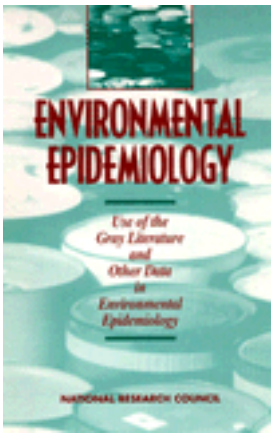

Environmental Epidemiology, Volume 2: Use of the Gray Literature and Other Data in Environmental Epidemiology

Committee on Environmental Epidemiology, National Research Council

ISBN: 0-309-52453-9, 200 pages, 6 x 9, (1997)

ENVIRONMENTAL EPIDEMIOLOGY

Volume 2

Use of the Gray Literature and Other Data in Environmental Epidemiology

Committee on Environmental Epidemiology
and
Commission on Life Sciences

National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C. 1997

NATIONAL ACADEMY PRESS • 2101 Constitution Ave., NW • Washington, DC 20418

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competencies and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to prcedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

International Standard Book Number 0-309-05737-X
Library of Congress Catalog Card Number 91-28051

Cover photograph: LES MOORE/UNIPHOTO

Printed in the United States of America

COMMITTEE ON ENVIRONMENTAL EPIDEMIOLOGY

ANTHONY B. MILLER (*Chairman*), University of Toronto, Toronto, ON, Canada
DAVID BATES, University of British Columbia, Vancouver, BC, Canada
THOMAS CHALMERS, Department of Veterans Affairs and Harvard School of Public Health (deceased December 1995)
JOHN FROINES, UCLA School of Public Health, Los Angeles, CA
DAVID HOEL, Medical University of South Carolina, Charleston, SC
JAMES MELIUS, New York State Department of Health, Albany, NY
JOEL SCHWARTZ, Harvard University School of Public Health, Cambridge, MA

LYNN GOLDMAN, US Environmental Protection Agency, was a member of the committee until February 1994

Special Advisers

ROBERT MORRIS, Medical College of Wisconsin, Milwaukee, WI
PAUL SCHULTE, National Institute for Occupational Safety and Health, Cincinnati, OH
DIANE WAGENER, National Center for Health Statistics, Hyattsville, MD

Staff

DEVRA LEE DAVIS, Scholar in Residence until June 1993
LINDA MILLER POORE, Research Associate
AMY REDMON, Editor
NORMAN GROSSBLATT, Editor
PAULETTE ADAMS, Project Assistant

The Committee on Environmental Epidemiology (listed above) prepared the original version of this report. The Commission on Life Sciences of the National Research Council, whose membership is shown on the next page, completed the report. All members of the Committee on Environmental Epidemiology agreed to the present content of the report.

Preface

VOLUME 1 OF *Environmental Epidemiology* was published in 1991 and has helped to define a field that seeks to clarify the relationship between exposure to physical, biologic, and chemical agents in the environment and human health. That report examined and evaluated the published scientific literature on health effects that could be linked with exposure to hazardous-waste sites and presented recommendations about filling major data gaps in order to advance the field.

In preparing volume 2, the Committee on Environmental Epidemiology set out to address important issues that were introduced in volume 1, such as the use of biomarkers and principles for drawing inferences from epidemiologic studies. The effects of exposure to an environmental agent can be hard to detect. The populations that have been exposed to the agent at a specific site are often small and the amount of exposure hard to determine. Only small changes in incidence of a disease may have occurred—so small as to make it difficult to determine clearly whether an association exists between the environmental exposure and the effect observed. Yet it is of great public-health importance to know if effects *are* occurring. Large numbers of people at many different sites may be exposed to the same environmental agent. Small effects detected in a small population could mean that a larger number exposed in the total population are at risk. Volume 2 continues the discussion of environmental epidemiology by examining ways to improve the chances of detecting an effect if one is occurring. Thus, the report focuses on improving how we measure exposure and how we apply the standard methods of epidemiologic research. The committee also examines the so-called gray litera-

ture—reports that have not been published in journals after peer review but may contain valuable clues about possible hazards to human health. These gray literature reports are often from state and local public-health groups and have usually been reviewed locally.

The committee that prepared volume 1 also prepared the bulk of the report that follows. In particular, it did the long and arduous work of reviewing a collection of gray-literature reports to determine how useful these might be in helping to understand the effects of environmental exposure. They also prepared the original versions of all other chapters. Unfortunately, long delays occurred in the latter stages of the study process, and the original committee could not complete the report. Responsibility for completion was then assumed by the Commission on Life Sciences of the National Research Council, the oversight body for the Committee on Environmental Epidemiology. Thus, though the original committee deserves the gratitude of the scientific community for initiating the preparation of this report and for its initial work, the Commission takes responsibility on behalf of the National Research Council for the contents of the report along with the original committee. Special thanks are due to those members of the Commission—John Bailar, Malcolm Pike, and Jonathan Samet—who played the central role on behalf of the Commission.

We acknowledge the efforts of and thank the Committee on Environmental Epidemiology and the staff of that committee. Their names are listed in the front of this report. We also thank the Agency for Toxic Substances and Disease Registry, sponsor of the study, for its support.

THOMAS D. POLLARD, *Chairman*
Commission on Life Sciences

Contents

The National Academy of Sciences is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce Alberts is president of the National Academy of Sciences.

The National Academy of Engineering was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. William A. Wulf is interim president of the National Academy of Engineering.

The Institute of Medicine was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and upon its own initiative, to identify issues of medical care, research, and education. Dr. Kenneth Shine is president of the Institute of Medicine.

The National Research Council was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and of advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing service to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both academies and the Institute of Medicine. Dr. Bruce Alberts and Dr. William A. Wulf are chairman and interim vice chairman, respectively, of the National Research Council.

1

Environmental Epidemiology: The Context

MANY CHRONIC DISEASES ARE of unknown etiology but may be related to environmental exposures. This report, the second of the Committee on Environmental Epidemiology of the National Research Council's Commission on Life Sciences, considers what information is needed to determine the prevalence of these diseases and whether they are a result of environmental exposures. The report highlights research opportunities and methodologic advances that will improve the scientific base of the evolving field of environmental epidemiology and examines the contribution of the so-called gray literature to our knowledge about possible links between environmental exposures and chronic diseases. The gray literature is defined as literature that is not "white" (available and cataloged), and that is not "black" (not available, unknown, or not obtainable). Gray-literature reports are usually produced in small quantities, intended for limited audiences, and not widely known; they are not indexed in standard sources, such as Medline. In the field of environmental epidemiology, these gray-literature studies may include such items as state health-department reports and doctoral and master's theses. One of the committee's criteria for determining the quality of the information in papers in the gray literature was whether the papers would meet currently accepted peer-review standards for publication in journals. The twin focus on the gray literature and epidemiologic methods fulfills the charge to the committee for its second report to examine state health-department reports and similar analyses that are not generally available in the peer-reviewed literature. As undertaken by the committee, the examination of the gray literature involves developing criteria for

requesting a few state health departments to provide reports for examination, developing criteria that help to define adequate epidemiologic studies, applying those criteria to selected state studies, and assessing those studies.

INTRODUCTION

Most people in the developed and developing worlds endorse environmental protection (Dunlap, 1992). Public concern about the environment and its relation to human health is demonstrated by the public reaction to reports of contamination at places like Love Canal, Times Beach, and Stringfellow Pits in the United States.

The nature of environmental-health concerns differs considerably between the developed and developing worlds (Doll, 1992). In the former, cigarette-smoking constitutes a major cause of illness and death, and occupational hazards, environmental tobacco smoke, lead, and other air pollutants adversely affect public health. Levels of lead thought to be safe a decade ago are now believed to produce permanent damage to children's intellectual potential (ATSDR, 1990).

In the developing world, concern focuses on basic sanitation, pure air, and clean water. The problems are traceable to a large extent to infectious agents, but exposure to toxic substances plays a role. The World Bank estimates that 1 billion people lack safe water, 1.7 billion are without adequate sanitation, 1.3 billion are exposed to unsafe soot and smoke, and 700 million women and children are exposed to severe air pollution from cooking fires (World Bank, 1992).

The incomplete understanding of causes of many common chronic diseases in both developed and developing countries fuels interest in identifying avoidable environmental hazards. Thus, more than 60% of all cases of birth defects are of unknown or poorly understood etiology (NRC, 1989a), as are many cases of degenerative neurologic diseases (NRC, 1992a), adult-onset asthma (NRC, 1992b) and other chronic respiratory diseases (NRC, 1989b), and renal and hepatic diseases. With respect to reproductive health generally, an array of end points are of concern, ranging from effects on offspring to reproductive health in males and females, including sexual maturation, onset of menses, menopause, sexual functioning, and endometriosis. Although these events are often discussed in an atmosphere of high public concern, suspect environmental factors must be studied with strict adherence to scientific canons of independent, verifiable research.

Although genetic predisposition and poor nutrition are important risk factors for many chronic diseases and other adverse health outcomes, the possible contribution of preventable or controllable environmental fac-

tors to these health outcomes needs to be clarified. Etiologic patterns in time and place of many diseases often cannot be determined, because appropriate data, methods, and theoretical frameworks have not been developed to evaluate such patterns. Thus, a top priority for environmental-health research is to identify better ways to look for environmental factors that may contribute to disease.

DEFINITION OF ENVIRONMENTAL EPIDEMIOLOGY

Modern epidemiology, the study of disease patterns in populations, encompasses a broad array of subject matter, including subspecialties that concentrate on such domains as clinical trials of pharmaceutical agents; such outcomes as reproductive and developmental effects, infectious diseases, and chronic diseases; such risk factors as occupation, nutrition, and alcoholism; and special populations. Thus, epidemiology includes controlled clinical evaluations of different treatment methods; comparative assessment of lifestyle factors, such as smoking, drugs, and drinking habits; estimations of the risks of occupational factors; and cross-sectional and time-series analyses of factors that may affect health.

Whatever the subject of epidemiologic evaluation, the basic theoretical and general principles are concerned with evaluating the statistical and biologic importance of variations in the frequency of occurrence of illnesses and related phenomena of health and health care. Epidemiologic study involves examination of the extent to which observed rates of a given phenomenon differ significantly from those that would be expected under specified conditions (Miettinen, 1985).

Interest in the application of epidemiology to the study of environmental hazards is increasing because epidemiologic studies can validate the models used in predicting hazards and can characterize the actual and potential health effects of such exposures.

The same definition of *environmental epidemiology* is used here as in the first volume (NRC, 1991), that is, the study of the effect on human health of physical, biologic, and chemical factors in the external environment. By examining specific populations or communities exposed to different ambient environments, environmental epidemiology seeks to clarify the relation between physical, biologic, and chemical factors and human health.

PURVIEW OF THIS REPORT

As in the first report, this volume excludes or limits the attention it gives to some environmental-epidemiology topics that have been the subject of other National Research Council committees. Thus, little attention

is given to issues related to toxic air pollutants and pesticide exposure of children. We do not deal with voluntary behavioral exposures, such as smoking, or with subjects, such as radiation, that have been dealt with in recent NRC reports.

This report considers the needs for research in environmental epidemiology in general, extends the assessment of information on the health effects of exposures from hazardous wastes started in the first volume up to the middle 1990s, and includes selected reports in the gray literature. Most of the gray literature examined here comes from state-generated studies or, if they are available to the public, analyses conducted by researchers for use in legal proceedings. Some of these contain critical information on subjects of interest to the committee. Also, as part of our review of the gray literature, we consider recent reports from the World Bank and the Pan American Health Organization regarding the health consequences of environmental contamination in central Europe, Central America, and China. The committee recognized its obligation to subject such studies to its own peer review. Thus, at least two committee members, expert in the relevant field of a study but not authors of the study, examined each of the reports from the gray literature that are cited in this report.

The decision to include gray literature arises because the committee found in preparing its first report that reports from a substantial number of relevant studies had not appeared in the conventional peer-reviewed literature. As discussed in more detail in chapter 7 of this volume, the committee solicited certain key reports from selected states on the epidemiologic study of hazardous-waste exposures. The committee then reviewed these to see how well they met its criteria for acceptable gray literature. The criteria are that a report be about an epidemiologic study involving a community or residential population, that it have some sort of peer review or other evidence of quality control, that the authors tried to collect exposure data, and that it have been published since 1980.

The first report of the committee considered findings in the published literature about hazardous waste in relation to the legislative mandates of the Environmental Protection Agency (EPA) and the Agency for Toxic Substances Disease Registry (ATSDR). The committee concluded that, as a practical consequence of these mandates, EPA and ATSDR tend to focus chiefly on the general estimation of exposure from and management of hazardous-waste sites and only indirectly consider potential public-health implications of exposures from such sites. As this committee and the US General Accounting Office (GAO, 1991) have observed, this is because inadequate resources have been devoted to characterizing human health risks possibly associated with exposure to hazardous wastes. Of more than $4 billion spent annually on the Superfund Program, less than

1% has been applied to the development of an active public-health program. In particular, exposures potentially critical to public health have been sparsely documented.

STRUCTURE OF THIS REPORT

Chapter 2 reviews basic epidemiologic methods that can be applied to environmental problems and comments on some well-recognized problems that environmental epidemiology faces, including small numbers of persons exposed, agents that have not been well characterized, and concern with small increases in risk. It states principles for inferring causation and discusses the types of evidence needed for environmental-epidemiology studies.

Chapter 3 expands on the needs for improving exposure-assessment information for environmental epidemiology and discusses these vis-à-vis specific epidemiologic-study designs.

Chapter 4 identifies health outcomes that should be subjected to environmental-epidemiologic study, ranging from chronic diseases of poorly understood etiology to those for which some causes are known. It identifies research opportunities for using biologic markers to study environmental factors that may be relevant to several chronic diseases, as well as for improving exposure-assessment information.

Chapter 5 considers existing data systems that are relevant to the research needs of environmental epidemiology.

Chapter 6 discusses several areas where improvements in methodology will advance the field of environmental epidemiology.

Chapter 7 presents a summary of selected gray-literature reports on hazardous wastes. It reviews several state studies of reproductive outcomes to illustrate the constraints on state health departments in carrying out environmental-epidemiologic research.

Chapter 8 recapitulates the major conclusions and recommendations of the committee.

SPECIAL ISSUES FOR THE STUDY OF ENVIRONMENTAL EPIDEMIOLOGY

As indicated in the first report, an optimal analysis of potential public-health consequences of suspect exposures proceeds from an assessment of past, current, or future exposures to the formulation of testable hypotheses of effects to be studied in one or more specific populations. An ideal environmental-epidemiology assessment considers all possible adverse health effects in exposed and unexposed persons and includes relevant contributing factors, including those that could confound, or interfere

with, the interpretation of results. Few studies meet this ideal; this limitation is also common in peer-reviewed, published reports.

The committee relied on a combination of evidence from different sources to assess the impact on public health of exposures suspected of causing symptoms or disease. The types of information from these sources are discussed at length in chapter 2 of volume 1.

SMALL RELATIVE RISKS, BUT LARGE NUMBERS OF CASES

Increasingly, environmental epidemiology concerns the search for factors that might moderately affect the risk of common multifactorial diseases. The effect of an individual environmental exposure on the relative risk of a disease may be small, but this does not mean that it is inconsequential; it can affect very large numbers of people and thus be associated with large numbers of cases of disease. For example, the risk of death in males aged 45-74 years with a diastolic blood pressure of 95 mm Hg in the Framingham study was only about 1.15 times the risk in those with a diastolic blood pressure of 85 mm Hg, yet the amount of disease that could be prevented in the population by reducing diastolic pressures to 85 mm Hg would be substantial. Increased use of hypertension medication, along with improvements in diet and exercise, is thought to be responsible for some part of the substantial decline in cardiovascular mortality in the last 20 years.

Large sample sizes and long-term followup studies are generally necessary to demonstrate potentially serious effects that involve small increases in relative risk. The chronic effects of ozone exposure and the acute effects of particulate air pollution are instances in which the relative risk may be small, but the population disease burden may be substantial. Ozone levels are often excessively high in many urban and coastal areas in the summer, when millions of people are outdoors, so that even a 10% increase in relative risk will produce a large number of cases of disease. For further discussion on this point, readers are referred to chapter 4.

THE NEED FOR MORE-GENERAL MONITORING

Monitoring systems utilizing existing data sources, as we discuss in chapter 5, will be of increasing importance. The first report also noted that monitoring may be the only way to determine the extent to which disease rates have changed as a result of changes in environmental contamination.

The use of aggregate statistics is also critical to detect trends or patterns in environmental pollution and health consequences that are not apparent at a local level. For example, what was initially thought to be laboratory drift in the measurement of blood lead concentrations in the

Second National Health and Nutrition Examination Survey was later identified as a 37% decrease in the population mean blood lead level as a result of a decrease in the use of lead in gasoline (Annest et al., 1983). This 37% decrease nationwide was used to justify the elimination of lead from gasoline.

Detecting the effects of monitored changes in disease and exposure also often depends on the alertness of researchers. The 2.5-fold increase in mortality in the London smog episode in 1952 became apparent when an investigator compiled weekly mortality data. Similarly, hospital admissions for asthma in children were cut in half in a Utah valley when a steel mill closed down, and admissions returned to their previous level when it reopened; this was also not detected by clinicians or the state health department but required the examination of hospitalization data by an investigator (Pope, 1991).

DEVELOPING RELEVANT EXPOSURE GRADIENTS

Exposures to synthetic organic chemicals and other modern products cannot be accurately segregated by source, such as air, water, or soil. Rather, modern exposure scenarios often model multimedia, multitemporal levels of many complex chemical compounds. Further, physical and biologic characteristics of other environmental factors can influence uptake and total dose of chemicals. Thus, heat, meteorologic conditions, humidity, and particle size affect the extent of uptake of airborne contaminants, and water hardness, pH, acidity, volatility of contaminants, and other natural background factors affect exposure to materials in water.

Much work in the past has relied on assumed dichotomous, yes-no exposures, but it is not always possible to find and study populations that are entirely unexposed to some environmental agents. Epidemiologists must work closely with exposure analysts to generate meaningful gradients of exposures for such populations, including the use of models to improve exposure estimates. These models need to include environmental and biologic fate, population activity patterns, biomonitoring, and biomarkers. Wherever possible, models should be validated by monitoring carefully selected subsamples of the population under study. This will allow more-refined estimates of individual exposure to be used in population-based studies, as is discussed in chapters 3 and 6.

THE ROLE OF PUBLIC-HEALTH DEPARTMENTS IN ENVIRONMENTAL-EPIDEMIOLOGY RESEARCH

Many issues in environmental epidemiology are in the domain of departments of public health. A group of residents near a hazardous-waste

site, for example, may become concerned about odors from the site, seepage into their yards or basements, or various symptoms that they may attribute, in the absence of other information, to chemicals from the site. These concerns are likely to come to the notice of the local public-health department, usually with requests for reassurance in the form of a study. Thus, public-health departments may be asked to study diseases of unknown etiology where there may be insufficient evidence to incriminate the hazardous-waste site or an alternate source and for which there has often been insufficient time since initial exposures (if indeed the exposures were from a waste site) for the presumed latent period to be exceeded. The public tends to have unrealistic expectations of what an epidemiologic study can produce.

Unfortunately, most public-health departments are ill prepared to conduct epidemiologic research (Ozonoff and Boden, 1987), or other factors impede research:

- The staff of the public-health department may have little training in environmental epidemiology, environmental toxicology, or exposure assessment. Public concerns notwithstanding, most public-health practitioners will have been trained to cope with other important matters, such as the study of infectious-disease outbreaks, immunization, improvement of maternal and child health, or even the prevention of cancer, but these rarely have any direct relevance to the environment. These deficiencies in personnel and training can foster simplistic approaches to potential environmental-health problems that inappropriately apply the basic concepts of infectious-disease epidemiology rather than those appropriate to chronic disease.
- Even if outside specialists are found to perform an appropriate investigation, the resources available to such departments for investigation are usually insufficient. This may result in a limited investigation that is inconclusive, with the suspicion that, if a more thorough study had been conducted, a more definitive answer might have been achieved.
- Resources may also be lacking for adequate measurement of exposure. Regulators often do not take measurements for the primary purpose of assessing human exposure, but for other purposes, such as to determine compliance with an administrative requirement. A result is that the measurements, good for their intended purpose, may be inappropriate as a basis of studies of human health effects.
- Political and social considerations can impede the conduct of research. One group seeks assurance that no problem exists, while another seeks validation of its health concerns. Or the source of a possibly hazardous pollutant is an economic mainstay of a community.
- Even though it may be clear, as in a chemical spill from a railroad car in a populated area, that the public has been exposed to a possible

hazard, proprietary or "trade-secret" concerns may block access to information that is needed for prompt and appropriate remedial action.

- Finally, as emphasized in the first report, many exposed populations are so small, the period of observation so short, the exposure so poorly measured, or the outcomes so poorly defined that a verdict of "not proved" is all that can be attained. The situation is usually poorly suited for the conduct of research regardless of the efforts that are expended and the skills of the investigators.

Nevertheless, given the great and increasing variety of chemical, physical, and biologic pollution in the environment, the first indication of a hazard from a particular chemical or group of chemicals may still follow an investigation of some event, or cluster, by a public-health department or concerned citizens. Reports from state and local public-health groups may then enter the gray literature and in some cases be the starting point for research that is published in peer-reviewed journals. These locally initiated studies, although using small populations with poorly characterized exposures, may suggest an effect. Efforts to develop databases of such studies might serve 2 purposes: other parties interested in similar exposures could learn what studies are in progress, so as to increase sharing of information on study design; and appropriate analysis of the combined study findings might eventually become feasible, as discussed further in chapter 7.

The focus of this report is on environmental-health issues in the developed world, but it is recognized that in developing countries environment-related diseases occur along with the more predominant infectious and chronic diseases. Severe indoor particulate air pollution has been documented in the developing world at levels 100 times higher than the US standard of 150 $\mu g/m^3$ for particles less than 10 μm in diameter (PM_{10}), and outdoor pollution is also sometimes extreme. These high levels of air pollution, coupled with other disadvantages in the developing world, may account for the fact that acute respiratory disease is the second leading cause of death in children under 5 years of age in countries of the developing world (Leowski, 1986). Important risk factors for both infectious and chronic diseases include basic sanitation, living conditions and urban infrastructure, housing, air and water pollution, and working conditions. All of these are threatened by the sanitary burden that is being accumulated in developing countries. These are the areas in which hygienists have had a great impact in the past in developed countries.

CONCLUSIONS

For many chronic, degenerative diseases of potential interest in environmental epidemiology, data on rates of occurrence (incidence) in de-

fined populations are not routinely collected. In addition, exposure is rarely assessed in a manner compatible with the needs of epidemiologic investigation. Thus, it is often impossible to determine whether the incidence of a particular disease has changed in response to a new or changing environmental exposure. Where a gradient of exposure can be determined, the risk of disease can sometimes be related to dose in a specially designed study. However, in many instances, diseases of possible environmental etiology cannot be examined in relation to environmental factors until baseline disease incidences have been determined and appropriate measures or estimates of exposure have been developed.

Study of the health outcomes associated with environmental exposures suffers from a lack of sophisticated technology for assessing chronic effects, from basic methodologic limits of study designs, and from the highly charged climate in which such studies are at times conducted. Moreover, cross-sectional study designs, rather than the more-reliable cohort or case-control studies, are often required in environmental epidemiology. Until quite recently, federal and state support has focused on the need for rapid health assessments that do not necessarily comply with the requirements for environmental epidemiology. Few academic departments of epidemiology have concentrated on refinements in the techniques needed to improve environmental epidemiology, and those that do so must struggle with limited resources. The remainder of this volume outlines strategies that can remedy these problems.

REFERENCES

Annest, J.L., J.L. Pirkle, D. Makuc, J.W. Neese, D.D. Bayse, and M.G. Kovar. 1983. Chronological trend in blood lead levels between 1976 and 1980. N. Engl. J. Med. 308:1373-1377.

ATSDR (Agency for Toxic Substances and Disease Registry). 1990. Toxicological Profile for Lead. ATSDR/TP-88/17. Atlanta, Ga.: Agency for Toxic Substances and Disease Registry, Public Health Service, US Department of Health and Human Services.

Doll, R. 1992. Health and environment in the 1990s. Am. J. Pub. Health 82:933-940.

Dunlap, R.E., G.H. Gallup, Jr., and A.M. Gallup. 1992. The Health of the Planet Survey. Princeton, NJ: The George H. Gallup International Institute.

GAO (US General Accounting Office). 1991. Superfund: Public Health Assessments Incomplete and of Questionable Value. Report to the Chairman, Subcommittee on Oversight and Investigations, Committee on Energy and Commerce, House of Representatives. GAO/RCED-91-178.

Leowski, J. 1986. Mortality from acute respiratory infections in children under 5 years of age: global estimates. World Health Stat. Q. 39:138-144.

Miettinen, O.S. 1985. Theoretical Epidemiology: Principles of Occurrence Research in Medicine. New York: John Wiley & Sons. 359 pp.

NRC (National Research Council). 1989a. Biologic Markers in Reproductive Toxicology. Washington, DC: National Academy Press. 395 pp.

NRC (National Research Council). 1989b. Biologic Markers in Pulmonary Toxicology. Washington, DC: National Academy Press. 179 pp.

NRC (National Research Council). 1991. Environmental Epidemiology. Public Health and Hazardous Wastes. Washington, DC: National Academy Press. 282 pp.

NRC (National Research Council). 1992a. Environmental Neurotoxicology. Washington, DC: National Academy Press. 154 pp.

NRC (National Research Council). 1992b. Multiple Chemical Sensitivities: Addendum to Biologic Markers in Immunotoxicology. Washington, DC: National Academy Press. 202 pp.

Ozonoff, Boden, 1987. Truth and consequences: health department responses to environmental problems. Science, Technology and Human Values, 12:70-77.

Pope, C.A., 3rd. 1991. Respiratory hospital admissions associated with PM_{10} pollution in Utah, Salt Lake, and Cache valleys. Arch. Environ. Health 46:90-97.

World Bank. 1992. World Development Report 1992: Development and the Environment. New York: Oxford University Press. 308 pp.

2

Environmental-Epidemiology Studies: Their Design and Conduct

THIS CHAPTER DISCUSSES THE ORIGINS of epidemiologic study and summarizes common analytic techniques. After a brief discussion of study designs and the types of information they produce, this chapter notes several difficulties for studies of environmental epidemiology, including the problems of studying small numbers of persons or rare diseases. We recommend that research on study designs focus on the improvement of statistical power or probability of detecting an effect. Finally, we review principles for inferring causation in epidemiology.

ORIGINS OF EPIDEMIOLOGY

Although early epidemiologic studies often focused on infectious diseases and death, epidemiology today has a much broader application, as "the study of the distribution and determinants of health-related states and events in specified populations and the application of this study to the control of health problems" (Tyler and Last, 1991, p. 12). Traditionally, epidemiology has been linked with disease prevention, in that its results can indicate risk factors that can be modified in order to control or eliminate certain diseases.

As chapter 1 indicates, environmental epidemiology is a logical extension of the field, expanding the range of concerns to include biologic, physical, or chemical factors that may be related to patterns of health and disease in populations. In general, environmental epidemiology is an observational rather than an experimental science; scientific deductions are drawn from patterns of occurrence. Its principal aim is to identify risk

factors that can be averted or reduced so as to prevent or reduce the risk of future disease and promote public health.

TYPES OF STUDIES IN ENVIRONMENTAL EPIDEMIOLOGY

Environmental-epidemiologic studies can be classified broadly into 2 categories that are not mutually exclusive: descriptive and analytic. Typically, descriptive studies are most useful for generating hypotheses and analytic studies most useful for testing hypotheses, though each type of study can be used for both purposes. Whether a study is hypothesis-testing or hypothesis-generating depends more on the sequence of past studies and the present state of knowledge (i.e., whether a hypothesis currently under evaluation was suggested by a previous study) than on the study design. Recent innovations in descriptive studies sometimes permit refined assessments of dose-response relations and etiologic factors.

DESCRIPTIVE STUDIES

Descriptive studies include case reports, surveillance systems, ecologic studies, and cluster studies (WHO, 1983).

Case Reports

A case report is a descriptive study of a single individual or small group in which the study of an association between an observed effect and a specific environmental exposure is based on detailed clinical evaluations and histories of the individual(s). These reports require few financial or personnel resources other than those of clinical medicine, and they may indicate whether additional study of a larger group of persons with similar health problems and exposures should be undertaken. However, the value of case reports is often limited because they lack a context of the disease in unexposed persons, variables such as time and dose of exposure are generally not known, and controls are absent. They are most likely to be useful when the disease is uncommon and when it is caused exclusively or almost exclusively by a single kind of exposure. In spite of these limitations, many known human environmental toxicants (e.g., methyl mercury, asbestos, tobacco smoke, and radon) first came to attention in case reports and series developed by astute clinicians, pathologists, and health workers. Public-health agencies must often investigate clusters of cases that are reported to them by private physicians and others. While case reports may not lead to identification of new causes of disease, they are more likely to point to specific hypotheses and to biologically meaningful associations if either the disease or the exposure is relatively rare.

Surveillance Systems

These systems provide broad-scale information on specified populations for which epidemiologic analyses can be conducted. Surveillance systems are generally designed to attain complete or nearly complete coverage of every identified instance of certain defined conditions in a defined population. Thus, they can be used to estimate the background incidence and prevalence of adverse effects, and trends can be analyzed across time and between populations or geographic areas.

Surveillance systems can identify increases or decreases in the occurrence of deaths from specific diseases and thus suggest or test hypotheses related to environmental exposures. For example, observations of a decline in age-adjusted stomach-cancer rates over time in the United States have stimulated the development of hypotheses about changes in dietary habits in the population as a whole, as well as about changes in the use of food preservatives and refrigeration (Howson et al., 1986) that might explain these trends. Similarly, after postmenopausal estrogen use fell in the United States, rates of endometrial carcinoma declined in women over age 65, lending support to an inference drawn from case-control studies that postmenopausal estrogen use increased the risk of endometrial cancer (Austin and Roe, 1982). In another instance, surveillance data from the National Center for Health Statistics suggested that a fall in blood-lead levels in US children was linked to a drop in gasoline-lead levels (Annest et al., 1983).

As public-health agencies have expanded the scope of surveillance systems (see chapter 5), it has become feasible to study the relationships between disease patterns and variations in environmental factors. Surveillance systems are expected to become increasingly common because the quality of their data is rising, statistical methods are improving, and costs are declining. The Agency for Toxic Substances and Disease Registry (ATSDR) has devised several surveillance systems to monitor the health of persons believed to have incurred exposure to such substances as trichloroethylene and dioxin. No results are yet available from those systems. If these exposure registries are to produce valuable results, they will need to include sufficient numbers of persons over a long enough period for diseases of interest to manifest themselves in numbers sufficient to demonstrate that some problem exists or that the problem is unlikely to exist and be large enough to cause serious concern.

Ecologic Studies

Ecologic studies explore the statistical connection between disease and estimated exposures in population groups rather than individuals. They combine data from vital records, hospital discharges, or disease registries

with grouped data or estimates of exposure, such as factory emissions in a given geographic area, proximity to waste sites, or air or water pollution levels. Observed associations may provide support for further investigations. Ecologic studies suffer from serious weaknesses: they assign group exposure levels to all members of the group, fail to control for individual confounding factors, use necessarily crude estimates of exposure, and may not capture the relevant exposure at the time of disease induction. Although some population groups in ecologic studies may appear similar to "cohorts" (see below), they lack the individual data that permit their analysis in a cohort study. The "ecologic fallacy" refers to drawing inferences incorrectly from data on groups or about individuals in the groups.

Several advances have facilitated an increase in the number of ecologic studies, including the development of surveillance systems, improved environmental-exposure databases (e.g., by ASTDR), and increased availability of sophisticated tools, such as geographic information systems. The value of ecologic studies may be strengthened as methods for estimating exposure are improved. Where valid proxies for gradations of exposure and relevant confounding variables can be devised, the ecologic fallacy may be reduced or overcome.

Ecologic investigations have provided important clues about causal associations even though these studies can be difficult to interpret. For example, fluoride was first found to prevent dental caries on the basis of observed correlations between geographic variations in natural levels of fluoride and rates of tooth decay (Dean et al., 1942). Similarly, rates of cardiovascular disease and cancer among immigrants have been correlated with those of their newly acquired compatriots, suggesting that changes in dietary and other factors are involved. Further refinements in the parameters of interest in ecologic studies might permit these studies to generate more-precise indications of associations between risk factors and disease (Greenland, 1990).

Studies of health problems in relation to fixed sources of environmental exposure have often relied on either labor-intensive techniques, such as personal interviews, or much more general classifications, such as ZIP code or town of residence. The latter approach has the obvious problems of errors in classification of actual residence and of including people in the exposed category who live far from the site and have little opportunity for exposure. This difficulty has been partially overcome by including better geographic-location information as part of state and federal lists of potential sources of environmental exposure (e.g., Superfund sites) and by public availability of more-complete coding of geographic information in US census data.

This additional information, along with the availability of improved mapping software, has greatly improved our ability to link health data,

such as cancer incidence with residence near a source of environmental exposure. With good geographic coding, disease cases and controls can be readily and quickly located in relation to an environmental source so that various measures of distance and direction can be studied. Data on the number and characteristics of people living in the area can also be obtained from census data.

Cluster Studies

A cluster study is a descriptive study of the population in a geographic area, occupational setting, or other small group in which the rate of a specific adverse effect is much higher than expected. Further, the group is often defined after the fact; that is, the "cluster" comes to attention, and the group is then defined so as to include it. Thus, clusters usually have the drawbacks of small samples. Cluster studies suffer from a major tautology: the data that inspire a hypothesized relation between a given exposure and a specific health outcome tend to be used to test this hypothesis, and then exposure and risk-factor data may be generated for persons defined to be in the study group and, usually, in some control group. For example, a reported cluster of cardiac birth defects that occur near a hazardous-waste site may be "tested" by comparing the measured rates of these defects in the same given geographic area with those from outside the same area. This is a highly unreliable approach methodologically and statistically, as the sample being studied has not been randomly selected. Nonetheless, the approach can be useful when the relative risk is extremely high, and it can be useful in developing hypotheses for study with other data. Many occupational hazards were first identified because clusters of disease were detected in specific workplaces, and other environmental diseases may also be ascertainable through cluster analysis.

ANALYTIC STUDIES

In contrast to descriptive studies, analytic studies are based on more individually detailed data from individuals that can be used to control for confounding, and they are usually more costly and labor-intensive. Information from medical records, clinical or laboratory investigations, questionnaire results, or direct measures or estimates of exposures may allow analytic studies to explore hypotheses about suspected causes of disease or identify and measure risk factors that increase the chance that a given disease will occur. Analytic studies may also be a source of additional specific hypotheses, often leading to a sequence of studies, the more recent being designed to attempt to refute hypotheses raised by earlier studies.

