In any event, a strategy of an educational campaign and the carrying out of a definitive prospective study with the Institute of Medicine as the overseers was decided on. At the present time, that study is still underway, but the preliminary data are so strong that the Oversight Committee of the Institute of Medicine backed our original recommendation for the institution of risk management. As is apparent, this decision involves all aspects of the decision-making process. Public opinion was rapidly mobilized by the media, the scientific community was unable to come to consensus on the issue and the Congress jumped into the situation on both sides of the issue. As an aside, it should be pointed out that this was a rather unpleasant experience for the author even though the scientific community now agrees with the original decision. It resulted in some of the most vicious personal attacks that I received during my four years in Washington. One final note, since the institution of the educational campaign the death rate from Reye Syndrome has dropped significantly, as has the number of cases being reported, although reporting has been improved. Cases being

Brandt

-6-

reported now are being seen at an earlier stage in the disease when they can still be treated. I believe this program has been successful.

The final issue, of course, is the selection of which risk management procedure to utilize. Again, there are legal issues such as the Delaney Clause which specifies that the product must be banned, scientific issues such as the strength of the evidence and its consistency, and, finally, public opinion. This decision is a complex interaction of all of these factors as a general rule, but there are certain guidelines that can be used. The first, of course, is that the more severe the adverse health effects are, the stronger the risk management strategy to be employed. A second consideration is whether there are alternatives available for the substance in question. Finally, one must consider whether there are mechanisms for controlling the risk and, thereby, permitting safe use or whether the benefits are worth the risks. For example, cancer chemotherapeutic agents carry significant risks although, admittedly, it isn't clear in every instance how much of the adverse health effects seen are due to the drug and how much to the basic disease process. Nevertheless, many of these agents carry a high enough risk that they are permitted to be used only because the disease they are being used to treat is fatal. Hence, in this situation, the benefit does outweigh the risk. Needless to say, the greater the evidence and the more plausible the potential scientific explanation, the better. My own judgment is that you should implement the minimum risk management procedure that is required to keep the risk as low as possible.

In summary, the following points are offered:

First, public policy decision-making will rarely be based entirely on scientific data. Emotional, economic and political considerations will all play a role. However, that should not diminish our interest and efforts to

Brandt

-7-

improve the science of risk estimation and control. Good methodology for dealing with competing risks will lead to a greater role for science.

Second, legislative intervention alone is usually not a solution and may be harmful. An example is the Agent Orange bill which permits compensation based upon the presumption of an adverse health effect. In any event, legislation without public education is doomed to failure.

Third, public opinion is a very potent force. Unfortunately, it is often manipulated by the media and by political forces. I am convinced that public education, including scientific disagreements, accomplishes more than scare tactics. All of us are afraid of the unknown making us easy prey for those with secondary goals to be achieved. Furthermore, public education should also include full disclosure of the consequences of the application of risk management.

Fourth, public policy should always clearly point out whether the goal is to manage or control risk or to eliminate it. One of the major defects of the Superfund Law is that it does not make that point clear, and the public is being misled.

Let me leave you with two important questions. First, should the Delaney Clause be rewritten, and if so, how? Is it consistent with modern science?

Second, how should our society deal with hypersensitivity reactions? Given that they occur, are our risk management strategies adequate? Is our science well enough developed to be sure that the adverse effects are truly hypersensitivity?

In closing let me point out that this symposium should make major contributions to our understanding of epidemiology, risk assessment and public policy decisions flowing therefrom. I commend all of you for participating.

Brandt

-8-

Chapter 28

Emerging Needs in Epidemiology
and Risk Assessment

Leon Gordis

Emerging Needs in Epidemiology
and Risk Assessment

Leon Gordis

I would like to commend the speakers for an exciting symposium. Dr. Stallones referred to the inadequacy of the information we have and the need for additional information. From a practical standpoint we should develop a means of capitalizing on populations which are being monitored but for whom the appropriate data are not being systematically collected at present. In different environments and settings we might be able to make existing data collection arrangements more productive for generating meaningful exposure data.

Dr. Nelson drew an interesting comparison of the attributes, both positive and negative, of toxicology and epidemiology and this topic has arisen several times in the last few days. As might be expected at a symposium with the title this one has, we sense a kind of breast beating by epidemiologists as they lay out the various areas of vulnerability of the discipline. I think it an excellent idea to discuss the problems of epidemiology as long as we do it in a constructive fashion. What concerns me is that a message is implied regarding toxicology: that aside from the problem of having to extrapolate across species, toxicology is free of the problems with which unfortunately epidemiology is endowed. Since I am not a toxicologist, I feel very free to talk about toxicology. Many of my friends tell me I always speak best when unencumbered by data and information.