The classic designs of analytical studies are case-control and cohort studies. In addition, 2 "hybrid" designs—nested case-control studies and case-cohort studies—can be based on identified cohorts.

Case-Control Studies

Case-control studies compare exposures of individuals who have a specific adverse effect or disease with exposures of controls who do not have the effect or disease; controls generally come from the same population from which the cases were derived. There is an extensive literature on the design of case-control studies, including selection of controls, correction for confounding, statistical methods for analysis, and presentation of measures of effect, usually the odds ratio (Schlesselman, 1982). These studies generally depend on the collection of retrospective data. They may suffer from recall bias, i.e., the tendency of people who have a disease to remember putative causes more readily than those without a given disease. However, it is often possible in a case-control study to collect histories of exposure to many different factors and control for confounding more efficiently than in a large cohort study, where the costs of collecting substantial exposure data from all the members of the cohort may be prohibitive. It is likely that case-control studies will be conducted with increasing frequency as new ways of characterizing exposure through the use of biologic markers are developed (see chapter 3), mirroring the development that has occurred in the last 2 decades in other areas of epidemiology.

Cohort Studies

These studies identify a group of persons called a cohort, or sometimes several cohorts with differing kinds of the exposures of interest. Sometimes, a control group has zero exposure. The cohort study evaluates associations between the exposure(s) and 1 or more health outcomes in the cohort(s). In a cohort study, individuals with differing exposures to a suspected risk factor are identified and then observed for the occurrence of certain health effects over some period, commonly years rather than weeks or months. The occurrence rates of the disease of interest are measured and related to estimated exposure levels.

Cohort studies are of 2 kinds—retrospective and prospective—each with advantages and disadvantages. The retrospective (or historical) cohort study relates a complete set of outcomes already observed in a defined population to exposures that occurred earlier; data on both exposure and outcomes must be available at the time the study is undertaken. Prospective cohort studies, in which current exposure is directly measured

and individuals are then followed, have a potential for more-accurate measurements but may suffer from loss of subjects to followup or bias in ascertainment of end points. Also, it may be necessary to wait for many years or even for the time of followup to exceed the latent period between exposure and effect or for sufficient outcome events to occur.

Cohort studies can utilize questionnaires or laboratory tests to measure both exposure and outcome. One advantage over case-control studies is that multiple outcomes can be evaluated simultaneously in relation to the exposure data. However, the power to test associations will depend on the frequencies of the different outcomes considered, which in turn depend on the number of persons followed (see discussion below on power considerations).

One type of cohort study seeks to correlate time trends in outcome measures and environmental exposures. Such studies can be divided into 3 broad classes: those in which the outcome is estimated or measured relatively few times, those in which outcome variables are linked to episodic variations in exposure, and those in which long-term time trends in measures or estimates of health outcomes are linked with variations in monitored or estimated exposures. The first class is seen in some cardiovascular studies in which determinations of health status are made annually. Outcome measures are often continuous, as well as dichotomous. Other examples are those that correlate the development of chronic bronchitis with exposure to air pollution and prospective cohort studies that follow children's lead exposure and cognitive development from conception or birth. The second broad class examines changes in response to exposures that are episodic or of short duration. Studies that link peaks in air pollution to patterns of asthma fall into this category. The third broad class is similar to time-series studies often conducted in the social sciences. In such studies, both exposure and outcome measures are collected, perhaps on a daily basis, for periods of months or even years. Short-term fluctuations in those outcomes are correlated with short-term variations in environmental exposures. For instance, studies of changes in peak respiratory flow, respiratory symptoms, hospital admission, and daily mortality can be linked to changes in environmental air pollution. In most of these studies, the multifactorial nature of the outcome means that the explanatory power of each environmental variable is generally small. This has necessitated relatively large samples and careful modeling to avoid potential confounding.

Nested Case-control Studies

These studies are similar to ordinary cohort studies except that only a sample of controls (persons free of the disease) are studied in detail. They

generally use old cases in a defined cohort that has been followed long enough for sufficient outcome events to have occurred but only a random sample of cohort members who were eligible to become cases but had not developed the disease or died at the time the corresponding cases were diagnosed. Controls are often matched to cases on 1 or more potential confounders (e.g., age, sex, and smoking status) that the investigator does not wish to study. An individual selected as a control may become a case if the disease of interest develops. Nested case-control studies can be designed to have almost as much statistical power as the cohort study from which they are derived because of tighter experimental control, and they can be used to derive better inferences on exposure-disease associations. These studies may also be substantially more economical if the determination of exposure of the controls can be limited to a sample.

Case-cohort Studies

In this design, a random sample of the total cohort is drawn and taken to represent the exposure experience of the cohort. When the cohort has been followed long enough to accrue sufficient cases for analysis, the exposure experience of this subcohort is compared with that of the cases (who arise from the total cohort and might or might not be individuals in the subcohort who become cases). This design also provides economies in obtaining exposure data compared with a cohort study, but surveillance of the total cohort is still needed to identify the cases that occur.

SPECIAL CONSIDERATIONS

Many epidemiologic studies explore the relation between risk factors and health outcomes, often examining the relation between a single exposure and a single factor or disease. In environmental epidemiology, however, both exposures and outcomes are usually multiple. Many of the risk factors of interest derive from large-scale data sets on environmental pollution that involve continuous variables, as well as a variety of clinical health indicators. Much of cancer epidemiology has focused on studying specific anatomic sites of cancer and delineating important contributors to specific types of cancer, such as the link between occupational exposure to benzene and leukemia or that between asbestos and mesothelioma. Similarly, much of cardiovascular epidemiology has involved prospective cohort studies that concentrate on identifying a few specific risk factors.

CROSS-SECTIONAL DESIGNS

Many environmental-epidemiology studies are cross-sectional. In such designs, the relations between contemporaneous assessments of out-

come and exposure are studied; this can give rise to difficulties in determining the temporal aspects of an association. Often the exposure variable is measured continuously but with substantial error. The outcome is generally multifactorial, requiring a large number of covariates, and can include a wide range of health effects for which standard nomenclature, coding, and test systems do not exist. Examples of such outcomes include neurologic outcomes used in studies of lead toxicity, outcomes of some pulmonary-function tests, and diaries of activity level. Environmental epidemiology often relies extensively on a complex of study designs, such as cross-sectional designs that meld both analytic and descriptive studies, and often considers multiple health outcomes as well as multiple exposure variables.

MOLECULAR-EPIDEMIOLOGY STUDIES

Recent advances in molecular biology provide new ways to identify and measure markers of exposure or outcome, such as DNA adducts or oncogenes, that are identified through molecular biology. Such data can be used in any of the epidemiologic methods, so such studies have been designated "molecular epidemiology" by molecular biologists. The committee addresses the utility of biologic markers of exposure further in chapter 3 and biologic markers of outcome in chapter 4. However, the committee notes that the application of molecular biology to humans as distinct from experimental animals does not in itself justify the term "molecular *epidemiology*." For a study to be classed as molecular epidemiology, it is essential that valid epidemiologic techniques and study designs be used, including the selection of study subjects from a defined population. This field can develop only if epidemiologists and molecular biologists collaborate in the design and conduct of such studies. In the absence of adequate implementation of both aspects, the term *molecular epidemiology* should not be used.

CONSIDERATIONS OF THE POWER OF STUDY DESIGNS

Before any study is undertaken, sound epidemiologic practice requires careful consideration of statistical power, that is, the probability that a given research study will be able to detect a true positive effect if it exists. A study's power depends on many factors, including the increases in risk of exposed persons for the outcome under study, the size of the population to be surveyed, and, for cohort studies, the duration of follow-up. The higher the expected relative risk (RR), the smaller the population that needs to be surveyed. Conversely, the larger the population studied, the smaller the RR that can be detected. Most environmental pollution

includes relatively low levels of exposure to complexes of poorly defined materials. Thus, an environmental pollutant is likely to be associated with relatively small risks, though it could affect large numbers of people.

At any given level of statistical significance, there is a relation among study power, sample size, prevalence of exposure, and expected rate of a given outcome. In general, studies of larger numbers of persons over longer periods are more likely to yield positive results than those involving smaller populations for shorter periods. However, even large studies with long followup will result in uncertain findings if exposure is poorly measured or misclassified (see chapter 3). The sample size needed to achieve a given study power is also related to whether exposure is measured as a dichotomous or continuous variable, to the variability in distribution of the exposure, and to the effects of confounders and errors in the measure of exposure. In general, larger samples are needed when exposure measures are not continuous, when the effects of confounders and errors of measurement cannot be taken into account, and when the adverse outcome is a rare event (Greenland, 1983; McKeown-Eyssen and Thomas, 1985; Lubin et al., 1988; Lubin and Gail, 1990). Finally, all statistical-power calculations depend on the critical assumption that bias in both exposure and outcome can be ignored; this assumption may be rarely true in practice.

Statistical-significance testing is used to assess the likelihood that positive results of any given study represent a "real" association. However, no matter which statistical tests are employed, common research designs all produce studies with fixed, known chances of making a type I error, that is, of finding a positive result when one does not really exist. This probability is called alpha and is generally determined by a statistician at the time the protocol is drafted. It is commonly set at 5%.

Of equal importance for environmental epidemiology is a consideration of the probability that a failure to find an effect is a false negative, or type II error. This often occurs when small numbers of persons are studied and when relatively low risks are involved. Statistical tests cannot determine whether or not an error has been made but can indicate the probability that an error could occur, called beta, if the effect is of some hypothetical size specified by the investigator. The power to detect an effect of that size, defined as 1-beta, depends on the alpha level of significance testing and the unknown relative risk. Tables have been devised to help determine the number of observations required to have specified power to detect an effect of specified size if an association exists (Fleiss, 1981). For any specific size of effect, the power of a study increases as the study size increases.

Many episodes of environmental contamination involve low relative risks and small numbers of people, so environmental-epidemiology stud-

ies often lack sufficient power to detect important effects. This makes the development of innovative techniques to combine results an important priority for the field.

Pvalues are measures of random uncertainty alone and are dominated by sample size and other power considerations. In observational epidemiology, the primary sources of uncertainty about whether an effect is present are confounding, selection bias, and similar problems. In contrast, measures of the size of a possible effect, such as regression coefficients or odds ratios, may be less sensitive to sample size. If associations are due primarily to confounding, investigators may report considerable variation in measures of effect across different studies and populations. Hence, in modern epidemiology these measures of effect, and confidence intervals for them, are given greater attention than Pvalues. Consistency in these measures across studies with differences in exposures to potential confounders can provide valuable clues about whether observed associations indicate cause-effect relationships.

A very severe problem in environmental epidemiology is known as "multiple comparisons." If the probability of an error with 1 comparison (Pvalue or confidence bounds) is kept at the traditional value of 5%, a research study that includes more than 1 such comparison has a higher chance of making at least 1 error. While statistical methods exist to remove this effect, they have an unintended and often devastating effect on statistical power. This matter is dealt with in many statistical texts, so we do not expand on it here.

CAUSAL INFERENCE IN EPIDEMIOLOGY

The previous volume elaborated on criteria relevant to drawing inferences from epidemiologic studies (see NRC, 1991, for general guidance on these studies). They are summarized here as follows.

STRENGTH OF ASSOCIATION

The strength of association measures the size of the risk that is correlated with a causal agent (exposure). It is typically expressed as the risk of an exposed person's incurring a disease compared with that of a nonexposed person. The most-common comparison measures are the standard mortality ratio (SMR), the odds ratio (OR), and the relative risk (RR). The larger the ratio (SMR, OR, or RR), the stronger the association between the inferred link of exposure to disease for exposed individuals. For example, an RR of 1.4 for lung cancer after exposure to environmental tobacco smoke indicates that exposed persons are 40% more likely to develop lung cancer than are nonexposed persons. The strength of associa-

tion must often be considered in relation to the population at risk and intensity of the exposure. For example, an RR of 4 that affects a small population may have a much smaller public-health impact than does an RR of 1.2 that affects much larger numbers. Epidemiologists are sometimes concerned with attributable risk, which is a measure of the rate of disease above the background rate that can be attributed to exposure. This is more difficult to detect, study, and estimate in environmental epidemiology because it is difficult to determine a baseline rate. Problems with using strength of association as the principal criterion for causality include the fact that misclassification and other biases can profoundly change the strength of association.

SPECIFICITY OF ASSOCIATION

Specificity suggests that the suspected causal agent induces a single disease. While this may apply to a few associations between exposure and disease (e.g., vinyl chloride and angiosarcoma of the liver), single diseases (e.g., lung cancer) can have many causes, and single agents can cause many effects (e.g., lead at high-enough levels can cause increased blood pressure, neurologic symptoms, reproductive effects, and kidney damage). Specificity can be diminished by inappropriate or inaccurate grouping of diseases in a way that obscures a real effect (e.g., grouping some rare forms of cancer with other cancers).

CONSISTENCY OF ASSOCIATION

The observed relation between exposure and disease is seen rather regularly in independently conducted studies; the value of consistency is enhanced if the studies are of different types and in different populations. For example, a study of the association between lung cancer and passive smoking may produce an RR of only 2.0 or less, but this elevated risk has now been reported in over 30 studies carried out in 6 countries (NRC, 1986). Because of the variety in study protocols and populations, claims of bias in all the studies have little credibility. Studies not having statistically significant results can be combined with similar studies, as long as they all use sound methods. Studies that meet the standards for good epidemiologic practice can be grouped for meta-analysis, which allows for statistical pooling of different studies.

TEMPORALITY

The exposure should precede the development of symptoms or diseases of interest by an appropriate interval. The time between exposure

and disease should be consistent with biologic understanding of the time from exposure to the observed disease. For example, tobacco typically causes lung cancer 25 years or more after the beginning of regular exposure, though a few cases have been observed within 10 years of first exposure (Doll and Peto, 1978).

BIOLOGIC GRADIENT OF RELATION BETWEEN ESTIMATED EXPOSURE AND DISEASE

In general, a greater exposure should cause a stronger (though not always proportional) effect. For example, smoking more cigarettes increases the risk of lung cancer. Typically, dose equals the concentration integrated over time. In some cases, however, dosing patterns can be more important than the overall dose in the relation between dose and response. Also, the timing of the exposure can be critical in the dose-response relation.

EFFECTS OF REMOVAL OF A SUSPECTED CAUSE

If a causal relation exists, removing the causal agent should reduce or eliminate the effect; if the effect is irreversible in individuals already exposed, this reduction may not be apparent until the exposed generation is largely removed from the study population by death or in some other way (e.g., limitation to persons under age 65). If different causes are related to a single disease, then the principle applies only to the specific causal factor removed.

BIOLOGIC PLAUSIBILITY

The relation between the suspected causal agent and suspected effect should make sense, given the current understanding of human biology. Animal studies or other experimental evidence can strengthen or weaken the biologic plausibility of the relation by demonstrating mechanisms of disease or determining whether the association between exposure and disease holds in experimental situations. However, lack of a known mechanism does not invalidate a causal association. For many diseases, the underlying mechanisms are unknown.

REFERENCES

Annest, J.L., J.L. Pirkle, D. Makuc, J.W. Neese, D.D. Bayse, and M.G. Kovar. 1983. Chronological trend in blood lead levels between 1976 and 1980. N. Engl. J. Med. 308:1373-1377.

Austin, D.F., and K.M. Roe. 1982. The decreasing incidence of endometrial cancer: public health implications. Am. J. Pub. Health 72:65-68.

Dean, H.T., F.A. Arnold Jr., and E. Elvove. 1942. Domestic water and dental caries. V. Additional studies of the relation of fluoride domestic waters to dental caries experiences in 4425 white children aged 12-14 years, of 13 cities in 4 states. Public Health Rep. 57:1155-1179.

Doll, R., and R. Peto. 1978. Cigarette smoking and bronchial carcinoma: dose and time relationships among regular smokers and lifelong non-smokers. J. Epidemiol. Community Health 32:303-313.

Fleiss, J.L. 1981. Statistical Methods for Rates and Proportions. New York: Wiley. 321 pp.

Greenland, S. 1983. Tests for interaction in epidemiologic studies: a review and a study of power. Stat. Med. 2:243-251.

Greenland, S. 1990. Divergent Biases in Ecologic and Individual-Level Studies. Paper presented at the Second Annual Meeting of the International Society for Environmental Epidemiology, August 12-15, 1990, Berkeley, CA.

Howson, C.P., T. Hiyama, and E.L. Wynder. 1986. The decline in gastric cancer: epidemiology of an unplanned triumph. Epidemiol. Rev. 8:1-27.

Lubin, J.H., and M.H. Gail. 1990. On power and sample size for studying features of the relative odds of disease. Am. J. Epidemiol. 131:552-566.

Lubin, J.H., M.H. Gail, and A.G. Ershow. 1988. Sample size and power for case-control studies when exposures are continuous. Stat. Med. 7:363-376.

McKeown-Eyssen, G.E., and D.C. Thomas. 1985. Sample size determination in case-control studies: The influence of the distribution of exposure. J. Chronic Dis. 38:559-568.

NRC (National Research Council). 1986. Environmental Tobacco Smoke: Measuring Exposures and Assessing Health Effects. Washington, DC: National Academy Press. 337 pp.

NRC (National Research Council). 1991. Environmental Epidemiology. Public Health and Hazardous Wastes. Washington, DC: National Academy Press. 282 pp.

Schlesselman, J.J. 1982. Case-Control Studies: Design, Conduct, Analysis. New York: Oxford University Press. 354 pp.

Tyler, C.W., Jr., and J.M. Last. 1991. Epidemiology. Pp. 11-39 in J. M. Last and R. B. Wallace, eds. Maxcy-Rosenau-Last Public Health and Preventive Medicine. 13th ed. Norwalk, CT: Appleton & Lange.

WHO (World Health Organization). 1983. Guidelines on Studies in Environmental Epidemiology. Environmental Health Criteria 27. Geneva: World Health Organization.

3

Exposure Assessment in Environmental Epidemiology

EXPOSURES TO CHEMICAL AND PHYSICAL agents in the environment can produce a wide range of adverse health consequences. Environmental epidemiology attempts to determine whether a hazard exists—that is, whether there is a causal relation between exposure to certain chemical or physical agents and adverse health effects—and to measure and characterize any causal relations (to assess the exposure-response relationship). Typically, a continuum between level of exposure and the size or probability of health effects is assumed. Emphasis is placed on characterizing the associations across the continuum, and quantitatively defining the relation is a central feature of the epidemiologic investigation. Assessment of exposure is then a crucial component of environmental epidemiologic research. The estimation of exposure in relation to health effects is frequently difficult, and it has generally received inadequate attention. However, a field of exposure assessment is emerging.

Exposure assessment for purposes of environmental epidemiology may differ from exposure assessment for site remediation, mitigation, control, and risk assessment. The differences are sometimes subtle but may have substantial impact on the conduct of studies and associated allocation of resources. Investigations for the purpose of risk assessment, for example, generally include information on the source and identity of chemical agents, the concentration of each toxicant in various media, and the toxicity of identified toxicants as defined in experimental studies. Mathematical modeling may be used to define breakdown, transport, and ultimate location as well as the potential health risk. Environmental epidemiology, on the other hand, is more often hypothesis-based research

that seeks to examine specific populations or communities to clarify the relation between health and physical, biologic, and chemical factors.

Volume 1 of this report (NRC, 1991a) described some limitations and problems in the quantitative estimation of exposure when the focus of a study is possible adverse consequences of chemical exposure from hazardous-waste sites. This chapter reviews some aspects of exposure assessment or analysis that are important in environmental epidemiology and that illustrate the central role of exposure assessment. This chapter also discusses opportunities to improve analysis of exposure.

The importance of exposure assessment has been underscored in several reports (NRC 1988, 1991a,b). An International Society of Exposure Analysis has been formed, and the Science Advisory Board of the Environmental Protection Agency (EPA) recommended that EPA develop a 5-year program on exposure assessment (EPA, 1988).

The National Human Exposure Assessment Survey (NHEXAS) is a federal interagency program to design and implement an exposure surveillance system for the US population. The overall goal is to obtain periodic and systematic measurements of population exposures to multiple chemicals, including data on important environmental media, pathways, and routes, so that we can accurately determine current status, document historical trends, and predict possible future directions for exposures to hazardous chemicals (Sexton, 1991; Sexton et al., 1995) The NHEXAS has 3 specific objectives: (1) to document the occurrence, distribution, and determinants of exposures to hazardous environmental agents, including geographic and temporal trends, for the US population; (2) to understand the determinants of exposure for potentially at-risk population subgroups, as a key element in the development of cost-effective strategies to prevent or reduce exposures (risks) deemed to be unacceptable; and (3) to provide data and methods for linking information on exposures, doses, and health outcomes that will improve environmental health surveillance, enhance epidemiologic investigations, promote development of predictive models, and ultimately lead to better decisions (Sexton et al., 1995). As pointed out by Burke and Sexton (1995), "NHEXAS represents perhaps the most comprehensive exposure surveillance initiative ever undertaken . . . it has been designed to address the information needs of regulators and improve the scientific basis for risk assessment, risk management, and risk communication." The "consolidated report" of an EPA-appointed consensus team on NHEXAS concludes as follows: "The implementation of NHEXAS can be considered a turning point in environmental policy. It represents the first concerted effort to understand and track total individual exposures on a national scale" (Burke et al., 1992). A complete description of the NHEXAS phase I field studies, as well as a summary of the rationale and justification for

NHEXAS was published in the *Journal of Exposure Analysis and Environmental Epidemiology* (1995; 5:229-444).

EPA has made a significant contribution to exposure assessment by the issuance of guidelines for exposure analysis (EPA, 1992). The guidelines describe general concepts of exposure assessment and have application to risk assessment, trends analysis, and epidemiology. Those guidelines and the NRC report *Human Exposure Assessment for Airborne Pollutants* (NRC, 1991b) are major contributions to the assessment of the impact of toxic agents on potentially exposed populations. They provide a broad overview of the need for exposure assessment. This chapter is not intended to repeat the contents of those documents, but will focus on certain specific issues in exposure assessment for use in environmental epidemiology.

PRINCIPAL CONCEPTS THAT UNDERLIE THE CONTENT OF THIS CHAPTER

It is relevant to state the assumptions that the committee used as the basis and context for this chapter.

1. The effective application of exposure assessment methods may improve the results of any epidemiologic investigation. As in any line of epidemiologic investigation, an improvement in exposure assessment can reduce bias and improve statistical power to detect adverse effects associated with exposure to environmental contaminants.

2. However, important findings may derive from environmental-epidemiologic investigation even when the exposure assessment uses only simple and crude tools to characterize the exposure of a given population. Overreliance on sampling of exposure of individuals may not be cost-effective and may limit the size of the study, with little improvement over the findings based on indirect methods.

CONCEPTS AND METHODS IN EXPOSURE ASSESSMENT

This section reviews some of the basic concepts inherent in exposure assessment. For a more detailed discussion, the reader is referred to the NRC report *Human Exposure Assessment for Airborne Pollutants* (NRC, 1991b), the EPA guidelines for exposure assessment (EPA, 1992), and the Agency for Toxic Substances and Disease Registry (ATSDR) *Guidance Manual* (ATSDR, 1994). Epidemiologic research uses various exposure metrics. The choice of a specific metric will depend on the type of study in question, the resources available to the investigator, the conceptual framework behind the investigation, and above all, biologic considerations. In deciding which exposure metric is best in a particular study,

one must be clear about basic concepts of exposure analysis. Exposure assessment for use in environmental epidemiology must attend to 5 primary issues: (1) the definition and characterization of the potentially exposed population; (2) the collection of quantitative information on population exposure, temporal characteristics, and dose-response relations; (3) the medium and the microenvironment of principal concern in terms of exposure; (4) the use of information collected in one population in assessing potential risk to others; and (5) the biologic plausibility of any hypotheses based on mechanistic considerations that can assist and help guide the exposure assessment.

ATSDR (1994) has developed a definition of exposure as "an event that occurs when there is contact at a boundary between a human being and the environment with a contaminant of a specific concentration for an interval of time; the units of exposure are concentration multiplied by time." NRC (1991b) and EPA (1992) have also developed definitions of potential dose, applied dose, internal dose, and biologically effective dose for purposes of exposure assessment. The terms are illustrated in figure 3-1. *Potential dose* is the amount of the chemical ingested, inhaled, or in material applied to the skin. *Applied dose* is the amount of a chemical that is absorbed or deposited in the body of an exposed organism. *Internal dose* is the amount of a chemical that is absorbed into the body and available for interaction with biologically significant molecular targets. *Biologically effective dose* is the amount of a chemical that has interacted with a target site over a given period so as to alter a physiologic function.

A concept important to any type of study is that of total exposure. Assessment of total exposure has received considerable attention in recent years (Lioy, 1990; NRC, 1991a; Wallace, 1991; Wallace et al., 1986, 1987, 1988). Total-exposure assessment consists of estimating possible exposure from all media (soil, water, air, and food) and all routes of entry (inhalation, ingestion, and dermal absorption). NRC (1991a) and Lioy (1990) have developed a conceptual framework for human total exposure assessment that may serve as a guide for assessing human exposure to environmental contaminants. The framework is outlined in table 3-1. This framework accounts for all exposures to a specific agent or group of agents that an individual may have had, regardless of the environmental medium. Total-exposure assessment has particular relevance in environmental epidemiology insofar as it facilitates identification of the principal medium or microenvironment of concern and provides information on potentially confounding exposures.

NRC (1991a,b) described the different measurement and estimation techniques used in exposure assessment. Categories were defined that include direct measurement of exposure (personal monitoring, biologic monitoring, and biomarkers); indirect measures (microenvironmental

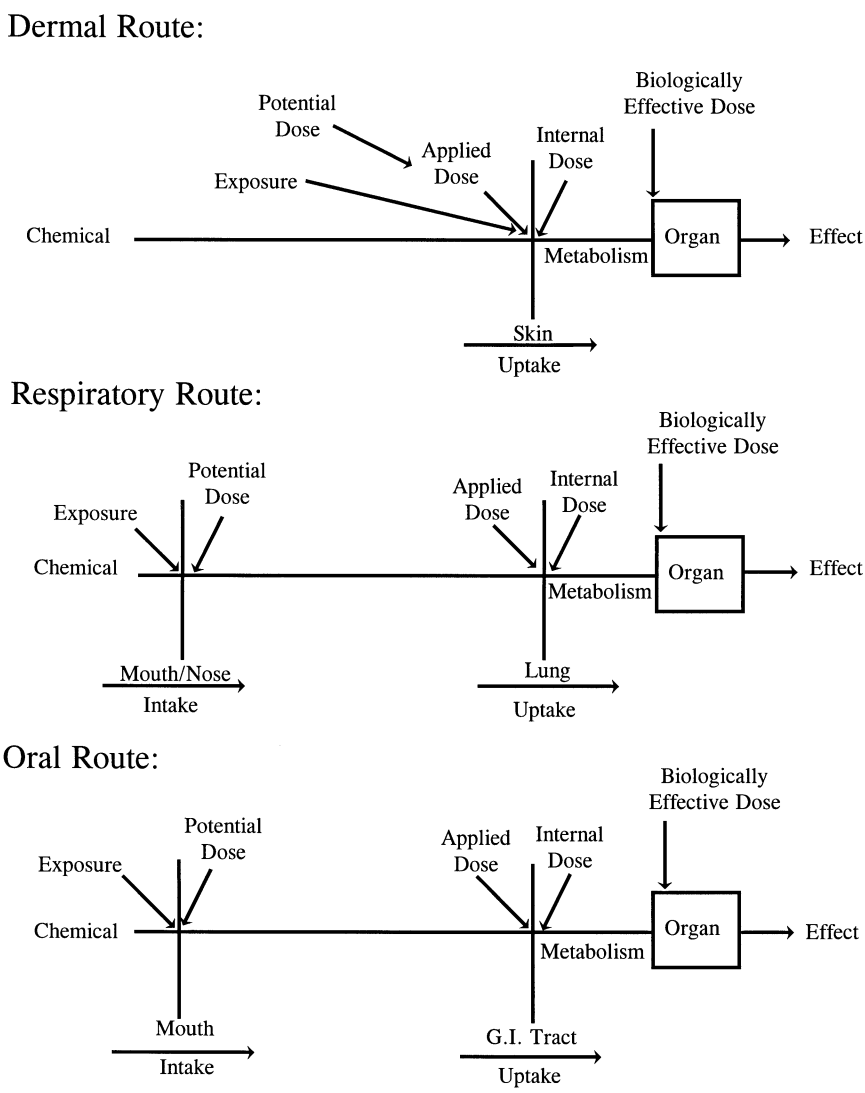

FIGURE 3-1 Schematic of dose and exposure.

monitoring coupled with exposure models, where microenvironmental monitoring is defined as the monitoring of contaminant concentrations in locations or media in which exposure occurs); that include mathematical modeling, questionnaire/diaries, or spatial factors, e.g., residence in a country or region or distance from a source of chemical contamination (figure 3-2). EPA (1992) has also provided examples of types of measure-

ments needed to characterize exposure-related media and associated parameters (table 3-2). Ryan (1991) has reviewed aspects of human exposure modeling that are useful for understanding the concept of exposure assessment.

When direct measurement of exposure is possible, it generally provides more-accurate information than indirect assessment of a particular individual's contact with a specific contaminant over time. The tradeoff is often between accurately measuring exposures over short periods, often outside the period of disease etiology, and indirect methods of assessing exposure over lengthier, more-relevant intervals. Personal monitoring has been widely used in workplace settings and can provide a measure of exposure across a range of microenvironments where individuals reside or work, though it is generally limited to a single chemical compound. However, personal sampling is often expensive, may demand extensive analytic capability and methodologies, and requires care in selecting study subjects. Biologic monitoring provides direct measures that integrate all routes of exposure to contaminants. Biologic monitoring may also provide more-precise estimates of target-organ dose, if appropriate toxicokinetic and metabolic information is available.

Indirect sampling uses exposure data available for defined areas or other microenvironments, generally from monitoring with time-activity information. Exposures in each microenvironment are weighted for the average time spent there and added to assess total personal exposure. Validation of the specific applicability of indirect monitoring is an important requirement for the successful use of this method. Some important epidemiologic studies have emphasized indirect measures of exposure as the primary linkage to health outcome. Some of these studies are reviewed below.

EXPOSURE-DATA NEEDS FOR EPIDEMIOLOGY STUDIES

NRC (1991a) has discussed exposure assessment in relation to the type of study being conducted: "The type of exposure assessment and the acceptable level of uncertainty in the data vary according to whether the assessment is designed to generate or test hypotheses about exposure, test instruments, make risk assessment decisions, or make regulatory decisions." Gann (1986) has asked, "What kind of exposure data do epidemiologists need?" He argues that the answer depends on the development of a well-defined research question. Bailar (1989) points out that the definition of dose-response relations is usually critical to establishing causality. It is expected that, when it is possible to examine ordered categories of exposure, higher doses will have greater effects so that "the dose-response relationship is monotonic." Thus, "departures from mono-

TABLE 3-1 Parameters Required to Calculate Potential and Internal Dose

Airborne contaminant	Water contaminant
I. Concentrations (μg/m^3,ppb) A. Microenvironments B. Personal	I. Concentration (μg/L,ppm) A. Tap water B. Water uses C. Effluent 1. Industrial 2. Commercial 3. Residential 4. Uncontrolled dumps
II. Patterns of exposure A. Intensity "episode" concentrations versus normal levels (average) B. Frequency and duration of contact	II. Patterns of exposure A. Drinking B. Swimming C. Cooking D. Bathing E. Laundry F. Showering
III. Transport A. Dispersion and advection B. Other meteorology related to removal rates (washout, fallout) C. Indoor ventilation and removal rates	III. Solubility of contaminant
IV. Chemistry A. Formation rates B. Transformation rates	IV. Volatility of contaminant
V. Deposition rate (μg/cm^2) A. Environmental B. Lung	V. Transport A. Groundwater B. Surface water C. Domestic supply
VI. Contact A. Inhalation (dependent on exercise regime)(m^3/time) B. Dermal deposition and permeability (μg/cm^2/time) C. Ingestion (food, soil)(μg/g/time)	VI. Chemistry A. Formation rates B. Transformation rates C. Degradation
VII. Absorption A. Within tissue B. Into the blood and other fluids	VII. Contact rate (μg/L/time) via exposure route A. Ingestion B. Skin C. Inhalation (volatilized)
	VIII. Absorption A. Dermal deposition and permeability B. Gastrointestinal tract

Source: Reprinted with permission from Lioy, 1990. Copyright 1990 American Chemical Society.

Soil and sediment	Food (commercial and homegrown produce)
I. Concentrations (μg/g) A. Dusts 1. Outdoor 2. Indoor B. Contaminated soil 1. Uncontrolled dumps 2. Airborne deposition 3. Landfills 4. Resuspension II. Patterns of exposure A. Frequency and duration B. Intensity of contact III. Percolation rate A. Soil composition B. Water table C. Solubility D. Transport IV. Volatilization A. Contaminant B. Soil composition C. Top soil and cover V. Contact rate via exposure route A. Dermal deposition and permeability B. Lung C. Gastrointestinal tract (pica) 1. Population 2. Abnormal ingestion behavior VI. Body parameter A. Lung volume B. Exposed skin surface (condition of skin) VII. Absorption A. Soil composition B. Contact and absorption rates All media: Can be supplemented by measuring a biological marker of accumulated single-medium or multimedia exposures in blood, urine, feces, and so forth. Many of these usually are nonmedia specific. Body weight: Used for lifetime exposure and dose calculation.	I. Concentrations (μg/g) A. Plants B. Vegetables and fruit C. Milk D. Animals and fish E. Cooked foods F. Beverages and water-based foods (μg/L) II. Patterns of exposure A. Rate (μg/L/time) (μg/g/time) B. Frequency C. Origin of food 1. Home grown 2. Commercial distribution 3. Local farms 4. Processed foods III. Source of contamination A. Naturally occurring contaminants B. Airborne deposition C. Fertilization D. Pest control E. Waste dumps F. Water supply G. Preparation and cooking techniques IV. Contact rate A. Gastrointestinal (GI) B. Inhalation (cooking only) V. Absorption through GI tract

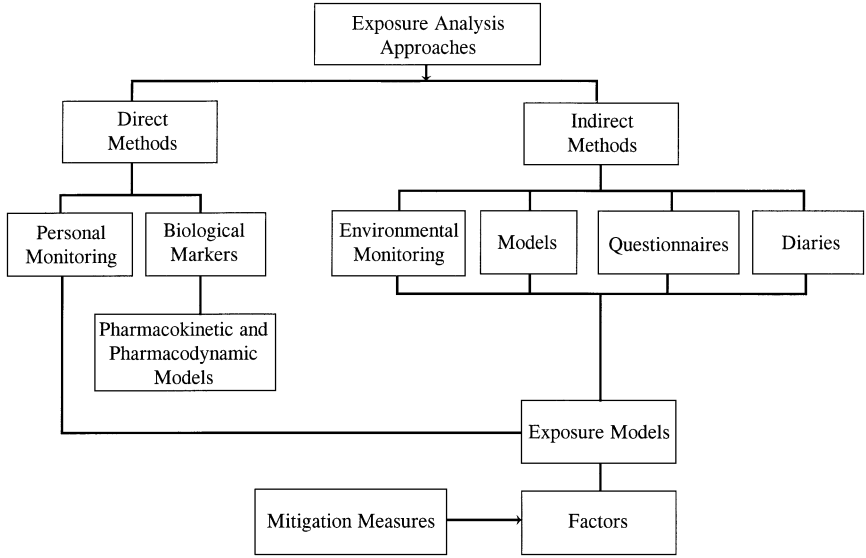

FIGURE 3-2 Possible approaches for exposure assessment.

tonicity raise questions about causality." Therefore, a key purpose of exposure assessment is often to support evaluation of dose-response relations.

DESCRIPTIVE EPIDEMIOLOGIC STUDIES

Exposure data for descriptive epidemiologic studies must fairly apply to the population from which the disease is arising. This requires estimation of the whole probability distribution of exposure, not just means, with special attention to accurate estimation of the upper end of exposures. This is especially important if many or most persons are thought to be exposed at levels below some "threshold" where effects first appear. This is best achieved by random sampling, and sampling theory will dictate the nature and size of the sample, as well as any repetitions over time or circumstance, from which exposure will be assessed.

Unless exposure information over time is already available or obtainable, data from an investigation that measures exposure "now" must be used to infer exposure levels at earlier times, e.g., the times when cases of the disease were induced. If the distribution of induction times is not known, the use of current exposures may be highly uncertain.

In some instances, disease rates in an exposed population are compared with disease rates in unexposed population units, but inferences are stronger if the investigator classifies risk by some population grada-

tion of exposure. Sometimes, if exposure is measured in only a few areas, analysis may have to be limited to simple "exposed-control" comparisons. Sometimes, as may often happen in studies of pollutants in air or diet, there are no unexposed population units, and low- (or lower-) exposure groups fulfill the role of a "baseline" group.

The sophistication of the exposure information collected will in general depend on resources. Environmental modeling may be necessary to deal with area and temporal variation. Analytic methods to cope with these problems are discussed in chapter 6.

ANALYTIC EPIDEMIOLOGIC STUDIES

The types of analytic studies used in environmental epidemiology were summarized in chapter 2. Here the particular exposure requirements of these studies are considered.

Case-control Studies

In case-control studies, the past exposures of cases and controls will have to be estimated, using historical records, if available, or current exposure measurements extrapolated back in time. Quantitative measures of exposure can reduce misclassification and allow the development of a dose-response curve.

Cohort Studies

Requirements for historical cohort studies are similar to those for case-control studies. For prospective cohort studies, there may be a need to estimate the extent of continuing exposure. In view of the large numbers required for cohort studies, resource constraints may make it impossible to do more than measure current or recent exposure.

As we pointed out previously (NRC, 1991a), the identification of potentially hazardous exposures in a group often results in cessation of exposure. This does not remove the need for characterization of past exposures.

Nested Case-control and Case-cohort Studies

Sometimes in a cohort study it is possible to collect specimens that could characterize exposure (e.g., biologic markers), where the expense largely resides in the analysis rather than specimen collection. When stored specimens are available (see NRC, 1991c, for a discussion of quality assurance associated with specimen archives) and can be analyzed after

TABLE 3-2 Examples of Types of Measurements to Characterize Exposure-Related Media and Parameter[a]

Type of measurement (sample)	Usually attempts to characterize (whole)
A. For Use in Exposure Scenario Evaluation:	
1. Fixed-Location Monitoring	Environmental medium; samples used to establish long-term indications of media quality and trends.
2. Short-Term Media Monitoring	Environmental or ambient medium; samples used to establish a snapshot of quality of medium over relatively short time.
3. Source Monitoring	Release rates to the environment from sources (facilities). Often given in terms of relationships between release amounts and various operating parameters of the facilities.
4. Food Samples (also see #11 below)	Concentrations of pollutants in food supply.
5. Drinking Water Samples	Concentrations of pollutants in drinking water supply.

Examples	Typical information needed to characterize exposure
National Stream Quality Accounting Network (NASQAN),[b] water quality networks, air quality networks.	Population location and activities relative to monitoring locations; fate of pollutants over distance between monitoring and point of exposure; time variation of pollutant concentration at point of exposure.
Special studies of environmental media, indoor air.	Population location and activities (closely matched to variations in concentrations); fate of pollutants between measurement point and point of exposure; time variation of pollutant concentration at point of exposure.
Stack sampling, effluent sampling, leachate sampling from landfills, incinerator ash sampling, fugitive emissions sampling, pollution control device sampling.	Fate of pollutants from point of entry into the environment to point of exposure; population location and activities; time variation of release.
FDA Total Diet Study Program,[c] market basket studies, shelf studies, cooked-food diet sampling.	Dietary habits of various age, sex, or cultural groups. Relationship between food items sampled and groups (geographic, ethnic, demographic) studied. Relationships between concentrations in uncooked versus prepared food.
Ground Water Supply Survey,[d] Community Water Supply Survey,[e] tap water.	Fate and distribution of pollutants from point of sample to point of consumption. Population served by specific facilities and consumption rates. For exposure due to other uses (e.g., cooking and showering), need to know activity patterns and volatilization rates.

continued

TABLE 3-2 *Continued*

Type of measurement (sample)	Usually attempts to characterize (whole)
6. Consumer Products	Concentration levels of various products.
7. Breathing Zone Measurements	Exposure to airborne chemicals.
8. Microenvironmental Studies	Ambient medium in a defined area, e.g., kitchen, automobile interior, office setting, parking lot.
9. Surface Soil Sample	Degree of contamination of soil available for contact.
10. Soil Core	Soil including pollution available for ground-water contamination; can be an indication of quality and trends over time.
11. Fish Tissue	Degree of contamination of edible fish tissue.
B. For Use in Point-of-Contact Measurement	
1. Air Pump/ Particulates and Vapors	Exposure of an individual or population via the air medium.

Examples	Typical information needed to characterize exposure
Shelf surveys, e.g., solvent concentration in household cleaners.[f]	Establish use patterns and/or market share of particular products, individual exposure at various usage levels, extent of passive exposure.
Industrial hygiene studies, occupational surveys, indoor air studies.	Location, activities, and time spent at monitoring locations. Protective measures/avoidance.
Special studies of indoor air, house dust, contaminated surfaces, radon measurements, office building studies.	Activities of study populations relative to monitoring locations and time exposed.
Soil sampling at hazardous waste sites.	Fate of pollution on/in soil; activities of potentially exposed populations.
Soil sampling at hazardous waste sites.	Fate of substance in soil; speciation and bioavailability, contact and ingestion rates as a function of activity patterns and age.
National Shellfish Survey[g]	Relationship of samples to food supply for individuals or populations of interest; consumption habits; preparation habits.
TEAM study,[h] carbon monoxide,[i] Breathing zone sampling in industrial settings.	Direct measurement of individual exposure during time sampled, including relationships between individual and the population. Exposures, relationships between concentrations at times sampled and at other times and relationship between sampled individuals and other populations. To make these links, activities of the sampled individuals compared to populations characterized are needed in some detail.

continued

TABLE 3-2 *Continued*

Type of measurement (sample)	Usually attempts to characterize (whole)
2. Passive Vapor Sampling	Same as above
3. Split Sample Food/ Split Sample Drinking Water	Exposures of an individual or population via ingestion.
4. Skin Patch Samples	Dermal exposure of an individual or populations.
C. For Use in Exposure Estimation from Reconstructed Dose:	
1. Breath	Total internal dose for individuals or population (usually indicative of relatively recent exposures).
2. Blood	Total internal dose for individuals or population (may be indicative of either relatively recent exposures to fat-soluble organics or long-term body burden for metals).
3. Adipose	Total internal dose for individuals or population (usually indicative of long-term averages for fat-soluble organics).
4. Nails, Hair	Total internal dose for individuals or population (usually indicative of past internal exposure in weeks to months range; can sometimes be used to evaluate exposure patterns).
5. Urine	Total internal dose for individuals or population (usually indicative of elimination rates); time from exposure to appearance in urine may depend on chemical.

[a]To characterize dose, intake or uptake information is also needed. [b]EPA, 1985a.
[c]EPA, 1986a. [d]EPA, 1985a. [e]EPA, 1985b. [f]EPA, 1985c.
[g]EPA, 1986a. [h]EPA, 1987a. [i]EPA, 1987a. [j]EPA, 1987a.
[k]EPA, 1987b. [l]EPA. 1986b. [m]EPA. 1986c. [n]EPA, 1987c.

Source: US EPA, 1992.

Examples	Typical information needed to characterize exposure
Same as above	Same as above.
TEAM study[j]	Same as above.
Pesticide Applicator Survey[k]	Same as above; Skin penetration.
Measurement of volatile organic chemicals (VOCS), alcohol. (Usually limited to volatile compounds.)	Relationship between individuals and population; exposure history (i.e., steady-state or not), pharmacokinetics (chemical half-life), possible storage reservoirs within the body, relationship between breath content and body burden.
Lead studies, pesticides, heavy metals (usually best for soluble compounds, although blood lipid analysis may reveal lipophilic compounds.)	Same as above; relationship between blood content and body burden.
NHATSF,[l] dioxin studies, PCBs (usually limited to lipophilic compounds.)	Same as above; relationship between adipose content and body burden.
Heavy metal studies (usually limited to metals).	Same as above; relationship between nails, hair content and body burden.
Studies of tetrachloropethylene[m] and trichloroethylene.[n]	Same as above; relationship between urine content and body burden.

those subjects have been followed long enough to classify them as cases or controls, precision almost as great as in the full cohort study may be obtained. The case-cohort design can allow similar efficiency in exposure assessment, often at far less cost than a full-scale cohort study.

ISSUES IN EXPOSURE ASSESSMENT

RISK ASSESSMENT AND ENVIRONMENTAL EPIDEMIOLOGY

Better measures of exposure are clearly desirable in epidemiology. Exposure data exist on a continuum ranging from measures of emission, to measures of ambient concentration, to microenvironmental measures weighted by time spent in each environment, to personal monitoring, to measures of internal dose and biomarkers. Pirkle et al. (1995) provide a summary with examples of 6 different uses of biomarker data: to identify priority exposures, to evaluate effectiveness of risk-mitigation strategies, to identify at-risk subpopulations, to recognize trends, to establish reference ranges, and to provide integrated dose measurements. Several of these measures may be available to evaluate a given exposure. Activity patterns may also allow extrapolation from microenvironmental data to aggregate risk, though with some additional error. While time-activity studies can often model the mean dose delivered to a population with reasonable accuracy, it is more difficult to predict accurately the distribution, or even the variance, about that mean. Epidemiologic studies that associate the ambient pollutant directly with the outcome inherently incorporate the distribution of population activity patterns.

More-accurate exposure assessment can increase the power of an environmental-epidemiology study to find an association. However, increased precision is likely to be more costly, and better information for each subject may reduce the number of subjects who can be studied. This tradeoff between sample size and precision per subject means that statistical power will not necessarily increase monotonically with improvements in the accuracy of the exposure assessment. Spending additional resources to obtain better measurements of risk factors other than the pollutant of interest may also increase the power of the study by reducing misclassification of confounders and modifiers, but the additional costs are another set of tradeoffs. For any given budget, the optimal tradeoff is usually not knowable in advance, though studies of multifactorial outcomes and low relative risks almost always require large sample sizes. This need for a large sample puts a premium on using inexpensive methods, such as questionnaire data on activity patterns and other exposure-modifying behavior, rather than expensive but more-accurate methods.

MEASUREMENT ERROR

The term *measurement error* is different from *exposure misclassification* because the former implies a continuous variable, while the latter suggests a dichotomous one. Exposure to environmental toxicants is a continuous variable in the real world, and one of the most important improvements in exposure assessment is for studies to move from dichotomizing exposure into continuous, or at least multilevel, measurements.

For example, consider a study relating air pollution to respiratory illness. The outdoor ambient-air pollution concentration is available from community monitoring. The analysis of these data will seek to correlate variations in air pollution with variations in respiratory outcome. Variations in recorded levels of air pollution may be thought of as having the following components: measurement error associated with the monitoring instrument, variation in the amount of time individuals spent outdoors, geographic variation in the outdoor concentration of the pollutant in the vicinity of the monitor, variations in the indoor/outdoor ratio, and individual variations in delivered dose. Critical issues include the size of each error component and the cost of reducing each component.

MISCLASSIFICATION

One of the most important advantages of improved exposure assessment derives from its impact on misclassification. Small errors in exposure assignment may have dramatic results on estimation of effect. Because of the limited scope of exposure assessments in most environmental epidemiology, misclassification is likely to be a substantial problem. In general, the internal validity of an epidemiologic investigation can be reduced by misclassification of study subjects. Copeland et al. (1977) emphasize that bias from misclassification will be "a function of the sensitivity and specificity of the classification procedure, the disease frequency, and exposure frequency." And in all case-control studies the bias depends on whether misclassification is the same or different in cases and controls, that is, nondifferential or differential in cases and controls.

In general, nondifferential misclassification causes measures of effect to be biased toward the null value. Such misclassification produces an underestimate of the effect, whereas differential misclassification can result in bias either toward or away from the null value. Copeland et al. (1977) argue that classification errors cannot be ignored and that investigators should attempt to estimate the magnitude of the errors. Dosemici et al. (1990) have shown that the predominant view that "nondifferential misclassification of exposure can only bias an estimate of a true positive odds ratio downward and not away from or beyond the null value" may

sometimes be wrong. These authors acknowledge that the problem they identify is not common, but they do suggest that caution is needed in interpreting results in the presence of nondifferential misclassification. (See Sexton et al., 1995, for a discussion of problems of misclassification associated with exposure measurement.)

THE NEED FOR IMPROVEMENT IN EXPOSURE ASSESSMENT

Better measures of exposure can improve the ability of a study to assess adverse effects from environmental agents. Such improvements lead to an increase in the power of the study and reduction in bias, but also to increased cost. The health outcomes and exposure analysis must be considered together to arrive at a balanced prioritization of study requirements.

A wide array of exposure-assessment tools is available to the epidemiologic investigator, ranging from personal monitoring to the use of diaries or other indirect means. All the tools have potential value when used logically and reasonably. A continuing dialogue is necessary between scientists whose emphasis is on exposure assessment and epidemiologic investigators. In this regard, Lioy (1991a) has stressed the need for continued dialogue to focus on critical questions that will reduce ambiguity in terminology and conceptual design and improve the experimental design of both health and exposure studies. The development of the EPA guidelines on exposure assessment (EPA, 1992) and NHEXAS (Burke et al., 1992) should stimulate further discussion on these issues.

Landrigan (1983) has commented on the advantages of improved exposure characterization by using individual versus grouped data in a study of the health effects of arsenic in drinking water. The average concentration of arsenic in well water was a poor indicator of individual exposure because some of the persons studied had supplemented their consumption or changed completely to drinking bottled water. When estimates of bottled water consumption were incorporated into individual-exposure assessment, the dose-effect relationship was strengthened.

Kennedy et al. (1991), in a cross-sectional study of pulp and paper workers exposed to chlorine gas, found no differences from workers in other industries. However, when those pulp and paper workers who had an acute gas exposure were considered, symptom and forced-expiratory-volume (FEV_1 and FEV_{25-75}) differences were found. The authors concluded that accidental chlorine or chlorine dioxide exposures in pulp mills are associated with increased respiratory symptoms and airflow obstruction, particularly among nonsmokers and former smokers.

Monster and Smolders (1984) studied teachers and pupils at a kindergarten near a factory with emissions of tetrachloroethane. The levels of

tetrachloroethane in the air exhaled by teachers and school-children were significantly greater than in a control group. The study demonstrated the applicability of biologic monitoring, where the exposure derives from the environment instead of the workplace.

Levine et al. (1985) monitored air at a hazardous-waste remedial-action site. This study demonstrated that, when workers remained in fixed job locations, "occupational inhalation exposure monitoring must consider contaminants generated upwind of the job location." Biologic monitoring would have been useful in this population of workers.

The selection of individual monitoring versus area or population monitoring of exposure is a matter of continuing concern. In particular, overreliance on "central-site monitoring" for assessment of exposure to air pollutants may be undesirable because of the possibility that personal exposures may differ from those estimated by central-site monitoring. Unrecognized variation among individuals in true air-pollution exposure contributes to variability among individuals in estimated slopes. While important findings have derived from air-pollution studies that make use of ambient-air monitoring data, investigations of more-precise questions such as determinants of variations in response may be improved through the use of personal monitors that are now becoming available.

The improved accuracy of personal monitoring of exposure generally comes at the expense of a substantial increase in cost and may therefore have its greatest value in the validation of other, less-expensive models of exposure. Freeman et al. (1991) have developed a location and activity log for assessing personal exposure to air pollutants and conducted pilot studies to validate the approach. Schwab et al. (1991) have suggested that the use of self-reported exertion levels in time-activity diaries has useful application to exposure assessment, particularly with respect to estimating the relation between exposure assessment and dose assessment. More studies of this nature are needed. The value of time-activity logs has been illustrated in detail by Sexton and Ryan (1989).

New instruments with the sensitivity and specificity necessary to conduct personal air-monitoring exposure assessments are also needed, and there have been a number of new developments of sampling instruments. Many of these instruments appear to be reasonable in cost (NRC, 1991b).

Hasabelnaby et al. (1989) used microenvironmental monitoring to characterize the exposure of preadolescent children to fine-particle air pollution. The authors concluded that microenvironmental monitoring is useful for estimating personal exposure in preadolescent children because they generally do not smoke and are not exposed occupationally. They also spend statistically significant periods at home, in school, or outdoors near their homes. These authors used data from the Six Cities Study to obtain more-accurate measures of exposure to passive smoking. They

included detailed information about a range of possible sources of exposures to environmental tobacco smoke. With the refined models developed in the study, a more-precise relation was identified between lung function and environmental tobacco smoke. This work illustrates the efficacy of using a population subset with improved exposure assessment in order to improve the sensitivity of the investigation.

Peters (1991) has conducted an epidemiologic investigation to identify the chronic effects of ambient air pollutants in southern California. The design for this 10-year study emphasizes exposure assessment. This emphasis is particularly warranted because of the magnitude of the health concerns and the potential cost of controls. The study is in 3 phases. Phase I considers the resources available to determine spatial and temporal patterns of pollutants, identify the locations that allow discrimination of pollutants, develop a sampling strategy, and determine cost-effective methods for estimating personal exposure. Phase II will consist of a cross-sectional study of health outcomes that can be related to personal exposure data. Five comparison groups are to be compared by their exposure to ozone, acid, and particles in varied exposure categories. Children will be the subject population. In phase III, the investigators will follow successive cohorts through either the duration of the study or high-school graduation. Changes in pulmonary function and incidence of disease will be compared with the exposures that have been assessed prospectively. This study will have a detailed exposure assessment for use in the evaluation of health effects associated with air-pollutant exposure. It will permit some estimation of the value of extensive exposure characterization in terms of both cost-resource use and the ability to define causal factors and dose-response relations in air-pollution studies.

Lioy et al. (1992) and Stern et al. (1992) have investigated population exposure to chromium waste, with particular reference to residential exposure. Lioy et al. sought to identify microenvironments that can lead to important chromium exposure. Chromium levels were determined for indoor air, outdoor air, and house dust. Surface dust was identified as the best index of potential chromium exposure. This study illustrates the role of total-exposure monitoring in the selection of exposure media for an epidemiologic study. Stern et al. (1992) compared chromium levels in urine with exposures estimated by environmental monitoring. The authors found an association between chromium in household dust and urine levels of chromium that was consistent with residential exposure to chromate-production waste. These 2 papers are excellent models of a systematic approach to exposure assessment.

Investigators from the National Cancer Institute have investigated methodologic issues in exposure assessment for case-control studies of cancer and herbicides-pesticides (Blair and Zahm, 1990a,b, 1992; Brown et

al., 1991). These authors conclude that improvements in exposure assessment are necessary if epidemiologic investigations are to provide reliable information on the relations between cancer incidence and pesticide exposure. These studies are noteworthy in their attempt to identify problems of reliability and validity of exposure assessments in case-control studies of cancer and pesticide exposure. For example, Brown et al. (1991) compared interview data from farmers with data from their wives or other surrogates and found excellent agreement between direct and surrogate interviews regarding the use of specific pesticides.

Blair and Zahm (1990a) concluded their study by stating, "Exposure misclassification undoubtedly occurs. Most errors from misclassification, however, are likely to be nondirectional in nature and would bias risk estimates toward the null and dilute exposure-response relationships. Methodologic investigations are needed to evaluate the reliability of current exposure assessment procedures and to develop new resources to improve assessment techniques." These recommendations are valuable, but it seems more likely that new resources are needed to enhance the best use of existing but underutilized techniques. These recommendations further illustrate the need for greater interaction between epidemiologists who recognize the importance of exposure assessment and investigators whose area of emphasis is exposure determination.

AIR-POLLUTION STUDIES AND EXPOSURE ASSESSMENT

For more than 4 decades, researchers have studied the impact of air pollution, most notably respirable particles and other priority pollutants, on human health. Carefully designed studies have provided a wealth of important information about relations between ambient exposures and adverse health consequences. A brief review of some of these studies illustrates the importance of microenvironmental monitoring and indirect measures of exposure characterization.

Studies by Abbey et al. (1991), Euler et al. (1987, 1988), and Mills et al. (1991) illustrate the benefit of ambient-air monitoring in epidemiologic characterization of health-related effects of air pollution. Rather than relying on the mean concentration of ambient air pollutants only, these authors evaluated the numbers of hours that ambient air pollutants exceeded thresholds. For each participant in the studies, the levels of air pollutants at ZIP-code centroids were estimated for each month of residence (Abbey et al., 1989).

Neas et al. (1991) related annual average nitrogen dioxide concentration, measured for indoor pollution as a continuous variable and as 4 ordered categories, to respiratory symptoms and pulmonary function in children. A 15-ppb increase in the mean concentration was associated

with an increase in lower respiratory symptoms. These authors discuss the importance of misclassification of exposure in diluting the findings of the effects of nitrogen dioxide exposure.

In a large-scale analysis of ambient daily air pollution and mortality, Schwartz and Marcus (1990) found a strong relation between air-pollution levels and daily mortality over the period 1958-1972 in London, England. Inclusion of terms for changes in temperature and humidity increased the strength of the relation between particles and mortality. The inclusion of weather data in the analysis illustrates how indirect measures of exposure can strengthen existing associations.

Information derived from daily health diaries can be particularly important. For example, diaries can capture variability in the occurrence of symptoms, which can be related to variation in air-pollutant concentrations. Schwartz et al. (1991) described the advantages of linking ambient-air measurements and individual indicators of household exposure to daily diaries of respiratory symptoms. Diary studies record the health status of each study participant repeatedly over time and can define the impact of short-term changes in the environment on human health. In this analysis of data from the Six Cities Study of Air Pollution and Health and from a nurses' diary study in Los Angeles, exposure information was obtained from air-monitoring stations as well as from data on the type of cooking stove and history of parental and roommate smoking. A statistically significant relation between air pollution and reported symptoms remained after the effects of autocorrelation and heterogeneity were addressed. The reader is referred to volume 1 of the present report (NRC, 1991a, pp. 166-167) and the report *Human Exposure Assessment for Airborne Pollutants* (NRC, 1991b, pp. 157-159) for a more detailed discussion of the use of diaries.

The Health Interview Survey (HIS) of a sample of thousands of US residents was used by Ostro and Rothschild (1989) to assess the relation between acute respiratory morbidity and air pollution. The large number of subjects in this study aided in the identification of statistically significant associations. The exposure information was derived from EPA's SAROAD monitoring network of ambient-air pollution. Portnoy and Mullahy (1986) have also used the HIS data to advantage. In each study, statistically significant findings were documented despite the limited exposure assessment from ambient monitoring.

Some investigators (Bates and Sizto, 1987; Bates et al., 1990; Samet et al., 1981) have reported a relation between visits to hospital emergency departments for respiratory ailments and various types of air pollution. The statistically significant relation between ozone concentration and asthma visits was identified by a multiple regression analysis that controlled for temperature. These authors conclude that, among regions with

periodic accumulation of ozone in the ambient environment, an exposure-response relationship may be discernible.

EXPOSURE ASSESSMENT AT HAZARDOUS-WASTE SITES

Volume 1 of this report (NRC, 1991a) reviewed the exposure assessments conducted in environmental-epidemiologic investigations of hazardous wastes. The report concluded that repositories of potentially dangerous substances can be found at many hazardous-waste sites but that future risks to public health could be determined only with more-detailed information about human exposure. The report concluded that exposure assessment must be improved if we are to understand the associations among contaminants, exposures, and adverse health consequences.

One of the problems in epidemiologic studies of hazardous-waste sites has been the limited nature of the exposure assessments (NRC, 1991a). Most studies have dichotomized subjects into exposed and unexposed categories or used surrogates of exposure, such as distance from a site or residence in a defined geographic area. Only a few have estimated quantitative exposure of groups, and we know of none that have used individual exposure measurements. Problems in the accuracy of exposure data have been discussed in detail (Walter, 1991; Morgenstern, 1982; Greenland and Morgenstern, 1989). Clearly, there can be serious errors in making the assumption that the exposure level assigned to a geographic subunit applies to everyone in that unit.

Investigations of hazardous-waste sites have generally focused on which chemicals are present at or under the site (NRC, 1991a), especially the possible contamination of groundwater, and have given little attention to evaluating chemical movement from the site. Thus, information is skimpy on exposure to the population in homes or businesses near the site in question. The routine monitoring of groundwater for purposes of site remediation and problem mitigation entails significant costs (e.g., hydrogeologic characterization, boring of wells, and soil and water sampling). Government agencies have rarely assessed population exposure to chemicals from the sites.

ASSESSMENT OF PAST EXPOSURE

In environmental epidemiology, information on past exposures is usually not available, and current exposure may not fairly represent the past because of technologic developments, public or government awareness of a possible problem, or other changes. Because recall bias or other types of information bias may be important in these studies, documented information on past exposure is particularly useful. In some studies,

water-consumption records have been used to estimate past exposures (Lagakos et al., 1986; Wrensch et al., 1990 a,b; Whorton et al., 1988). Mathematical modeling, toxicokinetic models, and biomarkers may improve our ability to estimate past exposures, especially where the body burden of xenobiotic chemicals is related to toxic insult.

In occupational epidemiology, it is sometimes possible to reconstruct an industrial environment for exposure estimation; exposure-monitoring records of workers may be available, and even employment records by job category may be of value. (See *Applied Occupational and Environmental Hygiene*, June 1991). Checkoway and colleagues (1991) have discussed methods to assess or rectify misclassification of historical exposures in occupational epidemiology (table 3-3). This work examines some indirect methods for determining possible bias in estimates of the health effects that result from nondifferential misclassification of exposure and considers misclassification of confounders. Checkoway et al. (1991) suggest that such direct approaches as simulating past exposures are often infeasible. They prefer indirect approaches for evaluating the effects of exposure and confounder misclassification.

COMPLEX MIXTURES

Waste sites may expose persons to multiple chemicals (NRC, 1991a). Three important issues arise in studying exposures to complex mixtures in epidemiologic investigations: (1) quantification of exposure, (2) characterization of potential combined or interactive effects associated with exposure to multiple chemicals, and (3) identification of subpopulations that are especially sensitive to exposure from certain complex mixtures. These problems are multiplied where the mixture may vary from one site or time to another or when the mixture is not well characterized. Both problems are common in waste sites. An entire issue of the journal *Toxicology* (volume 105, 1995) is devoted to an examination of chemical mixtures.

Workplace studies often focus on exposure to a single chemical agent, e.g., lead in battery manufacturing or silica in a foundry. However, for the most part, epidemiologic studies of the environment must address the issue of complex mixtures. NRC (1988) has reviewed the issues and approaches to assessing the health impact of complex mixtures. One approach involves characterization of complex mixtures by toxicologic investigation. To date, toxicologic characterization of complex mixtures has received inadequate attention as a complement to environmental investigation. The advantage of toxicologic investigation is that exposure can be set by the investigator (Ozonoff and Wartenberg, 1991). Toxicologic studies are needed to characterize both mechanisms and interactive effects and to quantitate the exposure-effect relations of complex mixtures.

Studies of exposure to the complex mixtures found in indoor air have implications for other aspects of environmental epidemiology. Samet and Lambert (1991) have pointed out that a full understanding of the health effects of indoor air pollution will require information on the effects of pollutant mixtures: "Studies of complex mixtures need to be designed with consideration of the potential patterns of combined effects of the combined pollutants. The biological effect of one pollutant may be modified by the presence of other pollutants."

Hammond (1991) and NRC (1988) have discussed the use of markers to measure exposure to complex mixtures. The study of interactive effects in environmental settings will be assisted by toxicologic evaluation and by the use of biomarkers of exposure and effect.

INDEXES OF EXPOSURE

Volume 1 of this report (NRC, 1991a) discussed the concept of "dose" in the context of exposure assessment. In general, inadequate attention has been given to evaluation of the appropriate time relations between biologic models derived from mechanistic considerations and the exposure-assessment strategy. The appropriateness of determining the time course of exposure (e.g., peak exposure versus cumulative exposure) is an important methodologic issue that has received some attention in occupational epidemiology (Checkoway and Rice, 1992; Wegman et al., 1992). However, there has been little effort in environmental epidemiology to relate the levels and effects of exposure to biologic mechanisms. This may be because information on biologic mechanisms is limited or absent. It can force the analyst into ad hoc approaches to exposure assessment.

Checkoway and Rice (1992) reviewed indexes of exposure in occupational epidemiology and characterized dose surrogates by exposure intensity, exposure duration, and cumulative exposure. The appropriateness of any of these measures as dose indicators depends on the pathogenesis of the disease under consideration. Acute health outcomes have generally been associated with peak exposure intensity, whereas cumulative exposure has often been used to address chronic disease. Research using toxicokinetic models has demonstrated nonlinear relations between exposure and effect because of nonlinearities in the biologic processes (Hattis, 1990; Smith, 1992). Smith (1992) drew several conclusions about the use of cumulative exposure as a surrogate of dose for insoluble respirable dusts: (1) short, intense exposures can produce substantially higher long-term tissue doses than is implied by their cumulative exposures, and this may partially account for the observation of a disproportionately high pulmonary-disease risk for short-term workers; (2) populations with widely ranging combinations of intensity and duration may

TABLE 3-3 Crude Odds Ratios For Symptom Incidence: Exposed Compared with Control Populations

Symptom	McColl(1)	OII[a](2)	Del Amo(3)	Montrose (3)	Springfellow(5)	Purity(4)
Nervousness	1.6–5.5*	–	–	–	–	–
Headache	1.6–7.1*	1.8–4.6[b]	2.2[b]	1.2	1.1, 1.2	1.05
Sleeplessness	1.7–7.9*	1.9–5.3[b]	2.2[b]	1.1	–	–
Fatigue	2.6–7.0*	1.8–3.1[b]	3.0[b]	1.2	1.3, 1.2	–
Dizziness	2.4–8.0*	1.0–2.2[b]	3.3[b]	1.5	1.7, 1.4	–
Nausea	2.1–24.5*	2.1[b]–3.9[b]	2.9[b]	1.6	2.2[b], 1.9[b]	–
Loss of appetite	1.5–17.3*	2.2[a]–5.1[b]	1.5	1.0	–	–
Stomachache	1.1–10.2*	–	–	–	–	–
Sinus congestion	1.4–4.4*	1.4–2.7[b]	3.3[b]	2.1[b]	1.1, 1.1	–
Blurred vision	–	–	–	–	1.6, 2.2[b]	–
Eye irritation	1.6–4.8*	1.4–3.7[b]	3.3[b]	1.8[b]	1.2, 1.3	–
Nose irritation	2.0–7.5*	–	–	–	–	–
Runny nose	2.8–5.6	–	–	–	–	–
Sore throat	1.8–5.9*	1.7–2.9[b]	3.5[b]	2.1[b]	–	–
Cough	1.6–4.0*	–	–	–	1.2, 2.1[b]	–
Asthma	–	1.3–2.8[b]	1.9	1.8	–	–
Allergies	1.4–4.2*	–	1.9	1.8	–	–
Wheezing	2.8–15.5*	–	–	–	1.2, 0.9	–
Skin irritation	1.1–5.0*	2.2[b]–3.1[b]	3.4[b]	2.3[b]	–	1.7

Chest pains	1.7–4.4*	–	–	–	1.2, 1.3	–
Earaches	1.4–3.8*	1.5–3.1[b]	3.5[b]	1.6	1.6, 2.2	–
Frequent urination	–	–	–	–	1.7, 1.7[b]	–
Difficulty breathing	–	1.7–3.3[b]	1.7[b]	1.2	–	0.8
Toothache	–	1.5–2.3[b]	2.3[b]	1.4	–	1.0
Muscle aches	–	2.1–3.7[b]	1.9[b]	1.4	–	1.1
Weak in extremities	–	–	–	–	1.9[b], 1.5	–
Numbness in limbs	–	1.3–2.9[b]	2.2[b]	1.4	1.8, 1.1	–
High environmental worry	9%	32%	18%	18%	–	–
Worry followed illness[c]	0.5%	7%	8%	8%	–	–
Number in control area	354	928	212	194	203	1,801
Total number in "exposed" area	703	1,349	444	430	402	157

[a]OII, Operating Industries, Inc.
[b]Lower 95% confidence limit was > 1.
[c]Worry arose because of illness.
*χ^2 trend $r < 0.05$.

(1) Satin et al., 1983.
(2) Satin et al., 1986.
(3) Satin et al., 1987.
(4) Smith and Rigau, 1988.
(5) Baker and Greeland, 1986.

Source: Neutra et al., 1991.

show elevated overall risk, but no dose-response relation with cumulative exposure, because of misclassification in the high-dose categories; and (3) populations that have many subjects with low-intensity exposures may show an increased risk only in the highest group because there is no overloading of the clearance process at low exposure levels.

Smith (1992) argues that the exposure-dose relation should be examined for nonlinearity before cumulative exposure is used as a dose index. This matter has generally not been addressed in environmental studies. For example, where bioactivation of reactive molecules produces toxic metabolites, effects may show poor correlations with cumulative exposure and strong correlations with years of exposure, because metabolism is generally saturated and metabolite output is independent of exposure intensity.