Neither discipline has a monopoly on truth. In my view, toxicology shares many of the same problems that epidemiology has. I would like to mention a few examples. The first is the issue of sample size. The number of animals that are tested for any given dose is usually quite small. The number of doses at

which animals are tested may be very small, often a high, low and medium dose. So it may become extremely difficult to extrapolate from the findings both because of small sample sizes and limited number of ooses. The time needed for observing the effect may be a lifetime and this makes many animal studies very expensive and difficult.

We recognize that human studies are often afflicted with the problem of confounding and epidemiologists look with great yearning to the area of animal studies which do not seem to have confounding. But this is not really true. Confounding affects animal experiments to a serious extent. For example, in carrying out toxicologic studies, there may be great variability in the temperature and humidity at which animals are kept, and the quality of the air in the animal room, and the chow that the animals are fed often comes from different lots. These factors and the general problem of confounding usually are not mentioned in published papers. There may be differences in route of administration of the exposure, the vehicle in which the agent is administered and the stability of the agent in different experiments. So when data from different toxicologic experiments are pooled, the result is many of the same problems with which we deal in human population studies.

In addition, animal studies often use genetically inbred strains and are able to achieve, to a large extent, genetic homogeneity. However, this also poses a problem regarding the generalizability and meaningfulness of the findings regarding basic biologic mechanisms since the observations are made in highly inbred strains.

These are not criticisms of toxicology. The bottom line is that we have two disciplines: toxicology and epidemiology, and they complement each other. What we need is a greater collaboration between the two. It is important for people involved in policy and regulation to recognize that each discipline

has certain limitations and that the two together can be much greater than the sum of the two individually.

Dr. Omenn and others have commented on the issue of genetic variability . I believe that the genetic variability of a population is one of the very important problems that has not been dealt with adequately in epidemiology. We assume a uniformity of the denominator when we calculate rates and the fact of the matter is that we come up with average risks for large populations which really may have no relevance for people who want to know, "What is my risk?". We are not dealing with subgroups that are sufficiently refined on the basis of genetic characteristics. The dilemma is a trade-off: we lose power as we deal with smaller groups that share certain genetic characteristics and so we consequently trade off statistical significance for possible biologic significance. This is a problem that methodologically has to be addressed by epidemiologists and statisticians.

It has become quite clear in this symposium that epidemiology is extremely important in providing helpful data for regulatory policy. It is also important to use epidemiology for assessing the impact of regulation and policy once these policies have been initiated. If projections and risk assessments have been made, we should use the tools of epidemiology for following up populations after new policies are put in force and see whether indeed the original projections were justified. In that way we may learn how to make more rational policies in the future if we can see what we have accomplished and what mistakes we may have made.

Dr. Brandt talked about which studies should be done and certainly there are often political and legal reasons why studies are done. I would like to

suggest, and it may be a somewhat heretical, that certain epidemiologic studies should probably not be done. I do not believe that there is a fundamental need to study any association of any exposure with any outcome. Therefore, what should be the reason for doing a study aside from political and legal reasons? I believe there should be a rationale for doing the study on the basis of our current knowledge of the biology and epidemiology of the disease. To just grab some independent and dependent variables out of thin air, does not serve a valuable function and in fact has many negative implications. One can quote the aphorism of an anonymous sage who stated a long time ago, "If a study is not worth doing, it's not worth doing well".

Finally, in my view, epidemiologists have to be more concerned with the consumers of the epidemiologic data they are generating. It is not enough to say that our results diverge from those of A, B, and C and to leave it at that. The epidemiologist should be asking himself or herself, "Why do my results differ from the other studies? What does the regulator need to know to make effective use of the data that I am producing in the context of other data that are available in the literature?" If we look at epidemiology only as an intellectual discipline and disregard its applications, we are doing a disservice to the area of health policy. The training of epidemiologists must be brought closer to the implications and applications of epidemiologic data so that we can produce graduates who are sophisticated methodologically but in whom this sophistication rests on a strong biologic and public health foundation.

I would like to thank the sponsoring agencies for their foresight and for their generosity. Without their financial support this symposium could not have been held. I would also like to thank the Planning Committee and Program Committee and of course, the excellent speakers. Finally, I would like to thank you the audience for your interest, your enthusiasm, and your participation. In the spirit of this meeting, I wish you a zero risk trip home. Thank you.

Gordis-Closing