Checkoway and Rice (1992) considered cumulative exposure to silica and concluded that "disentanglement of peak and cumulative exposure effects should be accomplished most effectively in investigations of relatively short-term sequelae of substances with short retention times in the body."

SUBJECTIVE SYMPTOMS AND EXPOSURE ASSESSMENT

An important issue in epidemiologic studies of hazardous-waste sites is the relation between positive findings as evidenced by increased subjective symptoms in a defined population and the general lack of exposure assessment that would link the findings to specific exposures. Some authors have suggested that the increased prevalence of subjective symptoms may derive from recall bias or psychogenic factors, whereas others have argued that symptoms may result from exposures of particularly susceptible members of a population. In volume 1, it was suggested that longitudinal studies of symptoms in response to changes in exposure could help to resolve this problem, if subjects have no knowledge of their level of exposure that could bias reporting.

Several studies have found that subjective symptoms are associated with exposures from hazardous-waste sites (Baker et al., 1988; Neutra et al., 1991). Issues of cumulative versus peak exposure appear to be relevant in these studies. In general, investigation of the averaging time of exposure indexes should be a priority in environmental epidemiology. Unfortunately, when environmental epidemiology is driven by citizen-generated concern, the data available to the investigator may be so limited that precise questions cannot be addressed. The actual chemicals to which a community population is exposed are often poorly or inadequately identified, complex mixtures may be the rule rather than the exception, and exposure routes are not well defined, so the ability to ad-

dress such complex issues as dose rate is seriously compromised (NRC, 1991a).

Neutra et al. (1991) examined the results of 5 epidemiologic studies of symptom rates observed around hazardous-waste sites. Table 3-3 lists the 5 studies and the odds ratios for symptoms with increased prevalence that reached statistical significance in these studies. The health complaints were often subjective, and the circumstances leading to the studies resulted in intense media scrutiny and even litigation. Table 3-4 lists 8 hypotheses that have been suggested to explain the higher prevalence of symptoms in exposed persons, including reporting bias. Neutra et al. (1991) conclude that excess symptoms are reported by residents who complain of odors or are worried about environmental chemicals, and they suggest the possibility that "autonomic, stress-mediated mechanics or behavioral sensitization is active in the genesis of these symptoms."

One report on the Stringfellow hazardous-waste site (Baker et al., 1988) suggested caution in the interpretation of the data: "Our experience indicates the fundamental need for health studies of toxic waste disposal sites to be based on environmental monitoring and modeling of past exposures sufficient to identify potential exposure to specific chemicals at an individual or household level."

The study by Neutra et al. (1991) is one in a series that have examined the health effects associated with hazardous-waste disposal sites and the relation between odor and increased symptoms. A study at the Casmalia hazardous-waste site in California indicated that respiratory effects were associated with airborne releases of odorous materials from the site (Breslow et al., 1989). A report from Finland (Jaakkola et al., 1990) indicated that malodorous emissions from kraft paper pulp mills are associated with eye, nasal, and respiratory symptoms, although reporting bias could not be ruled out (NRC, 1991a). However, a study from Lowell,

TABLE 3-4 Suggested Causes of Higher Symptom Rates Near Hazardous-Waste Sites

Classical toxic reaction
Immunologic or other physiogenic "hazardous-waste syndrome"
Behavioral sensitization
Psychosomatic reaction to stress
Mass psychogenic illness
Reporting bias
Confounding factors
Odor, as an effect modifier

Source: Neutra et al., 1991.

Mass., that reported an increased prevalence of respiratory and constitutional symptoms concluded that recall bias was not a factor (Ozonoff et al., 1987).

Neutra et al. (1991) proposed protocols to test hypotheses that symptoms are associated with low-level chemical exposures and that these symptoms are early warnings of serious immunologic and neurologic dysfunction. They suggest that these would be "million-dollar studies" because they test paradigm-breaking hypotheses, and they would need to be replicated to be believed by the community in question (if that were ever to occur). However, million of dollars in tort liabilities may be associated with these sites, and the issues require resolution.

These issues clearly require careful attention to exposure evaluation and characterization. The cautions raised by Baker et al. (1988) are entirely appropriate. Improved exposure assessment is an absolute requirement if these costly and scientifically important issues are to be resolved.

USE OF BIOLOGIC MARKERS OF EXPOSURE

The use of biologic markers has been reviewed in detail by NRC (1991a,b). Three types of biologic markers may be used to provide information on exposure (Hulka and Wilcosky, 1988; Schulte 1989; Stevens et al., 1991): markers of internal dose, such as blood lead; markers of biologically effective dose, such as blood DNA and protein adducts; and markers of biologic effects, such as chromosomal micronuclei.

The central, critical issue in the use of biologic markers of exposure, not yet adequately addressed, is criteria for their validation. NRC (1991b, p. 129) has reviewed the use of biologic markers and concluded: "A major limitation of using biological markers for exposure assessments stems from the fact that most are in a developmental stage and not fully validated or field-tested." NRC (1991b) also raised issues of the ambiguity of many markers, the variability of markers, and the difficulty of establishing links between exposure and effect. Criteria governing the validation and use of biologic markers were described in detail in that report.

DESIGNING AND CONDUCTING A STUDY WITH BIOLOGIC MARKERS

When environmental-monitoring data are available for classifying individuals, it may not be necessary to perform biologic monitoring. Biologic monitoring of a sample of subjects can be used to validate environmental-monitoring variables. If there is a close correspondence between exposure and the biologic marker, then further use of the more-costly biologic monitoring may not be needed and the exposure classification scheme may be considered validated. If the correspondence is not close,

it may be that exposure from diverse sources or by various rates is not covered by environmental monitoring. For example, classification of residents by distance from an arsenic smelter may not adequately reflect the arsenic concentrations in their diets, and environmental monitoring may not provide good estimates of total arsenic exposure.

Biologic markers are often assumed to be good indicators of exposure because they represent the integrated exposure from various sources and through various routes. However, to assess this assumption requires correlation of the marker with the potentially less-adequate environmental measure. There is no "gold standard." Perfect correspondence between the marker and the exposure could mean that neither is better than the other or that there are no other routes, sources, or host factors that intervene. On the other hand, it may mean that the marker is not an accurate reflection of these other intervening factors. It is important to determine whether the marker shows an exposure-response relation, whether all potential routes are accounted for, and whether susceptibility or host factors are addressed.

Host factors, including behavioral factors and genetic characteristics, may influence the amount of a toxic agent that interacts with critical macromolecules in cells and tissues. This is the "biologically effective dose." The biologically effective dose assesses exposure from all routes and sources as well as some aspects of effect modification, possibly including host characteristics for uptake, metabolism, absorption, and excretion. However, the marker may not necessarily encompass all these factors. Thus, even when biomarkers are useful, the best appraisal of exposure may still include ambient and environmental measurements as well as biologic measurements.

Numerous biomarker-related issues may arise during the conduct of studies, including questions of specimen collection, transport, storage, and assay; measurement error of technical variables in the assay; biologic variability; and assay interpretation and communication of results.

In cohort studies, biologic markers may be measured in subsets of populations, such as in a nested case-control or case-cohort approach, to assess etiologic questions and mechanisms and to identify high-risk subpopulations. In these situations, biologic markers of exposure may be useful to (1) distinguish exposure subgroups, (2) determine whether there is a relation between exposure and dose, or (3) evaluate the relation between exogenous exposure and internal or biologically effective dose.

Biologic markers may also be useful to identify the effect of an intervention. For example, does reduction of environmental emissions result in a reduction in the level of DNA adducts? Research studies to assess interventions need to include assessment of baseline levels of biomarkers in order to interpret the effect of the interventions.

INTERPRETATION AND GENERALIZATION OF STUDY RESULTS

Biologic markers of exposure can be of use after a study has been completed. For example, if researchers wish to see how well the results of a completed study might apply to a broad population, they may sample the population for the distribution of a particular marker to determine whether exposures are constant over a wide range of geographic conditions, demographic descriptors, and occupations. Even if the original study did not measure such factors, biologic markers may clarify what exposures such target groups may have experienced.

It may also be possible to perform individual risk assessments using biologic measures of exposure. A classic example of individual risk assessment is the use of serum-cholesterol measurement to predict disease risk (Truett et al., 1967). With more-recent technology, one might attempt precise individual risk assessment by studying an individual's specific spectrum of gene mutations from specific exposures to a carcinogen.

DATA GAPS, RESEARCH RECOMMENDATIONS, AND RESOURCE LIMITATIONS

Few biologic markers of exposure have been validated. As indicated above, validation of a marker of exposure requires an understanding of the dosimetric characteristics pertaining to the time between exposure and the ascertainment of the markers, the degree to which the marker represents exposure, and the nature and shape of the exposure-marker relation (Schulte 1989; Stevens et al., 1991). Little is known about the prevalence, range of variability, persistence, and confounding factors of many candidate exposure markers. Such information must be collected before these can be used with confidence in environmental-epidemiologic studies. Markers of the biologically effective dose require additional research to assess the role of host factors, particularly genetic susceptibilities, as effect modifiers.

DOSIMETRIC MODELING

More attention is being devoted to characterizing the quantity and timing of toxic chemical agents' reaching target tissue, because concentrations measured by microenvironmental monitoring or even personal dosimetry may not accurately reflect target-tissue dose. This has led to greater emphasis on mathematical models derived from biologic mechanisms of toxicity. This approach has historically been the focus of pharmacologists who have sought to develop appropriate models to characterize the relation between drug efficacy and dose. Mathematical

modeling of tissue dose, biologically effective dose, or internal dose is receiving greater attention in addressing toxicologic issues. However, there has been little attention to the relation between these models, hereafter referred to as dosimetric models, and classic epidemiologic models that estimate disease risk. Kriebel (1991) has discussed the importance of this approach to epidemiologic investigation: "Often it is difficult or impossible to accurately estimate the exposure experience of each member of a cohort, and so various kinds of proxy variables must be used. It is well known that the use of these proxies can introduce misclassification of exposure, often leading to underestimation of the magnitude of exposure-disease associations. Even when accurate exposure data are available, serious bias may still occur if these data are used in a mis-specified epidemiologic model to estimate an exposure-response relationship."

Kriebel discusses several tenets for the use of dosimetric models. These tenets result in a 2-phase approach to epidemiologic investigation: first, is the development of a mathematical model to estimate individual doses; and second, the use of epidemiologic models to estimate the risk of disease associated with these estimated doses, with appropriate control for confounding. The advantage of this approach for environmental epidemiology is that the dosimetric model can quantify a hypothesis about uptake processes and metabolism of the chemicals in question and may provide insight into the biologic mechanisms of effect. This facilitates the overall design of the epidemiologic model and subsequent analysis. That is, the dosimetric model informs the exposure assessor and epidemiologist and provides a way to reduce misclassification and improve the precision of the study. These concepts have been applied by Hattis (1990), Smith (1992), Hodgson and Jones (1990), Vineis and Terracini (1990), Vincent and Mark (1988), Pinto et al. (1978), and Kriebel and Smith (1990).

Smith (1985) developed a compartmental dosimetric model of dust deposition for an occupational-epidemiologic study of pulmonary fibrosis in silicon carbide workers. In discussing the Smith model, Kriebel (1991) asks whether the considerable effort that goes into the construction and use of such a model was justified. He raises 3 criteria that could be used to evaluate any particular model: (1) the better model will fit the data better, (2) the better model will accommodate such secondary characteristics of the exposure-response relation as interactions with other agents (effect modifiers) and time (such as latency), and (3) the predictions of the better model can be generalized to other exposure situations. Smith (1991) also discusses the use of toxicokinetic modeling for epidemiologic purposes and argues that the use of toxicokinetic models can differentiate among hypotheses about the mechanisms that underlie the relation between exposure and effects.

The greatest overall contribution made by a 2-stage approach to epi-

demiology (as described by Kriebel, 1991) is to provide a framework within which all available data, including toxicologic information and experimental data, can be viewed. In fact, animal data may be essential to the development of some models. Ultimately, dosimetric models are mathematical expressions of formal hypotheses about the underlying physiologic processes that are the basis of the exposure-response relation (Kriebel, 1991).

TRAINING IN ENVIRONMENTAL-EXPOSURE ASSESSMENT

Human exposure assessment is inadequately addressed in most environmental-epidemiology studies, and one of the roots of this problem is the lack of training at the graduate level. There is a major need, in the United States and elsewhere, for the development of training programs in exposure assessment. There are extensive master's-degree programs for industrial hygiene as a result of the National Institute for Occupational Safety and Health's Educational Resource Center program and other extramural project grants for training. There is at least one similar training program that has environmental-exposure assessment as its focus—in the Department of Environmental Science at Rutgers University (Lioy, 1991b). Exposure assessment is addressed in courses on environmental risk assessment, but even there the context is more focused on risk assessment and site remediation than on epidemiologic investigation of public-health hazards. Training in exposure assessment must be multidisciplinary, with a multimedia approach, and should address all the major uses of exposure information—including risk assessment, epidemiology, environmental control, and exposure assessment—and industrial hygiene, toxicology, pollution prevention, and standard-setting. It would be useful to examine the relations among needs for training in these areas to define a new curriculum that would better address current and future needs.

Given the costs, resource requirements, and political sensitivity of many environmental-epidemiologic studies, the failure to provide training for environmental assessment will need to be addressed by policymakers and educators if we are to have substantial improvement in environmental epidemiology and risk assessment.

CONCLUSIONS

Exposure assessment is important in all environmental-epidemiologic studies. A wide range of exposure-assessment strategies and techniques are available for use in environmental-epidemiologic investigation. Associations have been clarified by improved use of exposure assessment even where indirect methods have been used.

Both direct measures (personal and biologic monitoring and biomarkers) and indirect measures (microenvironmental monitoring, diaries, and mathematical modeling) can be used for exposure assessment in environmental epidemiology. Each of these techniques has advantages and disadvantages. Their optimal use depends on the nature of the study, the biologic hypothesis, and resource constraints. No approach should be singled out as being the only acceptable strategy, e.g., personal monitoring. All approaches have validity and will improve the study if used appropriately.

Better exposure assessment in environmental epidemiology will increase the power of studies to find associations. However, within a fixed budget, spending more money on exposure assessment per subject will reduce the number of subjects who can be studied and hence could reduce statistical power. The tradeoff between precision and the cost of larger samples means that power will not increase monotonically with improvements in the accuracy of the exposure assessment. In studies of multifactorial outcomes and low relative risks, a large sample is almost always required. This means that inexpensive methods for modest improvement of rough and inexpensive exposure assessment may be more valuable than more-accurate but expensive methods. This includes very inexpensive methods, such as the use of questionnaire data on activity patterns.

Studies of large populations exposed to mixtures of air pollutants should incorporate detailed estimates of exposure, including detailed activity logs (including transit to work or school), the kind of air conditioning in the home and workplace, and the use of personal monitors to validate models in subsets of the population under study.

The problems of exposure measurement in persons living close to hazardous-waste sites were discussed in volume 1. Most studies have been structured around an "exposed-unexposed" classification or have used surrogates of exposure, such as distance from the waste site. Estimation of past exposures is particularly difficult and unreliable. Misclassification is likely to be a crucial problem in studies of this nature, and improved characterization of exposure is a priority.

The estimation of cumulative doses is an important component of many occupational-exposure studies, though such measures may not be valid even in occupational settings. The exposure-dose relation should be examined for nonlinearity before cumulative estimates are calculated. The relation between cumulative exposure and peak exposure is unknown in many environmental-epidemiologic studies, particularly those involving hazardous-waste site exposures or community exposures to episodic pollution.

The characterization of complex mixtures is a continuing problem for

exposure assessment. Four priorities in addressing complex mixtures are quantification of exposure to complex mixtures, characterization of combined or interactive effects, toxicologic characterization of the complex mixture in question, and identification of subpopulations that may be especially sensitive to one or more of the components of certain complex mixtures.

Biologic markers of exposure can strengthen environmental-epidemiologic studies. Unfortunately, few such markers are yet feasible in field studies, and few have been adequately validated. Efforts to improve and refine such indicators are important. The feasibility and value of banking blood samples for future analysis should be considered as studies are designed. Biomarkers of changes induced in the immune system of human subjects are needed.

Health effects are often subtle, and risks are difficult to estimate. As a result, more attention is being given to the estimation of target-tissue dose in ways that reduce misclassification and improve precision. Development of mathematical models to estimate target-tissue dose (toxicokinetic modeling) that may be combined with epidemiologic models to estimate risk is a new and important area of research.

Emphasis should be given to the development of training programs in environmental-exposure assessment. Improvement in the development and use of new techniques in exposure assessment is a high priority in environmental epidemiology.

Data should of course be generated and collected under rigorous conditions of quality control. Bias must be minimized, and variance must be both minimized and estimated when quantitative conclusions are to be drawn. Measures of and checks on data quality should be prominent in every manuscript and report, and authors must not be reticent in bringing out the mechanisms of their study—and there will always be weaknesses. Because of the difficulties of conducting epidemiologic studies, both descriptive and analytic, it is rare for any one study to be definitive, and this is especially true in environmental epidemiology. Every public presentation, written or oral, including reports to scientific colleagues, should contain prominent caveats about overinterpretation.

REFERENCES

Abbey, D.E., G.L. Euler, J.K. Moore, F. Petersen, J.H. Hodgkins, and A.R. Magie. 1989. Applications of a method for setting air quality standards based on epidemiological data. J. Air Pollut. Control Assoc. 39:437-445.

Abbey, D.E., P.K. Mills, F.F. Petersen, and W.L. Beeson. 1991. Long-term ambient concentrations of total suspended particulates and oxidants as related to incidence of chronic disease in California Seventh-Day Adventists. Environ. Health Perspect. 94:43-50.

ATSDR (Agency for Toxic Substances and Disease Registry). 1994. Environmental Data Needed for Public Health Assessments: A Guidance Manual. NTIS PB-179827. Atlanta: USDHHS, ATSDR, Division of Health Assessment and Consultation.

Bailar, J.C. 1989. Inhalation hazards: the interpretation of epidemiologic evidence. Pp. 39-48 in D. V. Bates et al., eds. Assessment of Inhalation Hazards, New York: Springer-Verlag.

Baker, D.B., S. Greenland, J. Mendlein, and P. Harmon. 1988. A health study of two communities near the Stringfellow waste disposal site. Arch. Environ. Health 43:325-334.

Bates, D.V., and R. Sizto. 1987. Hospital admissions and air pollution in southern Ontario: the acid summer haze effect. Environ. Res. 43:317-331.

Bates, D.V., M. Baker-Anderson, and R. Sizto. 1990. Asthma attack periodicity: a study of hospital emergency visits in Vancouver. Environ. Res. 51:51-70.

Blair, A., and S. Zahm. 1990a. Methodologic issues in exposure assessment for case-control studies of cancer and herbicides. Am. J. Ind. Med. 18:285-293.

Blair, A., and S. Zahm. 1990b. Herbicides and cancer: a review and discussion of methodologic issues. Pp. 132-145 in P. Band, ed. Recent Results in Clinical Cancer Research. Vol. 120. Occupational Epidemiology. New York: Springer-Verlag.

Blair, A., and S. Zahm. 1992. Epidemiology Studies Of Cancer Among Agricultural Populations. Presented at Third International Symposium: Issues in Health, Safety and Agriculture. Saskatoon, Saskatchewan, Canada, May 10-15, 1992.

Breslow, L., et al. 1989. Report of Santa Barbara Commission on Health Consequences of the Casmalia Resources Waste Disposal Facility. Report prepared for Santa Barbara Department of Health Services, California.

Brown, L.M., M. Dosemeci, A. Blair, and L. Burmeister. 1991. Comparability of data obtained from farmers and surrogate respondents on use of agricultural pesticides. Am. J. Epidemiol. 134:348-355.

Burke, T.A., and K. Sexton. 1995. Integrating science and policy in a national human exposure assessment survey. J. Exposure Anal. Environ. Epidemiol. 5:283-296.

Burke, T.A., J. Creason, P. Ryan, and M. Schwab. 1992. An Approach to the Design of the National Health Exposure Assessment Survey (NHEXAS): A Consolidated Report. Preliminary draft for NHEXAS Advisory Committee. 52 pp.

Checkoway, H., and C.H. Rice. 1992. Time-weighted averages, peaks, and other indices of exposure in occupational epidemiology. Am. J. Ind. Med. 21(1):25-33.

Checkoway, H., D.A. Savitz, and N.J. Heyer. 1991. Assessing the effects of nondifferential misclassification of exposures in occupational studies. App. Occup. Environ. Hyg. 6:528-533.

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

Dosemici, M., S. Wacholder, and J.H. Lubin. 1990. Does nondifferential misclassification of exposure always bias a true effect toward the null value? Am. J. Epidemiol. 132:746-748.

EPA (US Environmental Protection Agency). 1988. Future Risk: Research Strategies for the 1990s. SAB-EC-88-040. Washington, DC: US Environmental Protection Agency Science Advisory Board. 19 pp.

EPA (US Environmental Protection Agency). 1992. Guidelines for exposure assessment: notice. Federal Register, Part VI, Vol. 57, No. 104.

Euler, G.L., D.E. Abbey, J. E. Hodgkin, and A.R. Magie. 1987. Chronic obstructive pulmonary disease symptom effects of long-term cumulative exposure to ambient levels of total suspended particulates and sulfur dioxide in California Seventh-Day Adventist residents. Arch. Environ. Health 42:213-222.

Euler, G.L., D.E. Abbey, J.E. Hodgkin, and A.R. Magie. 1988. Chronic obstructive pulmonary disease symptom effects of long-term cumulative exposure to ambient levels of total oxidants and nitrogen dioxide in California Seventh-Day Adventist residents. Arch. Environ. Health 43:279-285.

Freeman, N.C., J.M. Waldman, and P.J. Lioy. 1991. Design and evaluation of a location and activity log used for assessing personal exposure to air pollutants. J. Expo. Anal. Environ. Epidemiol. 1:327-338.

Gann, P. 1986. Use and misuse of existing data bases in environmental epidemiology: the case of air pollution. Pp. 109-122 in F. Kopfler and G.F. Craun, eds. Environmental Epidemiology. Chelsea, MI: Lewis.

Greenland, S., and H. Morgenstern. 1989. Ecological bias, confounding, and effect modification. Int. J. Epidemiol. 18:269-274. [Published erratum appears in Int. J. Epidemiol. 1991 20:824.]

Hammond, S.K. 1991. The uses of markers to measure exposures to complex mixtures. Pp. 53-66 in L.M. Rappaport, and T. Smith, eds. Exposure Assessment for Epidemiology and Hazard Control. Chelsea, MI: Lewis.

Hasabelnaby, N.A., J.H. Ware, and W.A. Fuller. 1989. Indoor air pollution and pulmonary performance: investigating errors in exposure assessment. Stat. Med. 8:1109-1126; discussion 1137-1138.

Hattis, D. 1990. Pharmacokinetic principles for dose-rate extrapolation of carcinogenic risk from genetically active agents. Risk Anal. 10:306-316.

Hodgson, J.T., and R.D. Jones. 1990. Mortality of a cohort of tin miners 1941-86. Br. J. Ind. Med. 47:665-676. [Published erratum appears in Br. J. Ind. Med. 1990 47:846.]

Hulka, B.S., and T. Wilcosky. 1988. Biological markers in epidemiologic research. Arch. Environ. Health 43:83-89.

Jaakkola, J.K., V. Vilkka, O. Marttiila, P. Jappinen, and T. Haahtela. 1990. The South Karelia air pollution study: the effects of malodorous sulfur compounds from pulp mills on respiratory and other symptoms. Am. Rev. Respir. Dis. 142:1344-1350.

Kennedy, S.M., D.A. Enarson, R.G. Janssen, and M. Chan-Yeung. 1991. Lung health consequences of reported accidental chlorine gas exposure among pulpmill workers. Am. Rev. Respir. Dis. 143(1):74-79.

Kriebel, D. 1991. The Dosimetric Model in Epidemiology. Invited paper presented at 8th International Symposium: Epidemiology in Occupational Health, September 10-12, Paris, France. Organized under auspices of International Commission on Occupational Health, Scientific Committee on Epidemiology. 21 pp. Proceedings in press.

Kriebel, D., and T.J. Smith. 1990. A nonlinear pharmacologic model of the acute effect of ozone on the human lungs. Environ. Res. 51:120-146.

Lagakos, S.W., B.J. Wessen, and M. Zelen. 1986. An analysis of contaminated well water and health effects in Woburn, Massachusetts. J. Am. Stats. Assoc. 81:583-596.

Landrigan, P.J. 1983. Epidemiologic approaches to persons with exposures to waste chemicals. Environ. Health Perspect. 48:93-97.

Levine, S.P., R. J. Costillo, C.L. Geraci, and K.A. Conlin. 1985. Air monitoring at the drum bulking process of a hazardous waste remedial action site. Am. Ind. Hyg. Assoc. J. 46:192-196.

Lioy, P.J. 1990. Assessing total human exposure to contaminants: a multidisciplinary approach. Environ. Sci. Technol. 24:938-945.

Lioy, P.J. 1991a. Arguments on the Adequacy of Exposure Indices and Convergence with Health Indices of the Adverse Effects of Air Pollution. Paper given at the 91st Annual AWMA Meeting, Vancouver, British Columbia.

Lioy, P.J. 1991b. Human exposure assessment: a graduate level course. J. Expo. Anal. Care Environ. Epidemiol. 1:271-281.

Lioy, P.J., N.C.G. Freeman, I. Walnman, A.H. Stern, R. Boesch, T. Howell, and S.I. Shupack. 1992. Microenvironmental analysis of residential exposure to chromium laden wastes in and around New Jersey homes. Risk Anal. 12:287-299. [Published erratum appears in Risk Anal. 1992. 12:463.]

Mills, P., D. Abbey, W. L. Beeson, and F. Petersen. 1991. Ambient air pollution and cancer in California Seventh-Day Adventists. Arch. Environ. Health 46:271-280.

Monster, A.C., and J.F.F. Smolders. 1984. Tetrachloroethene in exhaled air of persons living near pollution sources. Int. Arch. Occup. Environ. Health 53:331-336.

Morgenstern, H. 1982. Uses of ecologic analysis in epidemiologic research. Am. J. Pub. Health 72:1336-1344.

Neas, L.M., D.W. Dockery, J.H. Ware, J.D. Spengler, F.E. Speizer, and B.G. Ferris, Jr. 1991. Association of indoor nitrogen dioxide with respiratory symptoms and pulmonary function in children. Am. J. Epidemiol. 134:204-219.

Neutra, R., J. Lipscomb, K. Satin, and D. Shusterman. 1991. Hypotheses to explain the higher symptom rates observed around hazardous waste sites. Environ. Health Perspect. 94:31-38.

NRC (National Research Council). 1988. Complex Mixtures: Methods for In Vivo Toxicity Testing. Washington, DC: National Academy Press. 227 pp.

NRC (National Research Council). 1991a. Environmental Epidemiology, Vol. 1. Public Health and Hazardous Wastes. Washington, DC: National Academy Press.

NRC (National Research Council). 1991b. Human Exposure Assessment for Airborne Pollutants. Washington, DC: National Academy Press.

NRC (National Research Council). 1991c. Monitoring Human Tissues for Toxic Substances. Washington, DC: National Academy Press. 211 pp.

Ostro, B., and S. Rothschild. 1989. Air pollution and acute respiratory morbidity: an observational study of multiple pollutants. Environ. Res. 50:238-247.

Ozonoff, D., and D. Wartenberg. 1991. Toxic exposures in a community setting: the epidemiologic approach. Pp. 77-88 in Molecular Dosimetry and Human Cancer. Boston: CRC Press.

Ozonoff, D., M.E. Colten, A. Cupples, T. Heeren, A. Sachatzkin, T. Mangione, M. Dresner, and T. Colton. 1987. Health problems reported by residents of a neighborhood contaminated by a hazardous waste facility. Am. J. Ind. Med. 11:581-597.

Peters, J. 1991. Epidemiologic Investigation to Identify Chronic Effects of Ambient Air Pollution in Southern California. Proposal to California Air Resources Board.

Pinto, S.S., V. Henderson, and P.E. Enterline. 1978. Mortality experiences of arsenic-exposed workers. Arch. Environ. Health 33:325-331.

Pirkle, J. L., L. L. Needham, and K. Sexton. 1995. Improving exposure assessment by monitoring human tissues for toxic chemicals. J. Exposure Anal. Environ. Epi. 5:405-424

Portnoy, P., and J. Mullahy. 1986. Urban air quality and acute respiratory illness. J. Urban Economics 20:21-28.

Ryan, P. 1991. An overview of human exposure modeling. J. Exp. Anal. Environ. Epidemiol. 1:453-474.

Samet, J.M., and W.E. Lambert. 1991. Epidemiologic approaches for assessing health risks from complex mixtures in indoor air. Environ. Health Perspect. 95:71-74.

Samet, J.M., F.E. Speizer, Y. Bishop, J.D. Spengler, and B.G. Ferris, Jr. 1981. The relationship between air pollution and emergency room visits in an industrial community. J. Air Pollution Control Assoc. 31:236-240.

Schulte, P. 1989. A conceptual framework for the validation and use of biological markers. Environ. Res. 48:129-144.

Schwab, M., A.P. Treblanche, and J.D. Spengler. 1991. Self-reported exertion levels on time/activity diaries: application to exposure assessment. J. Expo. Care Anal. Environ. Epidemiol. 1:339-356.

Schwartz, J., and A. Marcus. 1990. Mortality and air pollution in London: a time series analysis. Am. J. Epidemiol. 131:185-194.

Schwartz, J., D. Wypij, D. Dockery, J. Ware, S. Zeger, J. Spengler, and B. Ferris, Jr. 1991. Daily diaries of respiratory symptoms and air pollution: methodological issues and results. Environ. Health Perspect. 90:181-187.

Sexton, K. 1991. National Human Exposure Assessment Survey. EPA (Environmental Protection Agency). Position paper. Washington, DC: Office of Health Research.

Sexton, K., and P.B. Ryan. 1989. Human Exposure to Air Pollution: Methods, Measurements, and Models. In: A. Watson, RR Bates, and D. Kennedy, eds. Air Pollution, the Automobile, and Public Health. Washington, DC: National Academy Press.

Sexton, K., M. A. Callahan and E. F. Bryan. 1995. Estimating exposure and dose to characterize health risks: the role of human tissue monitoring in exposure assessment. Environ. Health Perspec. 103 (Suppl.): 13-29.

Smith, T.J. 1985. Development and application of a model for estimating alveolar and interstitial dust levels. Ann. Occup. Hyg. 29:495-516.

Smith, T.J. 1991. Pharmacokinetic models in the development of exposure indicators in epidemiology. Ann. Occup. Hyg. 35:543-560.

Smith, T.J. 1992. Occupational exposure and dose over time: limitations of cumulative exposure. Am. J. Ind. Med. 21:35-51.

Stern, A.H., N.C.G. Freeman, P. Pleban, R.R. Boesch, T. Wainman, T. Howell, S.I. Shupack, B.B. Johnson, and P.J. Lioy. 1992. Residential exposure to chromium waste-urine biological monitoring in conjunction with environmental exposure monitoring. Environ. Res. 58:147-162.

Stevens, D.K., R.J. Bull, C.H. Nauman, and J.N. Blancato. 1991. Decision model for biomarkers of exposure. Regul. Toxicol. Pharmacol. 14:286-296.

Truett, J., J. Cornfield, and W. Kannel. 1967. A multivariate analysis of the risk of coronary heart disease in Framingham. J. Chronic Dis. 20:511-524.

Vincent, J.H., and D. Mark. 1988. The Measurement of Aerosols in Relation to Risk Assessment. International Workshop on Exposure Assessment for Epidemiology and Hazard Control, Woods Hole, Massachusetts, 1988.

Vineis, P., and B. Terracini. 1990. Biochemical epidemiology of bladder cancer. Epidemiology 1:448-452.

Wallace, L.A. 1991. Personal exposure to 25 volatile organic compounds. EPA's 1987 team study in Los Angeles, California. Toxicol. Ind. Health. 7:203-208.

Wallace, L.A., E.D. Pellizzari, T.D. Hartwell, R. Whitmore, C. Sparacino, and H. Zelon. 1986. Total Exposure Assessment Methodology (TEAM) study: personal exposures, indoor-outdoor relationships, and breath levels of volatile organic compounds in New Jersey. Environ. Int. 12:369-388.

Wallace, L.A., E.D. Pellizzari, T.D. Hartwell, C. Sparacino, R. Whitmore, L. Sheldon, H. Zelon, and R. Perritt. 1987. The TEAM (Total Exposure Assessment Methodology) study: personal exposures to toxic substances in air, drinking water, and breath of 400 residents of New Jersey, North Carolina, and North Dakota. Environ. Res. 43:290-307.

Wallace, L.A., E.D. Pellizzari, T.D. Hartwell, R. Whitmore, H. Zelon, R. Perritt, and L. Sheldon. 1988. California TEAM study: breath concentrations and personal air exposures to 26 volatile compounds in air and drinking water of 188 residents of Los Angeles, Antioch, and Pittsburgh, California. Atmos. Environ. 22:2141-2164.

Walter, S.D. 1991. The ecologic method in the study of environmental health. II. Methodologic issues and feasibility. Environ. Health Perspect. 94:67-73.

Wegman, D.H., E.A. Eisen, S.R. Woskie, and X. Hu. 1992. Measuring exposure for the epidemiologic study of acute effects. Am. J. Ind. Med. 21(1):77-89.

Whorton, M.D., R.W. Morgan, O. Wong, S. Larson, and N. Gordon. 1988. Problems associated with collecting drinking water quality data for community studies: a case example, Fresno County, California. Am. J. Pub. Health. 78:43-46.

Wrensch, M., S. Swan, P.J. Murphy, J. Lipscomb, K. Claxton, D. Epstein, and R. Neutra. 1990a. Hydrogeologic assessment of exposure to solvent-contaminated drinking water: pregnancy outcomes in relation to exposure. Arch. Environ. Health 45:210-216.

Wrensch, M., S. Swan, J. Lipscomb, D. Epstein, L. Fenster, K. Claxton, P.J. Murphy, D. Shusterman, and R. Neutra. 1990b. Pregnancy outcomes in women potentially exposed to solvent-contaminated drinking water in San Jose, California. Am. J. Epidemiol. 131:283-300.

4

Researching a Broad Range of Health Outcomes

HEALTH OUTCOMES OF INTEREST to environmental epidemiologists range from well-characterized diseases, such as cancer, to syndromes or constellations of clinical measures of physiologic or neurobehavioral function. Changes in these health outcomes can indicate that environmental factors are involved. However, not all relations between health end points and environmental exposures have been recognized. In order to identify changes in chronic diseases, the expected or baseline rates of diseases need to be determined, as has already been done for cancer. This is difficult because case definitions for many chronic diseases and syndromes are not uniform or well defined, registries and codings are not consistent, many demographic variables may influence the outcome, and time trends may not be reliable.

Many environmental toxins, acting either independently or in combination with other agents, may affect several organ systems. We consider problems in defining and studying a variety of health end points, especially those not associated with cancer. We identify gaps in current knowledge, suggest several noncancer end points that require additional research, and highlight some of those research opportunities. Finally, we discuss biologic markers for these end points and considerations of special populations at risk.

A comprehensive assessment of the evolving literature on health outcomes that may be associated with environmental exposures is beyond the scope of this study. Readers are referred to reports from the National Research Council on some of these health outcomes (NRC, 1989a,b, 1991, 1992a,b,c, 1993).

Most environmental epidemiology has been concerned with relatively few health outcomes, such as the neurologic and mental outcomes associated with lead and methyl mercury, respiratory disease associated with air pollution, and malignant neoplasms associated with exposures to various chemical agents in the environment. However, there is increasing evidence that a much broader array of health outcomes—such as neurologic, respiratory, and reproductive end points—may also be associated with environmental exposures. Examples of the environmental associations that have been investigated are a link between congenital malformations and trichloroethylene in drinking water, increased hospitalization of children and adults for asthma associated with air pollution, and the neurologic consequences of exposures to lead and solvents in the environment (NRC, 1991). The factors causing many noncancer outcomes, the underlying mechanisms of causation, and relative contributions of the various causal agents have not been clearly delineated. Further, an individual or community may experience more than one health outcome. Persons may be exposed to multiple chemicals (sometimes from the same source), each producing different outcome events, or a single chemical can be responsible for different outcomes in the same or different individuals.

Research into many outcomes potentially related to environmental exposure is at an early stage of development. In part, this is because there has been little effort to measure the incidence and prevalence of many chronic diseases. It is difficult to identify, monitor, and study populations at risk. Finally, more-refined measures of disease that can be applied to groups of persons are needed.

This chapter provides a brief overview of health outcomes that have been associated with environmental exposure. We also review outcomes for which evidence from toxicologic or occupational studies suggests associations with environmental exposure, but where current environmental-epidemiologic data are inadequate to provide definitive information about risk.

RESPIRATORY OUTCOMES

A diverse and growing literature characterizes the effects of air pollution on human health. Progress over past decades has elucidated the link of an array of respiratory outcomes to air pollution. Adverse effects on the respiratory tract of air pollutants such as ozone or airborne particles are highly nonspecific and not easily detected clinically. Nevertheless, for many common diseases, even small relative risks (RRs) may translate into significant numbers of cases of disease because of the large size of the exposed population.

For example, studies of the effects on animals of long-term exposure to ozone at concentrations similar to those seen in some US cities showed chronic lung damage that increased monotonically with cumulative dose; there was no evidence of a threshold (American Thoracic Society, 1996a). If similar effects also occur in humans, a substantial proportion of the population in the United States is exposed every summer to ozone concentrations that may produce chronic lung damage. Analyses of data in 8 cities have detected associations of airborne particles with small increases in the risk of mortality (Schwartz, 1991). In a community study, the closing of a steel mill for a year was associated with more than a 40% decrease in hospital admissions for asthma in children; the next year the mill reopened, and hospitalization rates rose to their previous level (Pope 1989, 1991). Another recent study found that rates of acute bronchitis in children were about twice as high in a town with particle concentrations at the US air-quality standard as they were in a community with near background concentrations; intermediate communities had intermediate rates (Dockery et al., 1989). Braun-Fahrlander et al. (1992) found major changes in respiratory symptoms among schoolchildren at concentrations commonly seen in the US urban population.

ACUTE HEALTH EFFECTS

Acute health effects of air pollution have long been established, although primarily by studies at concentrations of pollutants far higher than those now typical in developed countries. Recently, however, exposure to air pollution has been associated with acute outcomes—such as reversible reductions in lung function, increased respiratory symptoms and illness, emergency-room visits and increased hospitalization, and deaths from respiratory and cardiovascular diseases—at exposure levels far lower than those at which earlier data were collected. These are discussed briefly below in order of increasing severity. Several recent reports provide more-complete treatments of the evidence on acute health effects (Dockery and Pope, 1994; American Thoracic Society, 1996b).

Exercising volunteers exposed to ozone at concentrations below current air-quality standards show reversible decrements in lung function (McDonnell et al., 1985). Studies of ozone-exposed children in summer camps (Spektor et al., 1988; Lioy et al., 1985) have found that they have similar responses at even lower levels, at least in the summer months, when they are outdoors during much of the day. Smaller but still significant effects of ozone on lung function have been reported in schoolchildren during the school year, when outdoor activity is less (Kinney et al., 1989). Studies of lavage fluid from the lungs of volunteers have also provided evidence of inflammatory processes in the lung following ozone

exposure (American Thoracic Society, 1996a). Short-duration exposures of asthmatics to SO_2 also reduce lung function; in the protocols, delivering the gas during exercise increases the amount of SO_2 received (American Thoracic Society, 1996b).

Short periods of moderately elevated particle concentrations have been associated with pulmonary-function deficits (Dassen et al., 1986; Brunekreef et al., 1991). Subsequent studies that have examined daily time series, rather than episodes, have also associated ambient PM_{10} concentrations with short-term variation in peak expiratory flow rate at concentrations below current national ambient-air quality standards (Pope and Dockery, 1992; Pope et al., 1991).

Daily symptom incidence and duration of respiratory illness have been linked to exposure to airborne particles. Supporting evidence includes results from the Six Cities Study conducted by Harvard investigators (Schwartz et al., 1989), in which particle concentrations never exceeded 75% of the current air-quality standard. In this study, daily diary responses concerning respiratory symptoms were significantly associated with particle concentrations. Similar findings come from Switzerland (Braun-Fahrlander et al., 1992) and Provo, Utah (Pope and Dockery, 1992; Pope et al., 1991). Acid-aerosol exposure has been associated with increased symptoms in a diary study of asthmatics (Ostro et al., 1991). For several of the respiratory symptoms, the odds ratios from Utah, the Six Cities Study, and Switzerland are similar.

Exposure to airborne particles has been associated with increased rates of bronchitis in children (Dockery et al., 1989) at concentrations below current standards and also with increased rates of croup attacks in children (Schwartz et al., 1991). These effects do not appear to be limited to children. For example, Ostro and Rothschild (1989) have reported an association between airborne particles and ozone and respiratory illness severe enough to restrict activity in adults. A meta-analysis has reported a significant association between NO_2 exposure and respiratory illness in children (Hasselblad et al., 1992).

Bates and Sizto (1987) reported that exposures to both ozone and sulfate were associated with increased incidences of hospitalization for respiratory illness in Ontario. Pope (1989), as mentioned above, found sharp changes in hospitalization rates for children when a steel mill closed and then reopened. Hospital admissions were also increased in the German smog episode of 1985 (Wichmann et al., 1989) at a time of sharp increases in both total suspended particles and SO_2. Sunyer et al. (1991) reported that airborne particles and SO_2 were associated with hospitalization for respiratory illness in Barcelona, Spain. Hospital emergency-room visits were associated with sulfates and SO_2 in a study in Vancouver, Canada (Bates et al., 1990), and respirable particles were associated with hospital

emergency-room visits in Israel (Gross et al., 1984). Schwartz et al. (1993) reported that inhalable-particle concentrations that never exceeded the current air-quality standards were associated with increased emergency-room visits for asthma.

Studies have associated airborne particles with increased mortality (Dockery and Pope, 1994; American Thoracic Society., 1996b). Those studies have shown similar patterns of disease in areas with different mean temperatures and climatic conditions and in areas with air-pollution peaks in both winter and summer. Other studies have used somewhat different methods, so their estimates of effect size are not directly comparable. However, Fairley (1990) and Schwartz and Marcus (1990) both reported that optical measures of airborne particles were associated with daily mortality. Qualitatively, the effect-size estimates as assessed by the regression coefficients seemed similar. In contrast, Hatzakis et al. (1986) reported that mortality in Athens, Greece, was primarily associated with SO_2 and not with particulate matter, though data for the winter season alone (Katsounnayi et al., 1990) showed a principal association with particulate matter. Kinney and Ozkaynak (1991) have reported associations of both ozone and airport visibility (as a proxy for particles) with daily mortality in Los Angeles, Calif. The ozone association was stronger in that study.

CHRONIC EFFECTS OF AIR POLLUTION

Spektor et al. (1991) have reported that long-term exposure of children to particulate air pollution was associated with impaired lung function. Data from the NHANES II survey (Schwartz, 1989) showed that chronic exposure to particles or ozone was associated with lower lung function in children; the ozone effect was stronger. Chestnut et al. (1991) have reported that long-term TSP exposure was associated with lung-function decrements in adults.

Some studies have associated differences in long-term exposure to air pollution with increased rates of chronic respiratory illness. For example, Euler et al. (1987, 1988) have reported that an index of cumulative exposure to TSP was associated with increased rates of chronic bronchitis in a cohort of Seventh-Day Adventists in California. The association remained when ozone exposure was also examined. A weaker association was found with ozone exposure. Detels et al. (1987) have reported differences in respiratory health among communities exposed to different levels of air pollution in the Los Angeles area.

Several studies (e.g., Lave and Seskin, 1977; Chappie and Lave, 1982; Lipfert, 1980; Ozkaynak et al., 1986) have sought to associate long-term differences in air-pollution concentrations with differences in age-adjusted

mortality rates across major urban areas. These studies have applied regression methods to data from multiple locations. Such studies are constrained by the difficulty of adequately controlling for potential confounding effects of other risk factors, such as smoking; this problem is avoided by time-series analyses within a single urban area. In general, research methods have improved over time, though much uncertainty remains. The magnitude of the excess mortality suggested by these studies is somewhat larger than that suggested by acute studies of daily mortality.

Studies have also been directed at specific diseases, for example, lung cancer. Archer (1990) examined deaths from lung cancer and other respiratory illnesses in 2 heavily populated valleys in Utah. Mortality rates were essentially identical in the 1940s and early 1950s. However, a steel mill was opened in one of these areas in the late 1950s. By the 1960s, a trend toward higher lung cancer and respiratory mortality was evident in this valley, and the trend was quite pronounced in the 1970s and 1980s.

NEUROLOGIC OUTCOMES

Neurologic and neurobehavioral changes can result from effects of agents on autonomic, peripheral, and central components of the nervous system. Thus, neurotoxins can have a wide range of effects, including changes in motor and sensory function or behavior, central nervous system damage, and cognitive, memory, and developmental changes. The effects on the autonomic system are primarily biochemical. Chronic exposures may affect the peripheral nervous system; even minor damage to myelin can affect nerve conduction velocities. Since the nervous system has only a modest capacity to regenerate, subtle damage can have serious and long-lasting effects (Tilson and Mitchell, 1992). Neurotoxic effects may be acute, chronic, or delayed. Time relationships may be more complicated for some neurotoxic exposures. For example, solvent exposure just before testing may interfere with performance tests for chronic neurotoxic effects (Melius and Schulte, 1981). Thus, the study design should use the proper exposure time range and the proper response time range for the situation being evaluated, e.g., chronic effects in a chronically exposed population. The full range of neurologic and behavioral effects of a toxic substance is rarely known (Xintaras et al., 1979).

During recent decades, increasing attention has been paid to subtle behavioral changes that may occur at doses of agents lower than those causing physical signs and symptoms (Gochfeld et al., 1991; Tilson and Mitchell, 1992). These changes may occur after inapparent and chronic exposures, and they may lead to functional impairment that would otherwise be subclinical. Such impairments, unnoticeable in an individual, can have a substantial impact on the population. For example, Needleman et

al. (1982) has estimated that a downward shift in IQ of 4 points would double the proportion of children with IQ less than 80. Environmental exposures have been suggested for such diseases as parkinsonism, amyotrophic lateral sclerosis, Alzheimer's disease, and several peripheral neuropathies (Tanner et al., 1987; Ngim and Devathasan, 1989; Kalfakis et al., 1991; Bos et al., 1991; McLachlan et al., 1991). Improved means to evaluate more-subtle neurobehavioral and neurophysiologic effects (Valciukas and Lilis, 1980; Xintaras et al., 1979) have led to major advances in assessing neurotoxic effects. These methods are often specific for certain types of neurobehavioral effects, e.g., visual versus auditory memory. Hence, a careful selection of tests is required for field surveys. Consideration must also be given to possible confounders, such as age, alcohol intake, and education. Sometimes a battery of tests may be needed to screen workers for the effects of a neurotoxic substance. Such a battery should be specific enough to measure functions related to known effects of the substance and heterogeneous enough to cover a variety of neurobehavioral functions.

Neurologic and behavioral effects have been assessed for exposures to some indoor air pollutants. Otto et al. (1992) exposed 66 healthy young male subjects with no history of chemical sensitivity to air, to clean air, and to a complex mixture of volatile organic compounds (VOCs). Participants reported more fatigue and mental confusion after exposure to the organic compounds. However, performance on 13 neurobehavioral tests was not affected. In another part of the study, eye and throat irritation, headache, and drowsiness increased or showed no evidence of adaptation during exposure, even though the intensity of odors decreased by 30% (Hudnell et al., 1992). The investigators concluded that these results indicate that irritation intensity and other symptoms are not related in a simple way to odor intensity. The findings suggest that the symptoms may not be a psychosomatic response to the detection of an unpleasant odor and that environmental odor pollution may affect neurobehavior. Instead, subthreshold levels of VOCs may interact additively and stimulate trigeminal nerve receptors. Many nonspecific-symptom clusters have an odor component. Noxious environmental odors might trigger symptoms by a variety of physiologic mechanisms, including exacerbation of underlying medical conditions, innate odor aversions or aversive conditioning, stress-induced illness, and possible phenomenal reactions. Whereas relatively consistent patterns of subjective symptoms have been reported among individuals who live near environmental odor sources, documentation of objective correlates to such symptoms would require the development of new research tools.

An example of a common environmental neurotoxicant that produces a variety of chronic effects is lead. Lead has been known to cause serious

cognitive damage to children since the pioneering work of Byers and Lord (1943). Historically, assessment of the neuropsychologic effects of lead has been hampered by inadequate markers of exposure; i.e., blood lead is a short-term marker, and levels may return to normal after exposure has ended (Needleman, 1986), even though subjects' past exposures have caused persistent physiologic effects. One approach to the exposure-assessment issue is use of lead in shed deciduous teeth as a marker of past exposures. In a cohort study, first- and second-grade students who were considered asymptomatic for lead were classified by dentin lead levels and then evaluated with a battery of neuropsychologic tests (Needleman et al., 1979). Children with high dentin lead scored significantly less well on the Wechsler Intelligence Scale for Children (revised), on 3 measures of auditory speech processing, and on a measure of attention (Needleman et al., 1982). The authors concluded that lead, at doses below those that produce clinical symptoms, is associated with impaired neurobehavioral functioning (Needleman, 1986).

Although recent studies have found negative associations between blood lead concentrations and the full-scale intelligence quotient (IQ), not all have been statistically significant after control for some potential confounders (Needleman et al., 1979; Fulton et al., 1987; Hatzakis et al., 1987; Fergusson et al., 1988; Yule et al., 1981; Lansdown et al., 1986; Schroeder et al., 1985; Hawk et al., 1986; Bellinger et al., 1987). These studies are supported by studies in animals. In primates, Rice and Wiles (1979) found learning and attention-deficit disorders. Studies in rodents have shown cognitive disorders and interference with the dopaminergic system in the brain (Cory-Schlecta et al., 1981). Inhibition of long-term potentiation in the hippocampal region by lead (Munoz et al., 1988) has also been demonstrated in rats at moderate blood lead levels, which is again consistent with learning disorders. Meta-analyses of the human data have reported evidence for the association (Needleman and Gatsonis, 1990; Schwartz et al., 1985). These associations may or may not indicate a cause-effect relation (see discussion in NRC, 1993).

Several lessons may be drawn from these studies. First, a toxicant may have numerous neurologic effects, though it may be difficult to measure real but subtle behavioral changes. Second, markers of neurotoxic effect are often obtained simultaneously with markers of dose, making it difficult to discern the temporal or pathologic sequence of exposure, dose, and response. Third, neurotoxicant-induced alterations of neurobehavioral, neurophysiologic, and neurochemical function are believed to precede morphologic evidence of toxicity and to be more sensitive. However, functional indicators can be compromised by the adaptive capacity of the individual, especially with moderate to low levels of exposure. If the function of the nervous system is viewed as an adaptive process, it is

logical to predict that at sufficiently high exposures the functional reserve of the individual will be depleted and performance will deteriorate (Tilson and Mitchell, 1983; Bleeker and Agnew, 1987). This implies a threshold at the individual level. However, if person-to-person variations in biologic effect are great, the overall group effect of exposure may have no detectable threshold and individual dose-response relationships may be lost because of influences by genetics, age, sex, prior experience, overall health, and adaptive capacities.

REPRODUCTIVE AND DEVELOPMENTAL OUTCOMES

The term *reproductive and developmental toxicity* refers to maternal, paternal, pregnancy, and fetal effects. The exposure of pregnant women to drugs or other toxicants has long been recognized to produce adverse pregnancy outcomes, and effects on both males and females of exposures before conception have also been reported.

FECUNDITY

Fecundity (biologic capacity to have a child, whether or not one does so) may be affected by several mechanisms, including decreased fertility (reproductive capacity expressed as number of liveborn children) of either partner or the desire to have children. A report that exposure to the pesticide dibromochloropropane affected the fecundity of exposed men (Whorton et al., 1977) spurred studies of more-subtle indicators to determine whether other chemicals may also adversely affect male reproductive capacity. Epidemiologic reports have linked paternal occupational exposures to adverse pregnancy outcomes (Davis et al., 1992). Data from the National Toxicology Program indicate that chemicals affecting the fecundity of male rodents usually affect the fecundity of the female as well. Research to assess the fecundity and fertility of women exposed to potential toxicants is growing.

Assessment of the effects of potential toxicants on male fecundity is conducted primarily through endocrine and semen analyses. The basic methods (Schrader et al., 1987, 1992) attempt to evaluate effects on the neuroendocrine system, testes, and accessory sex glands. Hormone levels are used to assess the neuroendocrine system. Sperm count, morphology, and morphometry are used to assess testicular function. Sperm-cell function, including motility and viability, and semen biochemistry are useful in evaluating accessory sex gland function.

Well-designed human field studies are needed to evaluate the many chemicals in our environment. Better laboratory methods are needed to assess semen, genetic damage in the sperm, and sexual function. Research

is also needed in recruitment strategies to increase participation rates in field studies, thus decreasing potential bias. Research is needed to understand the implications of animal data for human risk.

Gaps in knowledge about fecundity are even larger for females than for males. Data from well-designed human studies of potential exposures are needed. Data on the "normal" values and their interindividual and intraindividual variations are needed for designing future studies. Most early studies of women and reproductive toxicity were centered around adverse outcomes of pregnancy, with little interest in the women's fecundity. The data reported by the National Toxicology Program's Continuous Breeding Protocol indicated that many toxicants being tested were affecting female rat fecundity.

New methods are now being assessed to detect neuroendocrine function, biochemical changes, and ovulation in humans (Hughes, 1988; Wright et al., 1992). Research needs include the development of practical, noninvasive methods for field studies of exposed humans, including better epidemiologic methods for the recruitment and study of human populations. Further research should address female sexual function, accessory sex organs, and genetic damage to the ova.

PREGNANCY OUTCOMES

Adverse pregnancy outcomes include miscarriage and low birthweight, as well as malformations and functional abnormalities. Most of the findings on chemicals causing adverse pregnancy outcomes have been established in animal studies. Two major exceptions illustrate that human fetuses are at risk from chemical exposures: the therapeutic use of thalidomide (McBride, 1961) and environmental exposures to mercury in Minamata Bay, Japan (Koos and Longo, 1976).

A few noninvasive laboratory methods are useful for the assessment of adverse effects on human pregnancies. Recently developed "ultrasensitive" assays of human chorionic gonadotropin can detect pregnancy losses around the time of the first postconception menstrual cycle, and studies using this technology may provide insight into early pregnancy loss due to environmental exposures (NRC, 1989a).

BIRTH DEFECTS AND DEVELOPMENTAL EFFECTS

Developmental effects, including birth defects, have been studied as possible effects of toxic pollutants from point sources. This is because such effects can be readily monitored with either existing surveillance systems or special studies and because they have shorter latency periods than

cancer. Some of the studies evaluating exposures that might increase the risk of birth defects, congenital anomalies, and low birthweight were reviewed in volume 1 (NRC, 1991). One of these sets of studies, performed by the California Department of Health Services (CDHS), was updated in a special issue of *Epidemiology* (Swan et al., 1992; Deane et al., 1992; Zierler, 1992).

CDHS investigated cardiac defects among babies born in Santa Clara County in 1981 to September 1, 1982. The investigators compared the rates of cardiac defects in the study area (suspected of water contamination) with the rates for the rest of the county. Investigators found an increased risk of cardiac defects in the study area (RR = 2.6, p = 0.01). The results could not be due to recall bias, because the information is documented by charts. However, the investigators also compared the times when the birth defects occurred with when the water contamination occurred and concluded that the timing could not link them. Investigators mention other possible exposures, such as air contaminants and other contaminants in the water, that might have increased the rates of cardiac defects.

In a second study, CDHS compared the Los Paseos census tract, which received water from one contaminated well, with a control census tract that had demographic characteristics similar to those of Los Paseos. Women who had been pregnant in 1980 and/or 1981 were contacted by mail, telephone, or home visit and interviewed to determine rates of spontaneous abortion, congenital anomalies, and low birthweight, as well as various risk factors. After adjustment for confounding variables, the Las Paseos spontaneous abortion rate was 2.3 times that in the control community (odds ratio, OR = 2.3, 95% confidence interval, CI = , 1.3–4.2). Drinking tap water was also found to be associated with spontaneous abortions in both the control area and case area. The congenital anomaly rate was 3 times that of the control area (OR = 3.1, 95% CI = 1.1–10.4). However, the authors commented that the observed pattern is not consistent with a single teratogen, but rather with several teratogens affecting the fetus at different times during gestation. They did not find any low-birthweight babies in the Los Paseos area. Tests for chemicals in the control area did not show chemical contamination. The investigators concluded that the exposure data were insufficient to determine whether the leak into the contaminated well was a cause of the increased rates of spontaneous abortions and congenital malformations.

The relation between environmental exposures and birth defects remains important but difficult to study. An innovative study from New York state used existing data to detect some associations between congenital malformations (all types pooled) and residential proximity to hazardous-waste sites (Geschwind et al., 1992). This study examined unusu-

ally large numbers of subjects and used more than one exposure variable. The investigators linked a congenital-malformations registry and a hazardous-waste site inspection program to evaluate this relation. The congenital-malformations registry collects information from hospitals, medical facilities, and private physicians on all children born alive in New York state with a diagnosis before 2 years of age of congenital malformations, chromosomal anomalies, or persistent metabolic defects. Ongoing audits show that the data are 95% complete and accurate. They linked the information from this database to residence within 1-mile of 590 waste sites in New York state, excluding those in New York City and those in rural areas that were not divided into census tracts. Investigators examined data on 9,313 newborns with congenital malformations and 17,802 healthy controls. Regressions controlled for maternal age, race, education, complications during pregnancy, parity, population density, and sex of child. They found a statistically significant OR of 1.12 (95% CI = 1.06–1.18) for bearing children with any kind of congenital malformation. Within the 1 mile area, babies born near off-site chemical leaks had higher rates of malformation than did those in areas without chemical leaks (RR = 1.17, 95% CI = 1.08–1.27).

Some studies indicate that the development of children exposed in utero to polychlorinated biphenyls (PCBs) is moderately impaired. Groups of children in Michigan and North Carolina have been studied from birth to age 4 years. In Michigan, higher PCB levels in umbilical cords were linked to reduced head size, diminished chest girth, and shorter gestation; these effects also occur in children whose mothers have been exposed to PCBs occupationally (Tilson et al., 1990). Higher PCB levels were also associated with lower scores on standardized tests for infant development and reduced activity. In North Carolina, a study of 912 infants followed since birth has shown that levels of PCBs commonly encountered in the United States "produce detectable effects on motor maturation and some evidence of impaired infant learning" (Tilson et al., 1990).

Lead has long been linked to adverse pregnancy outcomes and, indeed, was used as an abortifacient early in this century. Fetal loss has been associated with lead exposure, although control for potential confounding factors in those studies was poor. More recently, occupational lead exposure has been associated with several kinds of adverse pregnancy outcomes (Schwartz, 1992). Prospective studies of pregnant women who were not occupationally exposed have associated higher blood or cord lead levels with lower birthweight, small size for gestational age, and shorter duration of gestation. Higher placental lead levels have been associated with pregnancy loss (Schwartz, 1992). Not all these outcomes have been seen in every study, however.

HEPATIC AND RENAL OUTCOMES

The liver and kidneys are directly involved in the body's handling of chemical exposures. The liver is the primary organ for metabolism of these substances, and the kidney is the primary organ for excretion of them and any toxic metabolic products. Both tend to concentrate toxic substances to levels well above those in the blood. Thus, the liver and kidney may have unusual patterns of exposure. The health effects of environmental toxicants on the liver can include enlargement, necrosis, fibrosis, cirrhosis, veno-occlusive disease, granulomas, lymphocyte infiltration, and cancer (Van Thiel, 1986; Tamburro and Liss, 1986).

Human exposures to low levels of many environmental and occupational chemicals can cause adaptive changes in the liver that may indicate exposure or early disease (Rubin, 1987; Van Thiel, 1986). In some situations, adaptive responses, though not toxic themselves, can indicate exposure. For example, various pesticides may have effects on liver function that indicate exposure. Several biologic indexes provide markers of hepatic function. The kidney has the dual role of separating unwanted from wanted substances and excreting the former. The high blood flow through the kidneys exposes them to the variety of chemicals absorbed as a result of breathing, ingesting, or skin absorption. Identifying and preventing environmentally induced renal disease is important because most types of renal disease are irreversible once substantial loss of renal function has occurred. Renal toxicants may produce glomerular nephritis, tubular necrosis, Fanconi syndrome, interstitial nephritis, hypersensitivity reaction, kidney stones, and cancer.

The influence of environmental exposures on the kidney has been widely described in the occupational literature (Littorin et al., 1984; Bernard et al., 1979; Friberg et al., 1985, Druet et al., 1982). A major impediment to recognition of early kidney disease is the lack of clinical and laboratory tests that are sensitive and specific (Goyer, 1987). The most-common clinical indicator of renal disease, serum creatinine, is not generally abnormal (between 1.5 and 2.0 mg/dL of blood) until 50% or more of renal function is lost.

IMMUNOLOGIC EFFECTS

A wide range of immunologic changes can result from exposures to xenobiotics. These have been extensively reviewed in recent reports from NRC and others (NRC, 1992a; Luster et al., 1989; Koller, 1987). The immune system defends the body by responding to exposure or stress. Distinguishing the changes that are defensive from those that are toxic is a current challenge.

As markers of adverse immunologic change are refined and validated, environmental-epidemiologic studies will be able to incorporate them in studies of exposure and effect. For further information on this developing field, readers are referred to *Biologic Markers in Immunotoxicology* (NRC, 1992a).

BIOLOGIC MARKERS IN ENVIRONMENTAL EPIDEMIOLOGY

Environmental-epidemiologic research of the future will include biologic markers of exposure, effect, and susceptibility, which were reviewed in volume 1 and in other NRC reports. These markers can reduce misclassification of exposure, refine the classification of disease, identify mechanisms of causation, and pinpoint high-risk populations and persons. The primary gap in existing information on various markers is the need for validation (NRC, 1989b). This requires that assays be assessed for sensitivity, specificity, and reproducibility. Once they are validated, laboratories must be able to conduct the relatively large number of assays generated by epidemiologic studies (Schulte, 1991; Hulka, 1991). In addition, costs of most biologic markers must be reduced before they will become relevant for epidemiologic study designs. Further, epidemiologic use will require information on how a marker varies with a wide range of factors, such as age, race, sex, time of day, season, drinking, smoking, use of medication, exercise, and concurrent diseases. With such information in hand, it will be easier to tell which markers to test for, how to use them, and how to interpret the findings.

Developments in molecular biology over the past decade may fundamentally alter the field, as well as provide a richer array of end points for statistical analysis. Axelson (1991) has argued that biomarkers could lead to new definitions of disease entities that combine clinical or histopathologic criteria with biochemical or genetic characteristics. Thus, recent work on the metabolism of debrisoquine, antipyrine, and other compounds has indicated that persons with some enzyme patterns have increased relative risks of cancer of the lung, pancreas, and stomach. Subjects who were extensive debrisoquine metabolizers and who also had occupational exposure to asbestos had an 18-fold increase in the rate of lung cancer.

There has been little research on indicators or markers of early or preclinical environmentally induced disease. Identification of such indicators can serve as both an early warning of potential problems and a guide to interventions for prevention or control. The key question in this regard is when a biologic change indicates a disease (Goyer and Rogan, 1986). From the epidemiologic perspective, a biologic change that is a candidate as an early indicator of disease must significantly differ between those

who will develop disease (in the absence of intervention) and those who will not.

The use of laboratory tests to assess biologic change requires the availability of validated biologic markers, the collection and storage of biologic specimens, and methods to integrate biologic measurements with observations, such as data from questionnaires. The validation of markers pertains to both the laboratory and the field. Before use in environmental epidemiology, population prevalence, sensitivity, and specificity of a marker need to be determined. This determination is also influenced by how markers vary among groups characterized by demographic, behavioral, or genetic factors. Once a study using biologic markers is planned, it is necessary to consider how specimens will be collected, transported, and stored, which can affect the ability to detect associations between exposures and outcomes.

Ultimately, biologic markers will be used in environmental epidemiology only if they are shown to contribute to the understanding of environmental hazards. Markers are not an end, but only a means to an end. Researchers should be able to identify what research questions each marker will answer and how the biomarker data will be used before collection of biologic specimens. Failure to consider these factors can lead to wasteful efforts.

Field studies of chemically exposed populations have made little use of immune-system responses to define exposures (NRC, 1989b; Karol, 1987; Grammer et al., 1988; Stejskal et al., 1986; Zeiss et al., 1977). There are several possibilities for developing the field of immunotoxicology (NRC, 1992). Better use of biomarkers of immunologic function requires an understanding of the interval between exposure and the appearance of the marker. Ideally, a specific exposure will cause the expression of a specific marker at a known time after exposure. However, human populations are generally exposed to mixtures, and a marker or battery of markers may be able to depict only qualitatively whether individuals have been exposed to immunomodulating substances (Biagini et al., 1986). Exposure assessment might be improved by using ordinal- and interval-scale characterizations of dose-dependent immunologic responses as measures of exposure. Karol (1983, 1987) demonstrated that an immunologic response (immunoglobulin G cytophilic antibodies) was dose-dependent over the range of 1-30 mg/kg in guinea pigs exposed to toluene diisocyanate vapors.

A marker cannot be used to estimate exposure without some data on reference levels in the unexposed population. There is a need to standardize the collection of such information and obtain appropriate demographic and other information about individual subjects. Even when information from the general population is available, data that assess immunologic markers of exposure or effect must be obtained from spe-

cific control populations that are adequately characterized in such variables as age, race, sex, geographic region, socioeconomic status, nutritional status, life style, and sexual habits.

The selection of immunologic markers to be used in a battery of tests to define exposure requires careful study. A battery of immunologic markers for assessing exposure might be quite different from one used to assess toxicity. Conversely, there may be value in developing a single battery that is less than optimal for specific tests but is familiar, as well as widely known and widely used. If the battery is to be comprehensive, it is often useful to employ some nonimmunologic markers as well. For example, the determination of a covalently bound DNA adduct, coupled with some measure of immunoreactivity, might provide a useful measure of exposure and immunologic response.

Since the immune system generally has some degree of functional reserve capacity, it is not likely that gross indicators, such as cell counts, will be useful as early indicators of exposure (Bick, 1985). One purpose of individual exposure assessment is to identify members of a possibly exposed population who are actually exposed.

The normal range of some immunologic parameters in humans includes extensive interindividual and intraindividual variability. In a 14-month study of healthy individuals between 21 and 70 years of age, the coefficient of variation for 13 of 16 immunologic parameters was greater than 30% (Dorey and Zighelboim, 1980). However, for all the parameters measured in a short period, intraindividual variability was significantly less than interindividual variability. The interindividual variability was due, in general, to subjects who consistently exhibited functional levels and cell numbers that were significantly higher or lower than the population mean. No significant correlation was found between age and most of the immunologic parameters examined, though this does not rule out significant age-related changes in other immune end points. Animal studies have shown age-related changes in immunity (Bick, 1985). Large intraindividual coefficients of variation determined for various immunologic parameters do not preclude the use of such markers to assess exposure or effect. In industrial and environmental exposure measurements, it is not unusual to find coefficients of variation as high as those found with immunologic parameters. Some implications of large coefficients of variation in immunologic markers that may be used to define exposure are that large samples must be used to reduce random variation, comparisons between exposed and nonexposed persons might involve stratification or adjustment for populations whose results significantly deviate from the mean, and construct validity must be established for multiple markers to ensure that confounders are distributed similarly between exposed and nonexposed groups.

In summary, immunologic markers of exposure can be used in several ways to distinguish exposed from nonexposed individuals and to distinguish among those with different levels of exposure within exposed groups. The use of immunologic markers should be considered in the context of other information, such as exposure history or environmental or breathing-zone measurements. The use of biologic markers may need consideration during the design of a study, during implementation, and during the interpretation of results. However, before any study, the investigator should ask what benefit biologic markers might provide over exposure assessment by ambient personal or environmental monitoring. If this question cannot be answered in a convincing way, biologic markers should not be used in environmental-epidemiologic research.

SUSCEPTIBLE POPULATIONS

The frequency of many diseases is likely to increase in certain population groups. For example, an increasingly large proportion of the population is over age 65 (Cooper et al., 1991). Older persons have an increased vulnerability to many stresses. Whether and to what extent an elderly person is more susceptible to the toxic effects of potentially hazardous compounds from the environment is largely unknown. Elderly individuals may be at increased risk from toxicant exposure because of age-related changes in the body's protective mechanisms (NRCb, 1989). Older persons have had a long period of exposure to chemicals, more time for latent adverse effects to manifest themselves, and more time for cumulative effects to emerge.

The role of aging in environmental toxicity has been a subject of extensive research. In aged populations, 2 classes of adverse effects may exist: those caused by aging alone and those caused by interactions between aging and toxic agents (Williams et al., 1987). It may not be possible to discern with confidence the health decrements from these. A problem, then, is the identification of experimental approaches that will give insight into the relation between age and toxicity (NRC, 1989b).

A major impediment to understanding the role of environmental agents in the aging process is the paucity of reliable and valid biomarkers of aging per se (Reff and Schneider, 1982; Ingram, 1988; McClearn, 1988). Harrison (1988) suggested 4 criteria for determining whether a particular physiologic assay may be useful as a biomarker of aging in an individual organism: the results should change significantly with age, the changes should be repeatable in the same individual, assays of independent physiologic parameters should give similar estimates of age for the same individual, and the degree of aging as determined by the assays should correlate with subsequent longevity.

Children also have heightened sensitivity to many environmentally induced diseases. The susceptible period of childhood can be viewed as beginning with influences on the sperm and egg and, hence, on preconception parental exposure and life styles, continuing with the impacts on embryos and fetuses, and including the particular susceptibility of neonates and infants. Research is needed to catalog environmental factors that influence each of these stages (NRC, 1989b). The immature enzymatic detoxification systems of embryos, fetuses, and neonates put them at increased risk of the effects of pharmacologic and environmental substances that pass through the placenta, such as heavy metals, PCBs, and immunogens.

Information is scanty about racial and ethnic differences in susceptibility to disease, though these also merit careful consideration. Lead poisoning affects children of all races but occurs with especially high frequency in inner-city African Americans. Other diseases also occur more commonly among different ethnic groups and reflect genetic, nutritional, and environmental factors that need to be systematically evaluated. Personal choice of life style, such as smoking or use of alcoholic beverages, and personal exposures, such as use of medication, may change susceptibility.

The recent and rapid increase in knowledge about molecular biology and genetics has identified, at the gene or molecular level, an increased number of susceptible groups. These have implications for research and for societal response to environmental hazards. The key to future epidemiologic research is the linkage of susceptibility markers to exposures and outcomes in ways that will answer questions about why, in similarly exposed groups, some people develop disease and others do not. Susceptibility markers can be used to differentiate subgroups of populations in order to identify risks.

The ability to distinguish sensitive subgroups in populations has legal, ethical, and social implications that need to be assessed in detail. For example, labeling individuals as susceptible has been alleged to have effects on psychosocial status and function, job security, property values, obtaining insurance, and finding a mate (Ashford et al., 1990). Recent evidence that each human has, on average, 5-10 genetic abnormalities may diminish the social and psychologic impact of knowing about a specific abnormality in a specific person.

RECOMMENDATIONS

The types, nature, and extent of environmental influences on humans are unclear. Substantial evidence associates air-pollution exposure with some respiratory outcomes. The toxicities of lead are increasingly well understood. The rate of environment-related risks of other outcomes is

less well understood, and further work is necessary to delineate environmental risks. Researchers should use a range of strategies to study possible relations of health outcomes to environmental exposures. These strategies include the following:

- Agencies concerned with the prevention of disease and the promotion of health should make a deliberate effort to identify diseases and syndromes that appear to be increasing or to have an important impact on public health. Specific causes of death and illness, as well as such functional afflictions as impaired vision and effects on reproductive health, need to be considered. Nomenclature and case definitions should be standardized for these diseases and syndromes.
- When time trends reveal recent shifts in disease patterns, researchers should explore possible etiologic explanations. Health-information databases should be designed to facilitate linkage to exposure information and other relevant factors, and vice versa, so that analyses can be extended from local patterns to relations with possible causal factors.
- Researchers should search for and study cohorts with exposures to toxicants that might be associated with specific health outcomes in order to test hypotheses about association. Whenever possible, high-exposure populations and those with a range of exposures should be selected.
- Opportunities should be exploited to add environmental assessments to existing research on end points where there is a potential environmental risk. This could be a cost-effective approach to epidemiology.
- Broad evaluations of symptom prevalence and disease history should be a part of selected environmental investigations, and findings that suggest effects on specific organ systems should be followed with more-focused studies.
- Tools to identify susceptible subpopulations need substantial further development and should be used to identify associations of exposures with specific effects. A broad range of criteria—such as age, race, metabolic phenotypes, and socioeconomic status—should be used.
- Data gaps exist in reproductive epidemiology, and we urge renewed research focus on this field. Important data gaps occur in the study of pregnancy, including early pregnancy loss, and extrapolations from animal studies have generally not been validated in humans. Data gaps also exist on male-mediated adverse pregnancy outcomes. An important research need is the development of measures and tools to identify toxic effects on fecundity.
- Longitudinal studies should be conducted to assess the significance of biologic markers and preclinical effects suggestive of disease, including their use in calculating and assessing past exposures.
- Biologic markers need to be validated for specificity for exposure

and health outcome. Many of the markers that have been identified and characterized in the laboratory have not been specifically tied to chemical, biologic, or physical exposures. Validation of these will strengthen the ability of environmental epidemiology to discern relations between exposure and disease. Research is needed to identify additional direct and indirect markers of exposure and disease.

- Environmental exposure to some substances is now routinely monitored. The planning of such monitoring should incorporate factors that will increase its utility for epidemiologic studies.

REFERENCES

American Thoracic Society, Committee of the Environmental and Occupational Health Assembly: R. Bascom, P.A. Bromberg, D.A. Costa, R. Devlin, D.W. Dockery, M.W. Frampton, W. Lambert, J. M. Samet, F. E. Speizer, and M. Utell. 1996a. Health effects of outdoor air pollution. Part 2. *Am. J. Resp. Crit. Care Med.* 153: 477-498.

American Thoracic Society, Committee of the Environmental and Occupational Health Assembly: R. Bascom, P.A. Bromberg, D.A. Costa, R. Devlin, D.W. Dockery, M.W. Frampton, W. Lambert, J. M. Samet, F. E. Speizer, and M. Utell. 1996b. Health effects of outdoor air pollution. Part 1. *Am. J. Resp. Crit. Care Med.* 153: 3-50.

Ashford, N.A., C.J. Spadafor, D.B. Hattis, and C.C. Caldart. 1990. Monitoring the Worker for Exposure and Disease: Scientific, Legal, and Ethical Considerations in the Use of Biomarkers. Baltimore: Johns Hopkins University Press. 224 pp.

Archer, V.E. 1990. Air pollution and fatal lung disease in three Utah counties. Arch. Environ. Health 45:325-334.

Axelson, O. 1991. Cancer and combined exposures to occupational and environmental factors. Rec. Res. Cancer Res. 122:60-70.

Bates, D.V., and R. Sizto. 1987. Hospital admissions and air pollutants in Southern Ontario: the acid summer haze effect. Environ. Res. 43:317-331.

Bates, D.V., M. Baker-Anderson, and R. Sizto. 1990. Asthma attack periodicity: a study of hospital emergency visits in Vancouver. Environ. Res. 51:51-70.

Bellinger, D., A. Leviton, C. Waternaux, H. Needleman, and M. Rabinowitz. 1987. Longitudinal analyses of prenatal leads exposure and early cognitive development. N. Engl. J. Med. 316:1037-1043.

Bernard A., J.P. Buchet, H. Roels, P. Masson, and R. Lauwerys. 1979. Renal excretion of proteins and enzymes in workers exposed to cadmium. Eur. J. Clin. Invest. 9:11-22.

Biagini, R.E., W.J. Moorman, T.R. Lewis, and I.L. Bernstein. 1986. Ozone enhancement of platinum asthma in a primate model. Am. Rev. Respir. Dis. 134:719-725.

Bick, P.H. 1985. The immune system: organization and function. Pp. 1-10 in J. Dean, M.I. Luster, A.E. Munson, and H. Amos, eds. Immunotoxicology and Immunopharmacology. New York: Raven.

Bleeker, M.L., and J. Agnew. 1987. The assessment of the dose-response relationship for low-level exposure to neurotoxicants in man. Pp. 508-515 in V. Foa, E.A. Emmett, M. Maroni, and A. Colombi, eds. Occupational and Environmental Chemical Hazards: Cellular and Biochemical Indices for Monitoring Toxicity. Chichester: Ellis Horwood.

Bos, P.M., G. de Mik, and P.C. Bragt. 1991. Critical review of the toxicity of methyl n-butyl ketone: risk from occupational exposure. Am. J. Ind. Med. 20:175-194.

Braun-Fahrlander, C., U. Ackermann-Liebrich, J. Schwartz, H.P. Gnehm, M. Rutishauser, and H.U. Wanner. 1992. Air pollution and respiratory symptoms in preschool children. Am. Rev. Respir. Dis. 145:42-47.

Brunekreef, B., P.L. Kinney, J.H. Ware, D. Dockery, F.E. Speizer, J.D. Spengler, and B.G. Ferris, Jr. 1991. Sensitive subgroups and normal variation in pulmonary function response to air pollution episodes. Environ. Health Perspect. 90:189-193.

Byers, R.K., and E.E. Lord. 1943. Late effects of lead-poisoning on mental development. Am. J. Dis. Child. 66:471

Chappie, M., and L. Lave. 1982. The health effects of air pollution: a reanalysis. J. Urban Econ. 12:346-376.

Chestnut, L.G., J. Schwartz, D.A. Savitz, and C.M. Burchfiel. 1991. Pulmonary function and ambient particulate matter: epidemiological evidence from NHANES I. Arch. Environ. Health 46:135-144.

Cooper, R.L., J.M. Goldman, and T.J. Harbin. 1991. Introduction: assessing environmental influences on the aging process. Pp. 1-6 in R.L. Cooper, J.M. Goldman, and T.J. Harbin, eds. The Johns Hopkins Series in Environmental Toxicology: Aging and Environmental Toxicology: Biological Perspectives. Baltimore: Johns Hopkins University Press.

Cory-Schlecta, D.A., S.T. Bissen, A.M. Young, and T. Thompson. 1981. Chronic post-weaning lead exposure and response during performance. Toxicol. Appl. Pharmacol. 60:78-84.

Dassen, W., B. Brunekreef, G. Hoek, P. Hofschreuder, B. Staatsen, H. de Groot, E. Schouten, and K. Biersteker. 1986. Decline in children's pulmonary function during an air pollution episode. J. Air Pollut. Control Assoc. 36:1223-1227.

Davis, D.L., G. Friedler, D. Mattison, and R. Morris. 1992. Male-mediated teratogenesis and other reproductive effects: biologic and epidemiologic findings and a plea for clinical research. Reprod. Toxicol. 6:289-292.

Deane, M., S.H. Swan, J.A. Harris, D.M. Epstein, and R. R. Neutra. 1992. Adverse pregnancy outcomes in relation to water consumption: a re-analysis of data from the original Santa Clara County study, California, 1980-1981. Epidemiology 3:94-97

Detels, R., D.P. Tashkin, J.W. Sayre, S.N. Rokaw, A.H. Coulson, F.J. Massey, Jr., and D.H. Wegman. 1987. The UCLA population studies of chronic obstructive respiratory disease. 9. Lung function changes associated with chronic exposure to photochemical oxidants; a cohort study among never-smokers. Chest 92:594-603.

Dockery, D.W., and C.A. Pope. 1994. Acute respiratory effects of particulate air pollution. An. Rev. Public Health 15:107-132

Dockery, D.W., F.E. Speizer, D.O. Stram, J.H. Ware, J.D. Spengler, and B.G. Ferris, Jr. 1989. Effects of inhalable particles on respiratory health of children. Am. Rev. Respir. Dis. 139:587-594.

Dorey, F., and J. Zighelboim. 1980. Immunologic variability in a healthy population. Clin. Immunol. Immunopathol. 16:406-415.

Druet P., A. Bernard, F. Hirsch, J.J. Weening, P. Gengoux, P. Mahieu, and S. Birkeland. 1982. Immunologically mediated glomerulonephritis induced by heavy metals. Arch. Toxicol. 50:187-194.

Euler, G.L., D.E. Abbey, A.R. Magie, and J.E. Hodgkin. 1987. Chronic obstructive pulmonary disease symptom effects of long-term cumulative exposure to ambient levels of total suspended particulates and sulfur dioxide in California Seventh-Day Adventist residents. Arch. Environ. Health 42:213-222.

Euler, G.L., D.E. Abbey, J.E. Hodgkin, and A.R. Magie. 1988. Chronic obstructive pulmonary disease symptom effects of long-term cumulative exposure to ambient levels of total oxidants and nitrogen dioxide in California Seventh-Day Adventist residents. Arch. Environ. Health 43:279-285.

Fairley, D. 1990. The relationship of daily mortality to suspended particulates in Santa Clara County, 1980-1986. Environ. Health Perspect. 89:159-168.

Fergusson, D.M., J.E. Fergusson, L.J. Horwood, and N.G. Kinzett. 1988. A longitudinal study of dentine lead levels, intelligence, school performance and behavior. III. Dentine lead levels and cognitive ability. J. Child Psychol. Psychiatry 29:793-809.

Friberg, L., C.G. Elinder, T. Kjellstrom, and G.F. Nordberg. 1985. Cadmium and health: a toxicological and epidemiological appraisal. 1:159-160.

Fulton, M., G. Raab, G. Thomson, D. Laxen, R. Hunter, and W. Hepburn. 1987. Influence of blood level on the ability and attainment of children in Edinburgh. Lancet 1:1221-1226.

Geschwind, S.A., J.A.J. Stolwijk, M. Bracken, E. Fitzgerale, A. Stark, C. Olsen, and J. Melius. 1992. Risk of congenital malformations associated with proximity to hazardous-waste sites. Am. J. Epidemiol. 135:1197-1207.

Gochfeld, M., N. Fiedler, and J. Burger. 1991. Neurobehavioral toxicology. Pp. 325-332 in J.M. Last and R. B. Wallace, eds. Maxcey-Roseneau-Last Public Health and Preventive Medicine. 13th ed. Norwalk, CT: Appleton and Lange.

Goyer, R.A. 1987. Biochemical and cellular indices of renal charges induced by exogenous chemicals. Report of session IV. Pp. 368-371 in Foa V., E.A. Emmett, M. Maroni, and A. Colombi, eds. Occupational and Environmental Chemical Hazards. Cellular and Biochemical Indices for Monitoring Toxicity. Chichester: Ellis Horwood.

Goyer, R.A., and W.J. Rogan. 1986. When is biologic change an indicator of disease? Pp. 17-26 in D.W. Underhill and E.P. Radford, eds. New and Sensitive Indicators of Health Impacts of Environmental Agents. Third Annual Symposium on Environmental Epidemiology, held April 26-28 in Pittsburgh, PA. Pittsburgh, PA: Center for Environmental Epidemiology, Graduate School of Public Health, University of Pittsburgh.

Grammer, L.C., P. Eggum., M. Silverstein, M.A. Shaughnessy, J.L. Liotta, and R. Patterson. 1988. Prospective immunologic and clinical study of a population exposed to hexamethylene diisocyanate. J. Allergy Clin. Immunol. 82:627-633.

Gross, J., J.R. Goldsmith, L. Zangwill, and S. Lerman. 1984. Monitoring of hospital emergency room visits as a method for detecting health effects of environmental exposures. Sci. Total. Environ. 32:289-302.

Harrison D.E., and J. R. Archer. Biomarkers of aging: tissue markers. 1988 Future research needs, strategies, directions, and priorities. *Exp. Gerontol*. 23: 309-325.

Hasselblad, V., D.M. Eddy, and D.J. Kotchmar. 1992. Synthesis of environmental evidence: nitrogen dioxide epidemiology studies. J. Air Waste Manage. Assoc. 42:662-671.

Hatzakis, A., K. Katsoynanni, A. Kalandidi, N. Day, and D. Trichopoulos. 1986. Short-term effects of air pollution on mortality in Athens. Int. J. Epidemiol. 15:73-81.

Hatzakis A., A. Kokkevi, K. Katsouyanni, K. Maravelias, J.F. Salaminios, A. Kalandidi, and A. Koutselinis. 1987. Psychometric intelligence and attentional performance deficits in lead-exposed children. Pp. 204-209 in S.E. Lindberg and T.C. Hutchinson, eds. Heavy metals in the Environment. Vol. 1. New Orleans/Edinburgh: CEP Consultants.

Hawk, B.A., S. R. Schroeder, G. Robinson, P. Mushak, D. Kleinbaum, and G. Dawson. 1986. Relation of lead and social factors to IQ of low-SES children: a partial replication. Am. J. Ment. Defic. 91: 178-183.

Hudnell, H.K., D.A. Otto, D.E. House, and L. Mohave. 1992. Exposure of humans to a volatile organic mixture. II. Sensory. Arch. Environ. Health 47:31-38.

Hughes, C.L. 1988. Monitoring of ovulation in the assessment of reproductive hazards in the workplace. Reprod. Toxicol. 2:163-170.

Hulka, B.S. 1991. Epidemiological studies using biological markers: issues for epidemiologists. Cancer Epid. Biomarkers Prevention 1:13-19.

Ingram, D.K. 1988. Key questions in developing biomarkers of aging. Exp. Gerontol. 23:429-434.

Kalfakis, N., D. Vassilopoulos, C. Voumvourakis, M. Ndjeveleka, and C. Papageorgiou. 1991. Amyotrophic lateral sclerosis in southern Greece: an epidemiologic study. Neuroepidemiology 10:170-173.

Karol, M.H. 1983. Concentration-dependent immunologic response to toluene diisocyanate (TDI) following inhalation exposure. Toxicol. Appl. Pharmacol. 68:229-241.

Karol, M.H. 1987. The development of an animal model for TDI asthma. Bull. Eur. Physiopathol. Respir. 23:571-576.

Katsounnayi, K., A. Hatzakis, A. Kalandidi, and D. Trichopoulos. 1990. Short-term effects of atmospheric pollution on mortality in Athens. Arch. Hellenic Med. 7:126-132.

Kinney, P.L., and H. Ozkaynak. 1991. Associations of daily mortality and air pollution in Los Angeles County (California, USA). Environ. Res. 54:99-120.

Kinney, P.L., J.H. Ware, J.D. Spengler, D.W. Dockery, F.E. Speizer, and B.G. Ferris, Jr. 1989. Short-term pulmonary function change in association with ozone levels. Am. Rev. Respir. Dis. 139:56-61.

Koller, L.D. 1987. Immunotoxicology today. Toxicol. Pathol. 15:346-351.

Koos, B.J., and L.D. Longo. 1976. Mercury toxicity in the pregnant woman, fetus, and newborn infant: a review. Am. J. Obstet. Gynecol. 126:390-409.

Lansdown, R., W. Yule, M.A. Urbanowicz, and J. Hunter. 1986. The relationship between blood-lead concentrations, intelligence, attainment and behavior in a school population: the second London study. Int. Arch. Occup. Environ. Health 57: 225-235.

Lave, L.B., and E.P. Seskin. 1977. Air Pollution and Human Health. Baltimore: Johns Hopkins University Press for Resources for the Future. 368 pp.

Lioy, P.J., T.A. Vollmuth, and M. Lippmann. 1985. Persistence of peak flow decrement in children following ozone exposures exceeding the National Ambient Air Quality Standard. J. Air Pollut. Control Assoc. 35:1069-1071.

Lipfert, F.W. 1980. Sulfur oxides, particulates, and human mortality: synopsis of statistical correlations. J. Air Pollut. Control Assoc. 30:366-371.

Littorin, M., M. Welinder, and B. Hultberg. 1984. Kidney function in stainless steel welders. Int. Arch. Occup. Environ. Health 53:279-282.

Luster, M.I., M.F. Ackermann, D.P. Germolec, and G.J. Rosenthal. 1989. Perturbations of the immune system by xenobiotics. Environ. Health Perspect. 81:157-162.

McBride, W.G. 1961. Thalidomide and congenital anomalies. Lancet 2:1358.

McClearn, G.F. 1988. Strategies for biomarker research: experimental and methodological design. Exp. Gerontol. 23:245-255.

McDonnell, W.F., 3rd, R.S. Chapman, M.W. Leigh, G.L. Strope, and A.M. Collier. 1985. Respiratory responses of vigorously exercising children to 0.12 ppm ozone exposure. Am. Rev. Respir. Dis. 132:875-879.

McLachlan, D.R., T.P. Kruck, W.J. Lukiw, and S.S. Krishnan. 1991. Would decreased aluminum ingestion reduce the incidence of Alzeheimer's disease? Can. Med. Assoc. J. 145:793-804.

Melius, J.M., and P.A. Schulte 1981. Epidemiologic design for field studies: occupational neurotoxicity. Scand. J. Work Environ. Health 7 (Suppl. 4): 34-39.

Munoz, C., K. Garbe, H. Lilienthal, and G. Winneke. 1988. Significance of hippocampal dysfunction in low level lead exposure of rats. Neurotoxicol. Teratol. 10:245-253.

Needleman, H.L. 1986. Studies of psychologic performance of children with elevated dentine lead levels. Pp. 139-140 in D.W. Underhill and E.P. Radford, eds. New and Sensitive Indicators of Health Impacts of Environmental Agents. Third Annual Symposium on Environmental Epidemiology, held April 26-28, 1992, in Pittsburgh, PA. Pittsburgh, PA: Center for Environmental Epidemiology, Graduate School of Public Health, University of Pittsburgh.

Needleman, H.L., A. Leviton, and D. Bellinger. 1982. Lead associated intellectual deficit [letter]. N. Engl. J. Med. 306:367.

Needleman, H.L., C. Gunnoe, A. Leviton, R. Reed, H. Peresie, C. Maher, and P. Barrett. 1979. Deficits in psychologic and classroom performance of children with elevated dentine lead levels. N. Engl. J. Med. 300:689-695.

Needleman, H.L., and C.A. Gatsonis. 1990. Low-level lead exposure and the IQ of children. a meta-analysis of modern studies. J. Am. Med. Assoc. 263:673-678.

Ngim, C.N., and G. Devathasan. 1989. Epidemiologic study on the association between body burden, mercury level, and idiopathic Parkinson's disease. Neuroepidemiology 8:128-141.

NRC (National Research Council). 1989a. Biologic Markers in Reproductive Toxicology. Washington, DC: National Academy Press. 395 pp.

NRC (National Research Council). 1989b. Biologic Markers in Pulmonary Toxicology. Washington, DC: National Academy Press.

NRC (National Research Council.) 1991. Environmental Epidemiology. Vol. 1: Public Health and Hazardous Wastes. Washington, DC: National Academy Press.

NRC (National Research Council.) 1992a. Biologic Markers in Immunotoxicology. Washington, DC: National Academy Press.

NRC (National Research Council). 1992b. Multiple Chemical Sensitivities: Addendum to Biologic Markers in Immunotoxicology. Washington, DC: National Academy Press.

NRC (National Research Council). 1992c. Environmental Neurotoxicology. Washington, DC: National Academy Press.

NRC (National Research Council). 1993. Measuring Lead Exposure in Infants, Children, and Other Sensitive Populations. Washington, DC: National Academy Press

Ostro, B., and S. Rothschild. 1989. Air pollution and acute respiratory morbidity: an observational study of multiple pollutants. Environ. Res. 50:238-247.

Ostro, B.D., M.J. Lipsett, M.B. Wiener, and J.C. Selner. 1991. Asthmatic responses to airborne acid aerosols. Am. J. Pub. Health 81:694-702.

Otto, D.A., H.K. Hudnell, D.E. House, L. Mohave, and W. Counts. 1992. Exposure of humans to a volatile organic mixture. I. Behavioral assessment. Arch. Environ. Health 47:23-30.

Ozkaynak, H., J. Spengler, and A. Garzd, et al. 1986. Assessment of population health risks resulting from exposure to airborne particles. In S. D. Lee, ed. Aerosols: Research, Risk Assessment, and Control Strategies. Chelsea, MI: Lewis Publishers.

Pope, C.A. 1989. Respiratory disease associated with community air pollution and a steel mill, Utah Valley. Am. J. Pub. Health 79:623-628.

Pope, C.A. 1991. Respiratory hospital admissions associated with PM_{10} pollution in Utah, Salt Lake, and Cache valleys. Arch. Environ. Health 46:90-97.

Pope, C.A., and D.W. Dockery. 1992. Acute health effects of PM_{10} pollution on symptomatic and asymptomatic children. Am. Rev. Respir. Dis. 145:1123-1128.

Pope, C.A., D.W. Dockery, J.D. Spengler, and M.E. Raizenne. 1991. Respiratory health and PM10 pollution. A daily time series analysis. Am. Rev. Respir. Dis. 144:668-674.

Reff, M.F. , and E.L. Schneider. 1982. Biological markers of aging. NIH Pub. 82-2221. Washington, DC: US Government Printing Office.

Rice, D.C., and R. F. Wiles. 1979. Neonatal low level exposure in monkeys (macac fascicularis): effect on two-choice non-spatial form discrimination. J. Environ. Pathol. Toxicol. 2:1195-1203.

Rubin, R.J. 1987. Biological indices of enzyme induction as markers of hepatic alterations. Pp. 127-136 in Occupational and Environmental Chemical Hazards: V. Foa, Cellular and Biochemical Indices for Monitoring Toxicity. E.A. Emmett, M. Maroni, and A. Colombi, eds. Chichester: Ellis Horwood.

Schrader, S.M., J.M. Ratcliffe, T.W. Turner, and R.W. Hornung. 1987. The use of new field methods of semen analysis in the study of occupational hazards to reproduction: the example of ethylene dibromide. J. Occup. Med. 29:963-966.

Schrader, S.M., R.E. Chapin, E.D. Clegg, R.O. Davis, J.L. Fourcroy, D.F. Katz, S.A. Rothmann, G. Toth, T.W. Turner, and M. Zinaman. 1992. Laboratory Methods for Assessing Human Semen in Epidemiologic Studies: A Consensus Report. Reproductive Toxicology

Schroeder, S.R., B. Hawk, D.A. Otto, P. Mushak, and R. E. Hicks. 1985. Separating the effects of lead and social factors on IQ. Environ. Res. 91:178-183.

Schulte, P.A. 1991. New opportunities for interdisciplinary cancer communication. Cancer Epid. Biomarkers Prevention 1:3-4.

Schwartz, J. 1989. Lung function and chronic exposure to air pollution: a cross-sectional analysis of NHANES II. 1989. Environ. Res. 50:309-321.

Schwartz, J. 1991. Particulate air pollution and daily mortality in Detroit. Environ. Res. 56:204-213.

Schwartz, J. 1992. Low level health effects of lead: growth, developmental, and neurological disturbances. Pp. 233-242 in H.L. Needleman, ed. Human Lead Exposure. Boca Raton: CRC Press.

Schwartz, J., and A. Marcus. 1990. Mortality and air pollution in London: a time series analysis. Am. J. Epidemiol. 131:185-194.

Schwartz, J., H. Pitcher, R. Levin, B. Ostro, and A.L. Nicholas. 1985. Cost and Benefits of Reducing Lead in Gasoline: Final Regulatory Impact Analysis. Washington, DC: Office of Policy Analysis, US Environmental Protection Agency.

Schwartz, J., D.W. Dockery, J.H. Ware, et al. 1989. Acute effects of acid aerosols on respiratory symptiom reporting in children. Air Pollut. Control Assoc. Preprint 89-92.1. Effects of inhalable particles on respiratory health of children. Am. Rev. Respir. Dis. 139:587-594.

Schwartz, J., D. Slader, T. V. Larson, W. E. Pierson, and J. J. Koenig. 1993. Particulate air pollution and hospital emergency risks for asthma in Seattle. Am. Rev. Respir. Dis. 147:826-831.

Spektor, D.M., M. Lippmann, P.J. Lioy, G.D. Thurston, K. Citak, D.J. James, N. Bock, F.E. Speizer, and C. Hayes. 1988. Effects of ambient ozone on respiratory function in active, normal children. Am. Rev. Respir. Dis. 137:313-320.

Spektor, D.M., V.A. Hofmeister, P. Artaxo, J.A.P. Brague, F. Echelar, D.P. Nogueira, C. Hayes, G.D. Thurston, and M. Lippmann. 1991. Effects of heavy industrial pollution on respiratory function in the children of Cubatao, Brazil: a preliminary report. Environ. Health Perspect. 94:51-54.

Stejskal, V. D., R.G. Olin, and M. Forsbeck. 1986. The lymphocyte transformation test for diagnosis of drug-induced occupational allergy. J. Allergy Clin. Immunol. 77:411-426.

Sunyer, J., J.M. Anto, C. Murillo, and M. Saez. 1991. Effects of urban air pollution on emergency room admissions for chronic obstructive pulmonary disease. Am. J. Epidemiol. 134:277-286.

Swan, S.H., R.R. Neutra, M. Wrensch, I. Hertz-Picciotto, G.C. Windham, L. Fenster, D.M. Epstein, and M. Deane. 1992. Is drinking water related to spontaneous abortion? Reviewing the evidence from the California Department of Health Services studies. Epidemiology 3:83-93.

Tamburro, C.H., and G.M. Liss. 1986. Tests for hepatotoxicity: usefulness in screening workers. J. Occup. Med. 28:1034-1044.

Tanner, C.M., B. Chen, W.Z. Wang, M.L. Peng, Z.L. Liu, X.L. Liang, L.C. Kao, D.W. Gilley, and B.S. Schoenberg. 1987. Environmental factors in the etiology of Parkinson's disease. Can. J. Neurol. Sci. 14 (Supp.3):419-423.

Tilson, H.A., and C.L. Mitchell. 1983. Neurotoxicants and adaptive responses of the nervous system: introductory remarks. Fed. Proc. 42:3189-3190.

Tilson, H.A., and C.L. Mitchell. 1992. Neurotoxicology. New York: Raven Press.

Tilson, H.A., J.L. Jacobson, and W.J. Rogan. 1990. Polychlorinated biphenyls and the developing nervous system: cross-species comparisons. Neurotoxicol. Teratol. 12:239-248.

Valciukas, J.A., and R. Lilis. 1980. Psychometric techniques in environmental research. Environ. Res. 21:275-297.

Van Thiel, D.H. 1986. New approaches for identifying hepatotoxins. Pp. 87-98 in D.W. Underhill and E.P. Radford, eds. New and Sensitive Indicators of Health Impacts of Environmental Agents. Pittsburgh, PA: University of Pittsburgh.

Whorton, D., R.M. Krauss, S. Marshall, and T.H. Milby. 1977. Infertility in male pesticide workers. Lancet 2:1259-1261.

Wichmann, H.E., W. Mueller, P. Allhoff, M. Beckmann, N. Bocter, M.J. Csicsaky, M. Jung, B. Molik, and G. Schoeneberg. 1989. Health effects during a smog episode in West Germany in 1985. Environ. Health Perspect. 79:89-99.

Williams, J.R., P.S. Spencer, S.M. Stahl, J.F. Borzelleca, W. Nichols, E. Pfitzer, E.J. Yunis, R. Carchman, J.W. Opishinski, and R.A. Walford. 1987. Interactions of aging and environmental agents: the toxicological perspective. Prog. Clin. Biol. Res. 228:81-135.

Wright, D.M., J.S. Kessner, S.M. Schrader, N.W. Chin, V.E. Wells, and E.P. Krieg. 1992. Methods to monitor menstrual function in field studies: attitudes of working women. Reprod. Toxicol. 6.

Xintaras, C., J.R. Burg, B.L. Johnson, S. Tanaka, S.T. Lee, and J.H. Bender. 1979. Neurotoxic effects of exposed chemical workers. Ann. NY Acad. Sci. 329:30-38.

Yule, W., R. Lansdown, I. Millar, and M. Urbanowicz. 1981. The relationship between blood lead concentration, intelligence, and attainment in a school population: a pilot study. Dev. Med. Child. Neurol. 23:567

Zeiss, C.R., R. Patterson, J.J. Rouzansky, M.M. Miller, M. Rosenberg, and D. Levitz. 1977. Trimellitic anhydride-induced airway syndromes: clinical and immunologic studies. J. Allergy Clin. Immunol. 60:96-103.

Zierler, S. 1992. Drinking water and reproductive health. Epidemiology 3:77-78.

5

Data Systems and Opportunities for Advances

THIS CHAPTER DESCRIBES SOME existing data systems that permit the characterization of personal exposure and health status. Given the scarcity of resources for studies in environmental epidemiology, researchers need to make the best use of existing data. It is beyond the scope of this chapter to cover all the pertinent data systems or to describe systems in detail. Rather, the focus is on classes of data-collection systems, some of the major systems in each class and their important features, and their use. The emphasis will be on data systems that are publicly available (often from the federal government). For a more-comprehensive list of federal data systems related to environmental exposure, see EPA et al. (1992); for other discussions of state and local data systems see, for instance, Health Officers Association of California (1986), National Governors' Association (1989), Frisch et al. (1990), and Sexton et al. (1992, 1994). There is a need for greater dissemination of the knowledge of the existence and availability of federal, state, and local systems. Many of these are limited in size, coverage, end points, completeness, or accuracy, but where they meet the investigator's needs, they can save much time and expense. A geographic information system can be very useful in investigation by providing an organizing framework for data on exposure and outcomes.

INTRODUCTION

The interest of the American public in environmental pollution seems to be driven primarily by concerns about health. People ask, "Have we

been exposed?" "Have we been affected?" "Will we be affected later?" They might well ask, also, "Do our management programs have any effect on the health of the public?" Table 5-1 outlines some epidemiologic research strategies that address these concerns. It shows that many types of epidemiologic studies and data can be used to determine the relation between the environment and human health. Although experimental studies of animals and laboratory studies of humans do provide some answers to these questions, epidemiologic research is essential to their resolution. Often, however, epidemiologic studies are neither available nor possible, and policy must be based on toxicologic evidence and animal studies.

Considerations of cost, urgency, and limited special expertise often require that officials rely on analyses of existing data that were gathered for other purposes. Epidemiologic studies of the classical kind involve the measurement of both the health status and the environmental exposure (or internal dose) of the persons being studied. However, such measurements cannot always be obtained. For instance, if historical exposures were not measured, the investigator may have to estimate them from other, less reliable, information. On the other hand, the exposure might be so extensive that no suitable control population remains. Individualized measures of exposure and health can also be infeasible or too expensive when a health effect occurs so infrequently that adequate study would require that a large number of exposed persons be evaluated in detail; such problems require other epidemiologic methods or the use of secondary data.

During development and implementation of public-health policy, analyses of secondary data are important at several stages. Intense study of selected small groups of people can provide useful information about risk that identifies a need for public policy. To determine the extent of potential exposure, the size and characteristics of the population exposed, or the background frequency of the health effect of interest, secondary-data analyses are useful. Information from existing data systems is useful in program planning and development when data are needed to validate findings of earlier targeted research studies.

During implementation, data from existing systems can provide additional insights for public-health policy. Existing data systems tend to reflect the programmatic and regulatory structure of government programs, so the identification of useful systems (or of their absence) might help to define the most appropriate needs for assessment and availability of the public-health response. This in turn allows for midcourse changes to reduce costs, improve response, or otherwise improve on-going programs.

Data systems are a primary mechanism for evaluating the impact of a public-health policy. For instance, a public-health program might target

TABLE 5-1 Issues of Major Concern to the Public and Methodologic Responses

	Methodologic Responses		
	Exposure Assessment	Applied (Response) Epidemiology[a]	Epidemiologic Study[b]
Are we exposed?	X		
Are we affected now?		X	
Did exposure cause a health effect?	X	X	X
Will we be affected later?	X		X
Did we improve health with a program initiative?			X

[a]Applied, or response, epidemiology refers to studies designed as a quick response to concerns expressed by a group of individuals regarding the potential for exposure or health effects. These are the basis of much of the "gray" literature and many of the studies performed by public-health agencies.

[b]Epidemiologic studies include classical case-control and cohort studies of targeted populations, in contrast with studies of the general population.

ozone because of its effects on several pulmonary health end points. However, the success of the program might be evaluated solely from the ambient concentrations of ozone in a polluted area. To evaluate the impact on public health, it is important to know the relation between the observed ambient concentrations of ozone and the frequency of various pulmonary health end points. Modification of public-health policies depends on knowledge of such relations, identified largely through analyses of secondary data.

DATA-COLLECTION SYSTEMS: WHAT THEY MEASURE

Evaluation of the relation between an environmental pollutant and human health requires data to characterize exposures to the pollutant, including concentrations in the environment, the probability and characteristics of human exposure, and the distributions of internal doses, as well as trends or differences in the health status of exposed people. Determination of risk-management alternatives requires, in addition, information on the sources and distribution of the pollutant. Data systems may address each of these needs. However, they have not necessarily been established with the goal of integration with other classes of data, and

Registries and Surveillance	Reference Surveys for Exposure	Reference Surveys for Health Effect	Risk Assessment
	X		
X		X	
X	X	X	X
X	X		X
	X		

most data-collection systems collect only one kind of data or data on one aspect of the general problem.

One distinguishing characteristic of a data system is how and when the responding units are sampled. For example, persons may be selected at random from a defined population but tested at a fixed time (8:00 am every day) or once at a haphazard time (when the laboratory is not otherwise busy). Some surveys are designed to obtain probability samples that accurately represent a reference group, such as a population or an occupational setting, but others obtain samples by convenience, such as collections of information from participating states or hospitals. A short-term survey may not be representative across time. A survey system might select sampling units that are characteristic or representative of larger reference groups. Characteristic sampling is based on selection from a list of strata; representative sampling is based on the distribution of strata in the population. For instance, in selecting monitoring sites, one might decide that several important types of environments should be evaluated. Monitoring sites can be selected to characterize those types of environments, as in the stratified sampling of air in urban areas and rural areas. Alternatively, one might select monitoring sites on the basis of a stratified probability-sampling scheme to yield data that are representative of the distribution of environments. Monitoring is expensive, and decisions about where to put monitors are generally considered carefully, but the decisions may not be optimal for a specific environmental-epidemiology study. For example, if budgets allow for only a few monitors to measure some chemical, should they be placed to obtain the most-representative

geographic coverage? In or near population centers? Where prior information suggests the levels are highest? Other?

A data-collection system can be either a compendium or a systematic survey. That is, it can consist of individual studies with similar but separate research designs and measurements, or it can collect data from many sources in a standardized fashion. Neither system is necessarily identical between study years or cycles. That is, the pollutant or health effect assessed by a systematic survey and how it is assessed may vary from time to time, from place to place, or in other ways.

Data systems with the characteristics mentioned above are useful for evaluating the relation between environment and health. The usefulness of any data system is limited by its characteristics, so it is important to understand the sampling and assessment characteristics of each data source before using it. (See the discussion below on bridging environmental and health issues.)

SOURCE OF POLLUTANT

The development of systems to collect information about discharges of pollutants (apart from occupational exposures) is a primary objective of the Environmental Protection Agency (EPA) (table 5-2). Its role as the principal environmental-risk management agency in the federal government requires data on the relative contributions of sources and on control options. The primary objective of EPA's data systems is to provide information pertinent to regulation, so they are designed to be comprehensive with regard to polluters and pollutants that have been identified as toxic. Many pollution-related data systems have emphasized the characterization of pollutant sources, rather than the distribution and fate of pollutants in the environment or the potential exposures of humans. Representative data (as opposed to comprehensive data) have little utility in assessing compliance with regulation of individual pollution sources, though such data can be useful in assessing needs for and monitoring the success of management programs.

Other data-collection systems characterize the amounts of a pollutant at its source. These include production volumes and emission inventories. These systems, too, are not directly concerned with the fate of pollutants in the environment. Data systems that contain location- and time-specific information can be used in analytic models to estimate the transport and fate of pollutants in the environment. However, few data systems contain both time-integrated information (for instance, yearly, periodic, or daily data on emissions) and geographic information (for instance, production volume at a worksite).

POLLUTANT CONCENTRATIONS IN THE ENVIRONMENT

The locations covered by most pollutant-concentration data systems are chosen to be characteristic rather than representative (table 5-3). Thus, most of the National Air Monitoring Stations or water-system quality sites of the National Stream Quality Accounting Network are in densely populated areas. These data systems contain detailed information on the location of the monitoring site, and samples are collected frequently enough to represent short periods. However, site selection is not based on detailed information about the population, the area, or the distribution of exposures among individuals, and the positioning of a station does not necessarily reflect the most likely route of human exposure. For instance, some air-monitoring stations are on the tops of buildings, and water-quality assessments are performed at the outflow pipes of water-treatment facilities, not at residential taps. Those locations might yield informative data on relative exposures, but may not represent either the distribution of concentrations in the environment or the actual exposures of people.

Pollutant-concentration data systems are probably underused for ecologic studies. These systems contain detailed geographic data, and, although few pollutants may be assessed, the analytic methods tend to be relatively stable over time, and exposure is generally measured at or integrated over short intervals.

Although most data on pollutant concentration are from monitoring systems, data from "response epidemiologic studies" are increasing. Response (or applied) epidemiologic studies are designed to respond quickly to expressed concerns regarding the potential for exposure or adverse health effects. Examples of response epidemiologic programs are the health-assessment studies of the Agency for Toxic Substances and Disease Registry (ATSDR) and the health-hazard evaluations of the National Institute for Occupational Safety and Health (NIOSH). In these studies, environmental concentrations of various pollutants are regularly assessed. However, study sites are often selected because the potential exposure is considered high or because of complaints about symptoms, so sites are not characteristic of ordinary population exposures. These studies do, however, attempt to characterize explicitly the scope or potential for human exposure in these presumably extreme settings, and they may contribute information on the relation between the environmental distribution of pollutants and human exposure or internal dose.

HUMAN EXPOSURE

Data on human exposure (table 5-4) are the least developed of the classes considered here, and generalizations to larger groups of people or

TABLE 5-2 Data-Collection Systems: Source of Pollutant

Data-System Name	Description
Production Volume Inventories	
Synthetic organic chemicals	Annual data on production and sales of synthetic organic chemicals produced in the United States
Site Inventories	
National pollutant-discharge elimination system	Permits for worksites that specify effluent concentration limits, monitoring, and reporting requirements
National Priorities List	List of the toxic-waste sites determined to be of immediate concern for remediation
Emission Inventories	
Toxic chemical release inventory	Annual estimates of releases from manufacturing facilities of minimal size and volume of chemicals per year
Integrated database	Information on spent fuel and radioactive-waste inventories for nuclear reactors, storage facilities, and mine tailings, among others
Sales Volumes	
Agricultural chemical use	Database of information on sales for agricultural purposes of fertilizers and pesticides, among others

particularizations to specific exposure situations are often difficult and uncertain. Much detail is required to make this class of data useful, but few detailed data systems have been developed. Detailed information on human exposures generally requires the use of personal monitors or structured activity questionnaires, but these tools are expensive and time-consuming. Thus, most systems contain information on small populations chosen to be characteristic, but not necessarily representative, of the target population. However, systems that exist generally have substantial extent and detail over periods as long as several years.

More data of this class could be gathered by brief structured activity questionnaires in large surveys. Brief questionnaires might provide less-

Primary Objective	Coverage/Sample + Design	Linking Data
Monitoring	National totals: comprehensive	None
Regulatory	National: comprehensive	Detailed geographic codes, river reach no., pollutant limits
Regulatory	National: based on reports from regions, comprehensive	Detailed geographic codes, environmental concentrations
Informational	National: comprehensive for defined worksites	Detailed geographic codes
Informational	National: comprehensive for defined sites	Facility name
Monitoring	National: characteristic farm sample	None

detailed information on human exposure patterns than personal monitors, but for a fixed total budget they can yield data on greater numbers of people. A combination of brief questionnaires for large numbers of people with validation and characterization of a subset using personal monitors might even be more useful.

INTERNAL DOSE

Like information on human exposure, information on internal dose is rarely collected systematically (table 5-5). Occasional studies of biologic markers of specific agents in small, defined populations are plentiful, but

TABLE 5-3 Data-Collection Systems: Environmental Concentrations

Data-System Name	Description
Monitoring Systems	
Aerometric Information Retrieval System	Ambient concentrations, emissions, and compliance data for airborne criteria pollutants
Microbiology and residue computer information system	Contaminant data from samples of meat and poultry at slaughtering establishments and from import shipments
Regulatory Systems	
Permit-compliance system	Information for tracking the permit, compliance, and enforcement of permittees under the Clean Water Act
Response Epidemiologic Studies	
Health assessments (ATSDR)	ATSDR assessments to identify potential health concerns among populations living near National Priority List sites
Microenvironment Settings	
Indoor air study	A pilot project to assess contaminants in indoor air

broad and systematic collections of data on biologic markers in the general population are few, and surveys have yielded little information with which to characterize the subjects' exposures. Direct measures of internal dose are not usually included in health-assessment studies (conducted by ATSDR) or health-hazard evaluations (conducted by NIOSH), but these sources could be modified to include internal-dose assessments.

ATSDR conducts public-health assessments to determine where, and for whom, public-health actions should be undertaken (ATSDR, 1992). Each assessment characterizes the nature and extent of hazards and identifies communities where public-health actions are needed. However, the assessment is largely or entirely a compilation and analysis of existing data, which rarely include internal doses of toxicants in the population of concern. The health-assessment format does not require the collection of

Primary Objective	Coverage/Sample + Design	Linking Data
Monitoring	National: air-monitoring stations in urban areas	Detailed geographic codes/point-source identifiers
Monitoring	National: Random sampling of meat products	No information on distribution of food
Regulatory, monitoring	National: comprehensive coverage of permittees	Detailed geographic codes, linked to Reach Pollutant Assessment System
Regulatory	National: all National Priority List sites	Detailed geographic codes, linked to environmental concentration data
Research	Selected sites: not sampled to be representative	None

new data, for at least 2 reasons. First, the objective of the ATSDR health-assessment study is to determine whether there is a potential for human health effects, not to determine the extent or magnitude of actual exposure. Second, many internal-dose assessments are invasive; this decreases participation rates and increases opportunities for bias. However, when a health assessment indicates a potentially significant risk to human health, ATSDR is obliged under the 1986 Superfund Amendments and Reauthorization Act (SARA) to consider a registry as a followup (ATSDR, 1988a), and registrants may be invited to participate in biologic testing for markers of exposure or effect (NRC, 1989).

There is also a need for studies that characterize a population with a well-defined sampling scheme. The National Health and Nutrition Examination Survey (NHANES), conducted by the National Center for

TABLE 5-4 Data-Collection Systems: Human Exposure

Data-System Name	Description
Time-Activity Patterns and Personal Monitoring	
Total-exposure-assessment methodology	Goals to develop methods to measure individual total exposure to toxic and carcinogenic chemicals
Surveys	
National Occupational Exposure Survey	Information on the probability of exposure to various chemicals based on job title
Registries	
National Exposure Registry	Identification of individuals with verified exposure to selected chemical, with followup studies to be performed on individuals in registry

Health Statistics, studies about 30,000 persons in the US population, chosen by random sampling (clustered, stratified, with deliberate oversampling of some subgroups). For study of general contaminants, such as lead or petrochemical oxidants, NHANES has been used as a data source. Given the followup capabilities of NHANES, detailed exposure data could be collected in subgroups of the entire sample that are identified as having received internal doses of particular interest. However, when the probability of exposure is small, the actual number of participants who could be studied to characterize the specific exposure would be small, possibly zero, and NHANES might not be sensitive enough. Specially designed surveys could be considered to characterize specific population exposures.

HEALTH STATUS

Most health-status information systems are not developed for the primary purpose of studying environmental health (table 5-6). Vital records are collected for legal reasons, hospital-discharge and cost information [e.g., Medicare provider analysis and review (MEDPAR)] is collected for economic or administrative reasons, and the National Health Interview

Primary Objective	Coverage/Sample + Design	Linking Data
Research	Selected sites: characteristic of urban populations	None
Research	National: characteristic of selected industries	Job-title codes
Research	National: selection based on reports, not probability sampling	Job titles, personal histories, residential geographic codes

Survey (NHIS) and NHANES are conducted for general US population health-monitoring reasons. Several other registries and surveillance systems are maintained to identify important risk factors, including environmental exposures, but in a general reporting system the amount of information that can be collected for each exposure of interest is limited. The information in such programs as the Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute, which collects cancer incidence and survival data from approximately 10% of the US population, and the Birth Defects Monitoring Program (BDMP) of the National Center for Environmental Health is gathered by abstractors trained in coding and abstracting from hospital records. The National Exposure Registry being developed by ATSDR will be an exception to the absence of environmentally focused health-status data systems (see "ATSDR Exposure Registries" below).

The range and quality of exposure data from surveys could sometimes be expanded by data obtained as an addition to routine followup but at additional cost. NHANES, NHIS, SEER, and the US national vital-statistics system all have followup capabilities. Some kinds of information collected in a followup survey could be only qualitative, such as whether a person was exposed or not, and others will be subject to problems of

TABLE 5-5 Data-Collection Systems: Internal Dose

Data-System Name	Description
National Health and Nutrition Examination Survey	An examination survey of a probability sample of the US population, with some toxic-chemical concentrations measured in blood samples
National Human Adipose Tissue Survey	Concentrations of various toxic chemicals in adipose-tissue samples collected from autopsied cadavers and surgical patients

poor memory and recall bias. Because of the mobile nature of the US population, collection of blood samples or examination data in a followup survey would require a mobile unit that would be used to assess only a few people in each location or would require transporting subjects to a central location. Tissue samples would be even more difficult to collect. These problems are magnified as the population size, geographic range, and length of followup period increase. Collecting good information is expensive.

Comprehensive collections of data on vital events, especially birth and death, are highly accurate and nearly complete because of legal requirements to document these events. Collecting data to adequate quality standards is harder for health characteristics that change from year to year or even from day to day (such as diseases, symptoms, and use of health services) or that are thought not severe enough to warrant professional care (such as symptoms or non-life-threatening diseases). For these, accurate and comprehensive surveillance of the US population is impossible. Many of the large federal surveys collect data on representative, rather than characteristic, samples of the US population, so even large geographic areas, such as states, may not provide reliable estimates of health status. Some systems, such as the Behavioral Risk Factor Surveillance System (of the National Center for Chronic Disease Prevention and Health Promotion), could be used to collect exposure and health-status information at the state level, but are limited to information that could be reliably solicited by the telephone interview method used.

Untapped resources that could provide systematic information on the health status of populations include routine medical examinations and school test-performance scores. However, when these data resources are not collected under common standards—as, for instance, a routine medical examination might be—the consistency of data may be poor.

Primary Objective	Coverage/Sample + Design	Linking Data
Monitoring	National: representative of civilian, noninstitutionalized	Residential geographic codes, personal histories, current job
Monitoring	National: Characteristic of urban populations	Geographic codes

BRIDGING ENVIRONMENT AND HEALTH

As noted, few data systems include information on both exposure and health. Even in those with both—such as NHANES, BDMP, or SEER—the focus is generally on health, and information on exposures is minimal. Health-status data systems may have self-reported information on only a few items related to possible exposures, such as occupation. Of the 3 surveys mentioned here, only NHANES has information on biologic markers based on blood or urine samples.

UNDIMENSIONAL STUDIES

Maps of cancer mortality have been used to infer environmental or occupational exposures (Fraumeni, 1987; Pickle et al., 1987) and to identify populations at high risk for specific exposures (Rothenberg et al., 1990; CDC, 1990). Clusters or trends of diseases have been used to identify populations that might have had unknown toxic exposures (Deane et al., 1989; Edmonds and James, 1990; Anto et al., 1989) and that could then be studied more intensively. Conversely, emissions data from the Toxic Release Inventory, data on ambient concentrations of various pollutants (EPA, 1991), and lists of toxic-waste sites (Commission for Racial Justice, United Church of Christ, 1987) have been used to identify populations of concern for high exposures. In those descriptive studies, information on population exposure is usually inferential and based on proximity to a source of pollution.

Because chance alone will create clusters (Rothman, 1990; Wagener, 1990), the observation of a cluster—even one that is quite striking—often does not signal an increased probability of illness because of some risk factor. Rather, the cluster is a rare, but predictable, event in a population

TABLE 5-6 Data-Collection Systems: Health Status

Data-System Name	Description
Vital Records	
National Vital Statistics Program	Data from birth, death, and marriage records
National Surveys	
National Health and Nutrition Examination Survey	Survey of a probability sample of the US population that includes interview, examination, and physiologic testing
National Health Interview Survey	Interview survey of a probability sample of the US population that includes rotating special topics, including knowledge of risk factors, such as radon, and occupational chemical exposures
Surveillance Systems	
Birth Defects Monitoring Program	Information sent by participating hospitals on birth defects diagnosed and recorded in the newborn period
Surveillance, Epidemiologic, and End Results Program	Demographic and diagnostic information on patients identified as having some form of cancer
Response Epidemiologic Studies	
Epidemiologic investigations	Centers for Disease Control and ATSDR; these studies in response to public concerns of potential exposures and health effects often include observational data on health status

that is not experiencing a change in the probability of an illness. The 1-in-1,000 event will occur, by definition, one time in 1,000, and if many thousands of clusters could be defined (by time interval, geographic location, specific health end point, etc.), then the observation of even multiple clusters may have little general meaning. Because of the great number of ways a population and an outcome can be subdivided, one is nearly cer-

Primary Objective	Coverage/Sample + Design	Linking Data
Monitoring	National: comprehensive	Residential geographic codes, job title
Monitoring	National: represents noninstitutionalized population	Residential geographic codes, personal history, job title
Monitoring	National: represents noninstitutionalized population	Residential geographic codes, job title
Monitoring	National: covers only participating hospitals	None
Monitoring	Participating geographic areas	Broad geographic codes; fine detail available in special studies
Health-hazard detection	National: compendium, based on reports	None

tain to find that *some* disease is more common in *some* segment than in others, with a low p value.

LINKING DATA SYSTEMS

Linking several data systems can improve information on both exposure and health.

Studies with Information on Environmental Concentrations, Health Status, and Nonenvironmental Risk Factors:

An environmental exposure is rarely the only potentially important determinant of risk, so a useful study must almost always include data on various other risk factors, such as certain behaviors, as well as the exposure and health status of the subjects. Such studies would involve linking data that could be used to calculate potential exposure to an environmental factor with subject-specific information on health status and behavioral risk factors. The following studies illustrate methods that could be used more widely.

Schwartz (1989) used health data from NHANES II and air-pollution measurements from the Storage and Retrieval of Aerometric Data (SAROAD) system, now referred to as the Aerometric Information Retrieval System, to evaluate the relation between lung function and chronic air pollution. Data on individual subjects were paired with air-pollution measurements based on the census tracts of the subjects' residences and on the locations of monitoring stations within 10 miles. Lung function varies with sex, age, and body size, estimated in this case by height and body-mass index. Smoking and respiratory conditions also have major effects on lung function. Information on all these risk factors was used in a multivariate analysis.

Ostro and Rothschild (1989) linked self-reported information on acute respiratory infections assessed in the NHIS to 2-week average air-pollution data based on metropolitan statistical areas from SAROAD. Information on other variables such as age, sex, race, smoking status, and chronic conditions was incorporated.

While outdoor monitoring provides only imprecise measures of personal exposure to air pollution, important relations may be identified if data from the monitoring stations are sufficiently correlated with exposure of the subjects. Relations might then be studied in greater detail in special studies. Broad correlations also provide indications of the impact of changes in the environment, as measured at monitoring stations, on changes in health status.

Studies with Information on Environmental Concentrations and Health Status, but Not Other Risk Factors:

Many studies that have examined correlations between measurements and health assessments have failed to adjust for other risk factors, environmental or otherwise, often because relevant data do not exist.

Linkage with mortality data is common and generally straightforward because of the virtually complete ascertainment of deaths and the use of

highly developed, standardized coding systems. Air-pollution data have been linked to mortality data on a local basis in numerous cross-sectional studies (Boyd, 1960; Buck and Wicken, 1967; Glasser and Greenburg, 1971; Lave and Seskin, 1973). Time-series analyses may require control for other variables, such as seasonality and autocorrelation (Ozkaynak et al., 1986; Mazumdar et al., 1982; Schwartz and Marcus, 1990). These approaches are ecologic; that is, a measure of the distribution of pollutant concentrations in an area is correlated with a measure of the distribution of health status in the area such as death rates.

An alternative is to study health measures in cohorts of known exposure status, such as mortality in occupational cohorts (Fraser et al., 1982; Wingren et al., 1991). The subjects can be persons at risk of exposure to toxic materials because they live near toxic-waste sites or for other reasons, e.g., inclusion in the National Exposure Registry of ATSDR.

Other assessments of the health of a population are based on hospital admissions, emergency-room visits, and calls for ambulances. Pope (1991) used regression methods to study the association between hospital admissions for respiratory conditions and measurements of PM_{10} in local areas, including control for temperatures based on month of admission. Although characteristics of individual subjects were not available, Pope noted strong associations between indicators of respiratory health, particulate pollution, and the operation of a nearby steel mill.

Studies with Information on Health Status and Nonconcentration Measures of Environmental Status:

The studies discussed above used ambient-concentration data from monitoring stations mainly as a surrogate for the probability that personal exposures were sufficient to affect health, but models can use other sources of environmental data. For instance, Frank et al. (1986) used the NHIS to evaluate the relation between chronic cardiovascular illness and exposure to carbon monoxide in the workplace. Information on occupational exposure from the National Occupational Hazards Survey was used to estimate the probability that people in specified jobs were exposed to carbon monoxide. Thus, the exposure data did not include direct measures. Information for individual subjects was linked to exposure probabilities on the basis of current occupation and humidity, based on county of residence. Additional risk factors evaluated in multivariate analyses included age, obesity, sex, demographic variables, and smoking status. Because information on health status is predicated on a subject's reporting that a medical professional had diagnosed a condition, Frank et al. incorporated economic variables—such as availability of health care and ability to afford health care—into the analyses.

The dichotomous classification of persons as potentially exposed or not is occasionally informative, as in the case of occupational exposures. However, without information on the extent of exposure, any health effects are likely to be underestimated. For instance, Lynch et al. (1989) demonstrated that misclassification of chlorination exposure could obscure a possible relation between exposure and risk of urinary-bladder cancer.

Other ambient measures of environmental status can also be informative. During NHANES II, Mahaffey et al. (1982) determined blood lead concentrations. Further analyses indicated that average exposure levels changed during the course of the study (1976 through 1980). Additional analyses quantified the relation between various sources of lead and changes in blood lead concentrations, especially in children (EPA, 1985; ATSDR, 1988b). Exposures to leaded paint, dietary lead, and leaded gasoline were considered, and only the change in total lead used in gasoline production was correlated with the change in blood lead (Annest et al., 1983). The correlation was particularly strong among the youngest children.

MONITORING OF ENVIRONMENTAL HEALTH EFFECTS

Monitoring is the continuing and systematic collection, analysis, and interpretation of data. If monitoring is linked to specific programs designed to prevent or control health outcomes, the activity is better termed surveillance. This section focuses on monitoring of health status.

Many of the causes of chronic diseases are unknown, though it is clear that multiple factors affect many specific health end points, including workplace, diet, place of residence, recreational activities, and ancestry. Sorting out the relative importance of causes of disease in humans remains daunting, given the multitude of exposures and uncontrollable factors that can affect health. It is suspected that some noncommunicable diseases (e.g., some cancers and diabetes) are increasing in frequency because of unknown factors in the environment. It is desirable, therefore, to monitor the occurrence of such diseases to provide a kind of early-warning system that would enable causes to be identified more readily than in the past.

Such monitoring systems must be designed in light of the problems of determining the contribution of environmental factors when some, but not all, causes are known. The problems include

- Difficulties in exposure identification and estimation.
- Followup and latency (time from exposure to the appearance of disease).
- Long-term nature of chronic diseases and the repeated use of health-care resources, such as laboratories and hospitals.

- Size of the affected population.
- Variable and imprecise symptomatology of some conditions.

To address those problems, monitoring systems have to be

- Large, i.e., cover a substantial population.
- Of long duration.
- Capable of producing information that can be combined with similar data systems, so as to increase sample sizes; this requires the collection of data in a standardized fashion.
- Capable of being linked to other data sources, e.g., exposure data, which would require personal identifiers and provisions for confidentiality.

A set of sentinel health events or health-status indicators needs to be identified for use in exposure, internal-dose, and health-status data systems. Lists for this purpose have usually focused on disease or syndrome end points (Rothwell et al., 1991; DHHS, 1991). For each health end point to be monitored, it is necessary to establish precise case definitions and gather baseline information on incidence and prevalence. The problems of definition and baseline determination are more complex if symptoms, rather than diagnosable disease, are considered.

Some monitoring systems for special purposes must be set up de novo, as has been the case for many cancer registries, but others can rely at least in part on existing data systems. Options include

- Special surveys (not designed for long-term monitoring, although surveys can be repeated periodically).
- Disease-reporting systems focused on incidence (as exist for several infectious diseases).
- Capture-mark-recapture systems (see below).
- Projects that link existing disease and exposure registries.
- Special surveillance mechanisms based on health-maintenance organizations or health-insurance systems, such as Medicare.
- Special record-linkage systems, such as that pioneered in Oxford, England, or under development in Manitoba, Canada. The Manitoba record-linkage system will evaluate the extent to which census data can be linked to the provincial health-insurance scheme to provide health information. This system, built on pioneering work by Roos et al. (1979), required a special agreement between Statistics Canada and the province of Manitoba. Strict conditions of confidentiality will be observed, and no individually identifiable information will be released (National Task Force on Health Information, 1991).

Several research methods that have been or could be used in special environmental-epidemiology studies could be extended to monitoring systems.

FOLLOWUP (COHORT) STUDIES

A common example of a followup study is the study of health effects among individuals living near a specific toxic-waste dump. Such studies are difficult and results are often uncertain because the numbers of persons involved is seldom large enough to provide statistically powerful results. As discussed in chapter 2, a hazardous-waste site may contain many different toxic chemicals, and it may be difficult to link a specific chemical to a specific disease. Further, of the large numbers of different chemicals, few have been characterized with respect to their ability to induce a specific disease, so such studies would have to be largely planned as "fishing expeditions" with extensive requirements for data collection and individual surveillance and with multiple end points. The multiple comparisons are then likely to produce some statistically significant associations by chance alone, so interpretation would be difficult.

Epidemiologic research is often expensive and time-consuming, especially where longitudinal studies of large populations are involved, so there is reason to consider "piggy-backing" needed research on other kinds of studies. For example, prospective cohort studies that are not directly related to the environment could possibly be inexpensively modified to collect additional data relevant to many of the objectives of environmental epidemiology. If this were to be done in a coordinated way for several such cohorts, a combined analysis might be informative.

REPEAT CROSS-SECTIONAL SURVEYS

Cross-sectional surveys to determine disease prevalence (or cumulated incidence) could be repeated (e.g., every 5 years) in populations known to have been exposed to environmental contaminants. Although less valuable in some ways than long-term followup of defined cohorts, because those who move away from the area (possibly because of known exposure or illness) would not be followed, they offer an alternative when resources for a large cohort study are not available.

DEATH-CERTIFICATE DIAGNOSES LINKED TO GEOGRAPHIC INFORMATION

Death certificates have been used to determine whether variations in an acute, fatal disease within and between geographic areas are consistent with exposures to an environmental agent. Death certificates are especially useful for diseases with a short course, a single cause, and a high case-fatality rate, such as infectious hepatitis. They are less useful for the identification of environmental causes of chronic and multifactorial disease.

CASE-CONTROL STUDIES

A difficulty with case-control studies is that the assessment of exposure involves extrapolation to the past. Replication of the findings in other settings is usually needed to provide the evidence required to infer causality.

ACTIVE SURVEILLANCE OF EMERGENCY ROOMS AND HOSPITAL ADMISSIONS

Changes in the frequency of emergency visits or hospital admissions for specific conditions can provide information about acute environmental insults. On a long-term basis, hospital admission rates measure the prevalence of serious disease in the community, though a major disadvantage is the inability of nearly all existing systems to distinguish between first and repeat admissions for the same condition. This is a strong argument for retaining personal identifiers in the basic data.

OUTBREAK INVESTIGATIONS

After an apparent cluster of cases is identified, investigators may go into the community to identify potential reasons for the cluster. Such investigations are often necessary to allay public anxiety, but typically they have little scientific value. It has been difficult to find direct links of environmental agents to the risk of disease, as many clusters are a result of chance, rather than an identifiable environmental agent (Rothman, 1990).

Poison centers have identified acute episodes of environmental contamination (e.g., Goldman et al., 1990) This is one type of outbreak investigation, with the problems just discussed, though with the advantage that the symptoms can sometimes be related to exposure to a specific substance.

HEALTH CARE FINANCING ADMINISTRATION DATA

Data are collected on nearly all chargeable episodes of disease in the US population aged 65 years and over by the Health Care Financing Administration. This resource is largely unexplored as a tool for disease-monitoring purposes (as it is exploited in Canada, where the provincial systems cover all ages). A disadvantage in the United States is the restriction to the elderly and the fact that the elderly, especially those of higher socioeconomic status, often move away from the area where they spent most of their lives. The full impact of environmental factors on many

chronic diseases may not be expressed until older ages. If diseases occurring in older members of the population can be linked to identified episodes of past pollution, they could provide information needed to prevent future disease in those who are now young, but such links are hard to discern in a mobile population.

For many chronic diseases, the time of first diagnosis may not be important. Rather, a measure of cumulative incidence (approximated by prevalence in nonfatal conditions) could be just as informative, thus increasing the usefulness of data that may mark the presence of chronic disease but not the date of diagnosis.

ATSDR EXPOSURE REGISTRIES

The Agency for Toxic Substances and Disease Registry was created by Congress by the Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (CERCLA) to address possible public-health effects of environmental exposures to hazardous substances from waste sites and chemical spills (NRC, 1991b). CERCLA requires ATSDR, in cooperation with the states, to establish national registries of persons who have been exposed to hazardous substances and later develop serious disease or illness (ATSDR, 1988a). While disease registries have not yet been established, the National Exposure Registry is further developed and is the focus of this discussion.

The stated purpose of ATSDR's Exposure Registry is "to aid in assessing long-term health consequences of exposure to Superfund-related hazardous substances" (ATSDR, 1988a, p. 7). To facilitate epidemiologic research, ATSDR intends to design and create its data systems for both hypothesis generation (identifying possible adverse health outcomes) and hypothesis testing (of suspected adverse health outcomes). Other goals are to facilitate state and federal health-surveillance programs and to provide information that can be used to assess the effects of an exposure on a population.

ATSDR's data system will contain subregistries created in 4 phases. First, it narrows down potential sites for inclusion to a workable number using criteria similar to those presented in table 5-7. Second, site files are requested from EPA, the US Geological Survey, and agency personnel associated with a remediation project. At this time, additional secondary criteria are evaluated, including assessment of participation, existing biomonitoring data, number of secondary or potential confounding contaminants, and reported health problems. During the third phase, site visits are conducted with local and state departments of health and environment and other interested officials. Affected neighborhoods are inspected, and special characteristics, including susceptible or transient

populations, are evaluated. Finally, on the basis of the above, a document presenting the rationale for selecting the site is prepared. The document is reviewed by ATSDR and, according to the resources available, the site is either approved or disapproved for establishment of an exposure registry. Final sites selected may also be based on the size of the population needed for the subregistry (JeAnne Burg, ATSDR, personal communication, 1993).

An individual is said to be *exposed* when 3 conditions are met:

- **A Contaminated Source**: Valid information indicates the presence of the contaminant(s) of interest in air, drinking water, soil, food chain, or surface water.
- **A Route of Transmission**: Evidence for that individual of one or more routes of entry (ingestion, inhalation, topical, or other parenteral routes) exists.
- **Indicated Transmission**: The contaminant traveled from the source via an appropriate route of entry to the body (ATSDR, 1988a).

TABLE 5-7 ATSDR Criteria for Setting Priorities for Sites

	Level	
Factor	Less Concern	Most Concern
Level of primary contamination	Below or close to standard (if known)[a]	Exceeds standard[a]
Toxicity of primary and secondary contaminant	Not a recognized human carcinogen, teratogen, neurotoxin, immunotoxin, etc., at levels present	Recognized as a human carcinogen, teratogen, neurotoxin, immunotoxin, etc., at levels present
Size of potentially exposed population	Small (<10 persons)	Large (>100 persons)
Current potential exposure	No	Yes
Past potential exposure (length)	Short term (<1 year)	Long term (>10 years)
Other considerations: particularly susceptible population and biomonitoring data indicating body burden		

[a]The standard is the level specified for that pathway.

Source: ATSDR, 1988a.

The data are collected by telephone, direct mailing, personal interviewing, or a combination of these. The method selected and the questionnaire or form are tailored to the specific subregistry. There will be a 6-month followup letter to confirm participation and address information and then annual or biennial updates of the core questionnaire via telephone interview (ATSDR, 1992). Participants may withdraw at any time.

The National Exposure Registry will contain multiple subregistries. The registry is to be maintained indefinitely, but the various subregistries may have a limited life (e.g., 5, 10, or 30 years) and a definite termination rationale (ATSDR, 1988a). The first 3 subregistries under development are for trichloroethylene (the most frequently encountered substance at proposed National Priority List sites (NRC, 1991b), benzene, and dioxin.

APPROACHES USED IN INFECTIOUS DISEASE EPIDEMIOLOGY

A major reason for the success of infectious-disease control has been the establishment of broad, effective surveillance systems to monitor patterns of disease. These surveillance systems are used to identify geographic areas having increases in incidence of disease, develop data to generate hypotheses about etiology, and test preventive health-control measures (Langmuir, 1963; Thacker et al., 1983; Thacker and Berkelman, 1988). With due note of the different time course and multifactorial etiology of many noncommunicable diseases, a similar surveillance mechanism could help to elucidate the contribution of environmental factors to disease.

CURRENTLY AVAILABLE INCIDENCE SYSTEMS

Many kinds of disease-incidence data systems exist. Communicable-disease surveillance has been established through public-health departments by making such diseases reportable. Public-health surveillance is a passive system in that physicians, hospitals, schools, laboratories, etc., are given the responsibility to report certain designated diseases to public-health units. In jurisdictions where a noncommunicable disease has been made reportable, such as cancer in many countries, mandatory reporting combined with imaginative use of existing data-collection systems has resulted in nearly complete national coverage. Mandatory reporting is not always necessary, however; the SEER program in the United States is voluntary but has virtually complete coverage. Some have even argued that making reporting mandatory might decrease the completeness and quality of reporting.

There are advantages and disadvantages in relying on passive public-health surveillance. The accuracy of case diagnosis may be poorer than

for disease registries, because registries often use standardized protocols for diagnosis and the initial reports are prepared by a relatively small number of persons who have been trained to use them (cancer registries, record-room librarians, etc.). The staff of registries is generally full time and dedicated to its purpose. In contrast, the accuracy and degree of ascertainment of public-health surveillance is low, because it depends heavily on the cooperation of large numbers of people who have other responsibilities and may rarely see the health outcome of interest. With both, there may be a tendency for bias, in that some socioeconomic groups may use the medical-care system less completely than other groups. However, the cost for the identification of cases is much lower for public-health surveillance than for registries of noncommunicable diseases. It is important to have a broad geographic coverage. With a broad coverage, areas of high incidence may be more readily identified, and, by evaluation of overall incidence patterns, it may be easier to determine whether clusters are merely the result of chance. Further, most environmental exposures are mixed, and multiple health outcomes are likely. For example, associations of air pollution with cancer, asthma, skin disease, and acute and nonacute respiratory symptoms are all plausible; therefore, concurrent monitoring of multiple disease outcomes is desirable. Even so, the amount of disease attributable to specific foci of environmental contamination is likely to be low. Thus, the effect of multiple other causes could easily overwhelm those from the environment, the "signal/noise ratio" being too low.

Public-health surveillance systems have not been used to any great extent for the monitoring of noncommunicable diseases that may be related to environmental exposures. Although such systems are generally inexpensive and with broad coverage, they can be incomplete, inaccurate, and misleading. With passive reporting, it has been estimated that only 10 to 50% of cases of serious communicable diseases are reported (Thacker and Berkelman, 1988). For communicable diseases, this is not a problem, as "outbreaks" of disease and changes in disease rates over time may represent a quadrupling in incidence in a very short period. In contrast, a "rapid" rise in a noncommunicable disease may represent less than a 10% increase in incidence over a period of years or decades. Passive reporting systems as currently constituted would not be able to identify such changes reliably; indeed, much-larger fluctuations would frequently occur through variation in reporting or by chance.

The ideal for surveillance of noncommunicable diseases is accurate incidence data across broad areas, long periods, and many diseases. It has been assumed that accurate incidence data will require virtually complete ascertainment of new cases within defined communities. However, society cannot afford registration systems for all diseases that may have

environmental causes. Further, the basic underlying assumption can be challenged. Although complete ascertainment of cases is ideal for accurate incidences, if the degree of ascertainment in relevant population segments is known, then this can be taken into consideration in estimating incidence. Thus, when complete registration is not feasible, disease reporting to designated public-health departments may be a means of providing data sufficiently accurate for the early-warning mechanisms that many members of the public are now demanding. An example is salmonella infections, for which reporting appears to be only about 1% of all cases (in part because many victims do not seek medical attention), yet that is sufficient to identify many food-related outbreaks.

POTENTIAL DIFFICULTIES WITH DISEASE MONITORING

Although disease reporting is relatively inexpensive, costs of reporting, analyzing, and interpreting the data may be substantial. There are probably too many end points and too many random departures from the baseline to slavishly use conventional probability testing methods. Time, money, and effort must be spent on chasing down each false lead, and this cost must be set against the value of the real leads that will be confirmed. Routine disease monitoring will not overcome the issues related to small populations exposed to many environmental hazards or, conversely, wide and almost uniform exposures of large populations. High relative risks could remain undetectable in the first instance, and large attributable risks in the second. The issues of confounders, biased reporting, and confidentiality are not solved by routine monitoring systems, but they may be brought into heightened visibility.

Routine monitoring is proposed as one mechanism by which to evaluate the causes of diseases of unknown etiology and to facilitate the detection of trends in diseases of environmental origin. However, this approach will often need to be supplemented and strengthened by other approaches.

CONFIDENTIALITY AND NEEDS FOR PERSONAL IDENTIFIERS

The study of interactions between environment and health poses questions of confidentiality and the need for personal identifiers. Over time, every person is exposed to a variety of potentially adverse environmental conditions and may experience a variety of adverse health effects. To evaluate complex interactions between environmental exposures and health effects, detailed and extensive information on individual subjects is often needed, and populations may have to be followed for long

periods. Linkages among data systems increase the difficulty of protecting the confidentiality of information.

The United States has 2 federal laws to protect the privacy of the individual from excessive government intrusion. One deals primarily with the collection and transfer of information, the other with its release. The latter, the Freedom of Information Act (FOIA) (5 USC 552), enacted in 1967 and amended in 1974, requires federal agencies to make most kinds of government records available to persons who request them. Health-data systems have not generally been seriously affected by FOIA, as it specifically exempts "personal and medical files and similar files the disclosure of which would constitute a clearly unwarranted invasion of personal privacy" (CDC, 1984, p. 6).

It is the Privacy Act of 1974 (5 U.S.C. 552a) that most affects health-information databases. The Privacy Act strictly limits what information government agencies can demand from the public and provides for legal protection of and safeguards on the use of personally identifiable information maintained in federal records systems. Congress has expressed some concern that the computerized databases in use today have outpaced the ability of individuals to protect their privacy when using the mechanisms set up to deal with the predominantly paper-record systems in use in 1974 (OTA, 1986). Specifically, the creation of record linkages between databases can run afoul of the Privacy Act which states that information may not be used for any purpose other than the purpose for which it was supplied (CDC, 1984). This can cause problems when researchers attempt creative and innovative linkages between databases that were intended for other purposes and do not have formal releases from the individuals to use their information for this purpose.

ATSDR's National Exposure Registry, for example, is subject to the Privacy Act. Although the registry is generally prohibited from disclosing personal information without written consent (which is routinely collected from participants through an informed-consent form), the Privacy Act does allow registry data to be released without consent in the following circumstances:

- To ATSDR personnel who maintain the registry.
- If required by FOIA (personal identifiers removed).
- For routine use. A routine use is defined as the use of a record for a purpose that is compatible with the purpose for which it was collected.
- To a recipient who has provided advance written assurance that the information released will be used solely for statistical research or as a reporting record. ATSDR requires that anyone seeking registry data for research purposes submit a study protocol for review to an agency review

panel that will in turn make recommendations to ATSDR. The final decision rests with ATSDR.

- To a person pursuant to a showing of compelling circumstances affecting the health or safety of an individual if upon disclosure to the requester notification is transmitted to the last known address of the individual.
- To Congress or the comptroller general.
- Pursuant to the order of a court of competent jurisdiction (ATSDR, 1988a, p. 31).

The ATSDR registry has the advantage of having been started well into the computer age and thus of being able to incorporate confidentiality protections into its system design. For older and other databases, however, the following specific issues must be considered:

- Can mechanisms be developed by which investigators can augment continuing longitudinal studies with new assessments in a timely fashion, perhaps by having the survey staff administer the tests so that more-detailed information could be provided to the investigator without risking the privacy of the subjects?
- Can statistical projects be established, whereby the information from several surveys could be augmented in a specific population? For instance, could a specific area be identified where the ambient concentrations of various pollutants are measured in more detail, the exposures of representative members of its population are evaluated in detail, and members of its population are subjected to internal dose assessments and health-status assessments? The registries of ATSDR are an example of such a mechanism, although they are not usually geographically circumscribed.
- Can the data systems of different agencies be linked, with return of the linked data to both agencies at the same level of detail as was provided?
- Should all data systems that obtain information on individual subjects collect personal identifiers in a consistent way and maintain them in a confidential data file for use later?

To maintain the confidentiality of data systems, statistical masks should be developed by agencies to protect the confidentiality of the data without distorting the relations among individual data items. Making linked data available on public-use data tapes is generally preferable to the agencies' releasing data and will help to maintain confidentiality. Linked data could be made available in a variety of formats—e.g., all demographic variables present but little geographic detail, or few demographic variables present but detailed geographic information—so that

each format could fulfill some analytic purpose without the possibility of investigators linking between formats and thus jeopardizing confidentiality.

Canada has had a National Mortality Data Base since 1950 and a National Cancer Incidence Reporting System since 1969. Many epidemiology studies have linked different data files, some collected originally for administrative purposes. These include evaluations of occupation and cancer (Howe and Lindsay, 1983), radiation and breast cancer (Miller et al., 1989), and pesticides in farmers (Wigle et al., 1990). The confidentiality issues have been solved largely by returning only anonymous data to investigators for analysis. However, when informed consent for linkage to vital-statistics data in the future had been obtained for a randomized trial of breast-cancer screening, individually identified information was returned to the investigators after linkage to the National Mortality Data Base (Miller et al., 1992).

DATA GAPS, RESOURCE CONSTRAINTS, AND RESEARCH OPPORTUNITIES

The utility of a data system in addressing an issue depends not only on the scope and quality of data but on the question being asked. Investigators and policy-makers all too often fail to recognize the multiplicity of questions, research designs, and data shown in table 5-1. There is a tendency for academic researchers to downgrade ecologic analyses and a tendency to discount data collected from regulatory-agency data systems. There is a tendency for policy-makers to consider the design or funding of data systems as though other data systems do not exist or as though regulation is their only purpose.

Only the federal government can coordinate the evaluation and linkage of many existing data systems and data-collection operations, and it should do so. The data systems should be supported by advisory groups of experts from all concerned agencies, and they should discuss system modifications that might enhance useful linkages. Existing systems should be evaluated with regard to their usefulness in estimating human risks associated with exposures, i.e., environmental-health end points. The federal government should also evaluate the data systems of the national environmental monitoring systems to determine whether some modifications might enhance their usefulness. Most federal environmental legislation does not require the collection of data needed to evaluate the health benefits of various environmental regulations. Hence, the environmental and health data systems have developed largely without consideration of environmental-health issues. Classical, targeted epidemiologic programs have not been buttressed with surveillance data that

would indicate the magnitudes of the environmental-health problems, and this hampers regulatory responses.

One way to begin to evaluate and identify modifications needed in existing systems is to develop a set of health-status indicators for inclusion in exposure, internal-dose, and health-status data systems. Attempts to list sentinel health events have usually focused on disease or syndrome end points (Rothwell et al., 1991; DHHS, 1991). However, many of the health end points are infrequent and therefore might not be useful in small populations. In studies that address public-health concerns—i.e., response epidemiologic studies—investigators often encounter symptom complaints and even collect information on symptoms, but symptoms are not regularly assessed by other health-data systems. Therefore, important data gaps are the lack of baseline health data on the frequency of certain rare diseases and conditions and the lack of data on the prevalence of symptoms in the general population.

Pollutant sources and ambient concentrations have been a focus of regulatory efforts. Assessment of the general health status of the population is usually a health-policy effort, largely independent of environmental health, so there is a paucity of data on human exposures and human internal doses. Existing and new data systems should be explored as sources for such data. One new data system is the National Human Exposure Assessment Survey, proposed by EPA. This survey was designed primarily to serve the interests of risk assessment, rather than to collect data to evaluate the effect of exposures on human health. However, relatively small changes in the design would materially increase the utility of this survey for environmental epidemiology. Such changes include the collection of personal identifying information, collection of data on other (nonenvironmental) potential confounders, and retention of data in a form that would permit linkage to outcome data sets, such as the National Death Index and population-based cancer registries. The committee urges EPA to cooperate closely with epidemiologists throughout the design of this survey, its implementation, continuing evaluation of the findings, and evaluation and possible modification of the study design.

An underused mechanism to collect exposure data is the brief activity questionnaire. The NHIS has been used to assess knowledge and protection practices relevant to environmental hazards such as radon and to occupational chemicals, but not to assess the duration and frequency of exposures of the general population. Brief questionnaires would, of course, provide less-detailed information on human exposure patterns than personal-monitoring studies.

Data to monitor the efficacy of various programs in decreasing body burdens of known toxicants (such as lead) are needed and, as new toxicants or data on new exposures (such as to mercury in paints) become

available, the distributions of body burdens need to be assessed to assist in the development of new regulations and public-health programs. These surveys will need flexible data-collection protocols because the toxicants to be assessed can easily change. Estimated body burdens of different chemicals should be periodically updated.

An important limitation that hinders greater use of linked data systems is investigators' lack of knowledge about potentially useful data systems and multiple kinds of data. Several innovations in recent years—such as electronic bulletin boards, commercial on-line systems, and distributed networks (Makulowich, 1993)—open up important new communication possibilities. These kinds of activities, as well as traditional inventories, should be encouraged.

Several problems limit the ability of investigators to link data from different systems. Often, the only linking variable available is geographic location. Information on location is often detailed for toxicant assessment but limited to broad areas, such as counties, for health assessments because of confidentiality concerns. Geographic information systems are overcoming the limitations on combining data with different types of geographic identifiers and from different sites. However, those systems are largely cross-sectional. The mobility of the population (and of pollutants) and the variable latent periods of health end points warrant longitudinal analyses. Existing systems and existing analytic procedures remain critical to the prevention or reduction of health problems from toxic exposures, but some simple, feasible, inexpensive changes would enhance the value of the data.

The ability of investigators to link data from multiple systems is also limited by the lack of comprehensive information on exposures and on health end points other than death. Morbidity data are generally collected through surveys with various sampling designs. Few data exist on exposures and internal doses that might be considered representative, let alone comprehensive, of even a circumscribed population. Although each environmental-epidemiologic issue could be addressed by a specially designed data system, such an approach would be prohibitively resource intensive, and ecologic analysis will often be the only feasible way to make general inferences about exposure and health. Even in a confined, compromise data system, detailed exposure patterns or the morbidity of each person simply cannot be obtained. Public-health policy decisions will depend on information from more-limited data systems, such as those described in this chapter, buttressed by studies of smaller populations that determine the validity and relevance of the information derived from larger population analyses.

The degree to which data systems represent or characterize a larger universe—i.e., a population or a well-demarcated region—should always

be made explicit. Many data systems cannot achieve comprehensive coverage, but there is a need to define sampling schemes better and to determine samples openly, rather than just to collect data from a sample of convenience. For some systems, a subset of environmental sites might be sampled to represent exposures of specific populations.

Another improvement would be the development of procedures to improve the comparability of data from different systems, that is, the use of common data modules or at least common data elements, including definitions of disease or health status. The National Exposure Registry is now collecting health-status data in a manner similar to that of the NHIS, and the prevalence of health end points determined through interviews with the registry population is to be compared with NHIS national population estimates for specific health-status indicators (ATSDR, 1992).

There is a need for many different types of cross-sectional, longitudinal, and followback studies to address environmental-health issues. Although some government agencies, such as the National Institutes of Health, regularly conduct varied studies to address etiologic issues, studies that address other needs in environmental health should be encouraged and conducted. Agencies involved with health promotion, such as ATSDR and the Centers for Disease Control and Prevention (specifically, the National Center for Environmental Health, the National Center for Injury Control and Prevention, and the National Institute for Occupational Safety and Health), need to conduct studies to address issues of health promotion and disease prevention.

The federal government should establish a mechanism by which to track the health impairments of populations for which data on exposures and other baseline measurements are available. The National Exposure Registry and the NHANES I Epidemiologic Followup Study are examples of such mechanisms. In addition, the National Death Index is a useful resource for the public-health community. Those mechanisms should be expanded, and this will require evaluation of confidentiality and other ethical issues, as well as careful review of the uses of data.

Improvements also could be made in the data systems that track inspections and compliance with regulations to enhance their utility for environmental-health assessment. For instance, the sampling period or geographic coverage around each site in the compliance-data systems could be extended. These systems typically collect environmental-concentration data until compliance is achieved. Collecting the data over an extended period would allow investigators to characterize the longer-term exposure patterns of sites known to contain pollutants. The choice of study sites where the nearby population distribution can be characterized would allow investigators to examine potential exposures with more assurance.

REFERENCES

Annest, J.L., J.L. Pirkle, D. Makuc, J.W. Neese, D.D. Bayse, and M.G. Kovar. 1983. Chronological trend in blood lead levels between 1976 and 1980. N. Engl. J. Med. 308:1373-1377.

Anto, J.M., J. Sunyer, R. Rodriguez-Roisin, M. Suarez-Cervera, and L. Vazquez. 1989. Community outbreaks of asthma associated with inhalation of soybean dust. N. Engl. J. Med. 320:1097-1102.

ATSDR (Agency for Toxic Substances and Disease Registry). 1988a. Policies and Procedures for Establishing a National Registry of Persons Exposed to Hazardous Substances (National Exposure Registry). Atlanta, GA: Agency for Toxic Substances and Disease Registry, Department of Health and Human Services, Public Health Service.

ATSDR (Agency for Toxic Substances and Disease Registry). 1988b. The Nature and Extent of Lead Poisoning in Children in the United States: A Report to Congress. Atlanta, GA: Agency for Toxic Substances and Disease Registry, Public Health Service, Department of Health and Human Services.

ATSDR (Agency for Toxic Substances and Disease Registry). 1992. The National Exposure Registry. Draft Document (5/19/92). [Atlanta, GA: Agency for Toxic Substances and Disease Registry, Department of Health and Human Services, Public Health Service.]

Boyd, J.T. 1960. Climate, air pollution, and mortality. Br. J. Preventive Soc. Med. 14:123.

Buck, S.F., and A.J. Wicken. 1967. Models for use in investigating the risk of mortality from lung cancer and bronchitis. Appl. Stat. 16:185.

CDC (Centers for Disease Control). 1984. CDC Staff Manual on Confidentiality. Atlanta, GA: Department of Health and Human Services, Public Health Service.

CDC (Centers for Disease Control). 1990. Guidelines for the determination of clusters. MMWR (no. RR-11).

Commission for Racial Justice, United Church of Christ. 1987. Toxic Wastes and Race in the United States. [New York]: Public Data Access, Inc.

Deane, M., S.H. Swan, J.A. Harris, D.M. Epstein, and R. Neutra. 1989. Adverse pregnancy outcomes in relation to water contamination, Santa Clara County, California, 1980-1981. Am. J. Epidemiol. 129:894-904.

DHHS (Department of Health and Human Services). 1991. Healthy People 2000: National Health Promotion and Disease Prevention Objectives. DHHS Pub. (PHS) 91-50212. Washington DC: US Government Printing Office.

Edmonds, L.D., and L.M. James. 1990. Temporal trends in the prevalence of congenital malformations at birth based on the birth defects monitoring program, United States, 1979-1987. MMWR CDC Surveill. Summ. 39(4):19-23.

EPA (US Environmental Protection Agency). 1985. Costs and Benefits of Reducing Lead in Gasoline. Final Regulatory Impact Analysis. EPA-230-05-85-006. Washington, DC: Office of Policy, Planning and Evaluation, US Environmental Protection Agency.

EPA (U.S. Environmental Protection Agency). 1991. National Air Quality and Emissions Trends Report, 1989. Research Triangle Park, NC: Office of Air Quality Planning and Standards.

EPA, NCHS, and ATSDR (U.S. Environmental Protection Agency, National Center for Health Statistics, and Agency for Toxic Substances and Disease Registry). 1992. Inventory of Exposure-Related Data Systems Sponsored by Federal Agencies. EPA/600/R-92/078. Prepared by Eastern Research Group, Inc., Arlington, VA, for US Environmental Protection Agency, National Center for Health Statistics, and Agency for Toxic Substances and Disease Registry.

Frank, R.G., M.S. Kamlet, and S. Klepper. 1986. The impact of occupational exposure to toxic material on prevalence of chronic illness. Pp. 59-63 in Proceedings of the 1985

Public Health Conference on Records and Statistics. DHHS Pub. (PHS) 86-1214. Hyattsville, MD: US Government Printing Office.

Fraser, P., C. Chilvers, and P. Goldblatt. 1982. Census-based mortality study of fertilizer manufactures. Br. J. Ind. Med. 39:323-329.

Fraumeni, J.F., Jr. 1987. Keynote lecture: etiologic insights from cancer mapping. Int. Symp. Princess Takamatsu Cancer Res. Fund 18:13-25.

Frisch, J.D., G.M. Shaw, and J.A. Harris. 1990. Epidemiologic research using existing databases of environmental measures. Arch. Environ. Health 45:303-307.

Glasser, M., and L. Greenburg. 1971. Air pollution, mortality and weather. Arch. Environ. Health. 22:334-343.

Goldman, L.R., D.F. Smith, R.R. Neutra, L.D. Saunders, E.M. Pond, J. Stratton, K. Waller, R.J. Jackson, and K.W. Kizer. 1990. Pesticide food poisoning from contaminated watermelons in California, 1985. Arch. Environ. Health 45:229-236.

Health Officers Association of California. 1986. Directory of Automated Information Systems in Local Health Departments. Sacramento: Health Officers Association of California.

Howe, G.R., and J.P. Lindsay. 1983. A follow-up study of a ten-% sample of the Canadian labor force. 1. Cancer mortality in males, 1965-73. J. Natl. Cancer Inst. 70:37-44.

Langmuir, A.D. 1963. The surveillance of communicable diseases of national importance. N. Engl. J. Med. 286:182-192.

Lave, L.B., and E.P. Seskin. 1973. Analysis of the association between U.S. mortality and air pollution. J. Am. Stat. A. 68:284-290.

Lynch, C.F., R.D. Woolson, T. O'Gorman, and K.P. Cantor. 1989. Chlorinated drinking water and bladder cancer: effect of misclassification on risk estimates. Arch. Environ. Health 44:252-259.

Mahaffey, K.R., J.L. Annest, J. Roberts, and R.S. Murphy. 1982. National estimates of blood lead levels: United States, 1976-1980: association with selected demographic and socioeconomic factors. N. Engl. J. Med. 307:573-579.

Makulowich, J.S. 1993. The use of electronic communications in environmental health research. Environ. Health Perspect. 101:34-35.

Mazumdar, S., S. Schimmel, and I.T. Higgins. 1982. Relation of daily mortality to air pollution: An analysis of 14 London winters, 1958/59-1971/72. Arch. Environ. Health 37:213-220.

Miller, A.B., G.R. Howe, G.J. Sherman, J.P. Lindsay, M.J. Yaffe, P.J. Dinner, H.A. Risch, and D.L. Preston. 1989. Mortality from breast cancer after irradiation during fluoroscopic examinations in patients being treated for tuberculosis. N. Engl. J. Med. 321:1285-1289.

Miller, A.B., C.J. Baines, T. To, and C. Wall. 1992. Canadian National Breast Screening Study. 1. Breast cancer detection and death rates among women 40 to 49 years. Can. Med. Assoc. J. 17:1459-1488.

National Governors' Association. 1989. The Potential for Linking Environmental and Health Data. Washington, DC: National Governors' Association.

National Task Force on Health Information. 1991. Implications of Privacy and Confidentiality Concerns on the Use of Health Information for Research and Statistics. Report of the Project Team to the National Task Force on Health Information. Ottawa: National Centre for Health Information, Statistics Canada. 43 pp.

NRC (National Research Council). 1989. Biologic Markers in Reproductive Toxicology. Washington, DC: National Academy Press.

NRC (National Research Council). 1991. Environmental Epidemiology. Vol. 1. Public Health and Hazardous Wastes. Washington DC: National Academy Press.

Ostro, B.D., and S. Rothschild. 1989. Air pollution and acute respiratory morbidity: an observational study of multiple pollutants. Environ. Res. 50:238-247.

OTA (US Congress Office of Technology Assessment). 1986. Federal Government Information Technology: Electronic Record Systems and Individual Privacy. OTA-CIT-296. Washington, DC: US Government Printing Office.

Ozkaynak H., J. Spengler, A. Garzd, et al. 1986. Assessment of population health risks resulting from exposure to airborne particles. In S. D. Lee, ed. Aerosols: Research, Risk Assessment, and Control Strategies. Chelsea, MI: Lewis Publishers.

Pickle, L.W., T.J. Mason, N. Howard, R. Hoover, and J.F. Fraumeni Jr. 1987. Atlas of U.S. Cancer Mortality Among Whites: 1950-1980. DHHS Pub. (NIH) 87-2900. Washington DC: US Government Printing Office.

Pope, C.A. 1991. Respiratory hospital admissions associated with PM_{10} pollution in Utah, Salt Lake, and Cache Valleys. Arch. Environ. Health 46:90-97.

Roos, L.L., J.B. Nicol, C.F. Johnson, and N.P. Roos. 1979. Using administrative data banks for research and evaluation: a case study. Eval. Quart. 3:236-255.

Rothenberg, R.B., K.K. Steinberg, and S.B.Thacker. 1990. The public health importance of clusters: a note from the Centers for Disease Control. Am. J. Epidemiol. 132(Suppl.1): S3-S5.

Rothman, K.J. 1990. A sobering start for the Cluster Busters' Conference. Keynote Presentation. Am. J. Epidemiol. 132(Suppl.1):S6-S13.

Rothwell, C.J., C.B. Hamilton, and P.E. Leaverton. 1991. Identification of sentinel health events as indicators of environmental contamination. Environ. Health Perspect. 94:261-263.

Schwartz, J. 1989. Lung function and chronic exposure to air pollution: a cross-sectional analysis of NHANES II. 1989. Environ. Res. 50:309-321.

Schwartz, J., and A. Marcus. 1990. Mortality and air pollution in London: a time series analysis. Am. J. Epidemiol. 131:185-194.

Sexton, K, D. S. G. Selevan, K. Wagener, and J. A. Lybarger. 1992. Estimating human exposure to environmental pollutants: availability and utility of existing databases. Arch. Environ. Health 47:398-407.

Sexton, K, D. K. Wagener, S. G. Selevan, T. O. Miller, and J. A. Lybarger. 1994. An inventory of human exposure-related databases. J. Exposure Anal. Env. Epidemiol. 4:95-109.

Thacker, S.B., and R.L. Berkelman. 1988. Public health surveillance in the United States. Epidemiol. Rev. 10:164-190.

Thacker, S.B., K. Choi, and P.S. Brachman. 1983. The surveillance of infectious diseases. J. Am. Med. Assoc. 249:1181-1185.

Wagener, D.K. 1990. Using biomarkers to assess exposure. In J.S. Andrews, Jr., B.O. Askew, J.A. Bucsela, D.A. Hoffman, B.L. Johnson, and C. Xintaras, eds. Proceedings of the Fourth National Environmental Health Conference: Environmental Issues—Today's Challenge for the Future. Held June 20-23, 1989 in San Antonio, TX. Atlanta, GA: Department of Health and Human Services, Public Health Service, Centers for Disease Control.

Wigle, D.T., R.M. Semenciw, K. Wilkins, D. Riedel, L. Ritter, H.I. Morrison, and Y. Mao. 1990. Mortality study of Canadian male farm operators: non-Hodgkin's lymphoma mortality and agricultural practices in Saskatchewan. J. Natl. Cancer Inst. 82:575-582.

Wingren, G., B. Persson, K. Thoren, and O. Axelson. 1991. Mortality pattern among pulp and paper mill workers in Sweden: a case-referent study. Am. J. Ind. Med. 20:769-774.

6

Opportunities for Methodologic Advances in Data Analysis

ASSESSING THE EFFECTS OF environmental agents poses well-recognized challenges. Relatively weak effects call for large samples and often complex designs; measurements of both exposures and outcomes may contain errors; the effect of the agent of interest may be confounded or modified by numerous other factors that may be unmeasured or even unrecognized; and the level of effect may vary with such time-dependent factors as age and duration of time since the exposure began. These methodologic issues have motivated the development of new study designs and analytic methods. Valid characterization of the effects of environmental agents and assessment of dose-response relationships often require the application of multivariate statistical models to control for confounding and to evaluate interdependence of effects. Models may also be needed to characterize temporal patterns of risk and to evaluate the consequences of errors in the measurement of the independent and dependent variables.

These challenges have been partially met by new statistical methods; advances in approaches to longitudinal data analysis have been particularly rapid. Although this chapter emphasizes advances in analytic methods, new epidemiologic approaches and the emergence and formalization of exposure assessment have contributed substantially.

This chapter reviews some of the methods, assumptions, and statistical techniques that can be applied to environmental analyses to improve and strengthen inferences about the relation between exposures and health outcomes and considers some of the many opportunities for further methodologic developments. New study designs for assessing ef-

fects of environmental agents were described in the report of the Health Effects Institute Environmental Epidemiology Planning Project (1993). Exposure assessment for studies of environmental agents was addressed in the first volume of this report. Thomas et al. (1993) have considered advances in contending with errors in exposure measures and their consequences.

INTRODUCTION

The most usual epidemiologic measure of effect of an environmental agent, the relative risk, is the ratio of the incidence of disease in those exposed to the agent of interest to the incidence in those not exposed. Categorization of subjects by exposure is straightforward in some types of epidemiologic research; for example, workers may be classified as exposed or nonexposed on the basis of personnel records and measurements of workplace contaminants. This classification of subjects into strata of exposed and nonexposed has analogy to experimental studies, such as clinical trials in which exposure status is controlled by the researcher. However, in epidemiologic studies of environmental agents, there may be no population that is entirely nonexposed, and the exposure may vary greatly from person to person in intensity, timing, and duration. In estimating the relative risk of disease associated with a particular environmental agent, the researcher may need to contend with multiple continuous and discrete variables, including the exposure of interest.

For example, such demanding data are encountered frequently in studying effects of environmental agents on respiratory health. Lung function, an outcome variable, is continuous whereas some predictors of interest may be (or be classified as) either discrete or continuous. Typically, this type of analytic problem is approached by modeling of the functional relations among variables. In general, a specific class of models is developed to examine functional dependence of outcome on risk factors, and a certain member or members of the class are identified as having adequately good fit to the data, or the class is rejected. For example, a common class is that of linear models, and the investigator may accept all linear models with coefficients in the calculated confidence ranges. The models include an associated distribution of differences between actual observations and those predicted by the models, and careful modeling always includes study of these differences. This approach allows for simultaneous consideration of the effects of multiple risk factors and the description of dose-response relationships for individual agents while controlling for the effects of other variables. Models in current use can accommodate both continuous and discrete risk factors.

Of course, inferences about relationships between predictor variables and outcome depend on the assumptions that a model requires regarding

the relationships among the variables, and it is likely that no model is absolutely correct in every detail. Nevertheless, an informative and biologically appropriate model extends the informativeness of data; a poor model may obscure true relationships between outcome and predictors; and both the efficiency and validity of inferences suffer if a model is seriously incorrect from either the biologic or statistical perspective.

Regression models, including linear regression, can be used to examine a specific proposed functional relation between a risk factor and an outcome. If a biologically inappropriate form of the relation is proposed, model findings may be misleading and incorrect.

In the past, linear models have been widely used to assess effects of environmental agents. Analysis of variance (ANOVA) and linear-regression models generally assume that the outcome varies linearly with functions of risk factors, that the individual observations are statistically independent, and that random differences from the model all have the same distribution, although models are available that relax each of these assumptions. For example, if the outcome seems to be approximately log-normally distributed, the investigator may assume that the natural logarithm of the outcome varies (approximately) linearly with continuous risk factors and that the errors of that model are (approximately) normally and independently distributed on a logarithmic scale. The outcome measures and risk factors are often assumed to be measured without error, although this assumption also can be relaxed. In any case, adherence to the assumptions underlying the chosen statistical model should be tested because violations can affect considerations of sample size as studies are designed and confidence bounds and hypothesis-testing as data are analyzed.

Thus, the linear-regression analyses that have often been used in studies of environmental agents typically carry strong underlying assumptions about the distribution of the data and the nature of the relationships being examined. There is a critical tradeoff here: the stronger the assumptions (if they are nearly correct), the more can be learned from a specific set of observations, but the greater the risk of a critical failure in one or more of the assumptions. Fortunately, most currently available statistical programs incorporate approaches to test compatibility with these assumptions, and some techniques allow analysis of data that violate one or more of these assumptions. Some of these methods are described below, with examples of their use, drawn largely from studies of the health effects of air pollution.

ANALYSIS OF DISCRETE OUTCOMES

Generally, counts, or discrete data, are assumed to follow some version of the Poisson distribution. The Poisson distribution does approach

the normal distribution as the mean of the counts gets large but the dependence of the variance on the mean remains. The classical Poisson distribution may also understate the variance of data. Extra-Poisson variability (i.e., greater variability in the counts than expected from a classical Poisson distribution) may exist in count data, and the possibility of such extra-Poisson variation needs to be examined in the modeling process. Less often, variances may be less than predicted from the Poisson distribution. Modeling the covariance structure of the data is discussed in more detail in the next section.

Data based on daily diaries, annual questionnaires of symptoms, and other outcomes considered in some respiratory studies are unlikely to meet all the assumptions that lie behind the common statistical approaches. Binary outcomes, such as the presence or absence of coughing, wheeze, or physician-diagnosed asthma, may be modeled as binomial data with logistic or probit regressions. Some investigators have analyzed such data using normal-theory, least-squares regressions. Such approaches make inefficient use of the data, and for rare events the linear model gives greater weight to the extremes than a logistic or Poisson model. Confidence bounds and significance tests for the effects of environmental exposures may be biased by incorrect application of these or other models.

ANALYSIS OF CORRELATED DATA

Measurements related in time and/or space, such as repeated measurements of the same population at successive times or measurements of persons from nearby geographic areas, are likely to be correlated, and their error terms may not be independent. For such data, the variance is unlikely to be characterized by a single dispersion parameter. Examples of such correlations include serial correlation (where correlations among measurements of the outcome at intervals are short relative to their variation over time), intraclass or intraindividual correlation (where multiple measurements in the same person over time are likely to show a similar deviation from the mean), and spatial correlation (where measurements in the same or nearby neighborhoods are likely to be correlated).

Correlations among data elements from either the same or different study units are common in many settings, and methods for dealing with the correlations are well developed. Similarly, methods exist to deal with nonuniform variance (heteroscedastic distributions). While such methods are now being used to assess effects of environmental agents, wider recognition of the problem and ways to deal with it is needed. The origins of the data often suggest the type of variance-covariance structures likely to be found. For convenience, the rest of this section is organized around

3 common types of correlation structures arising from different types of epidemiologic data. These are neither exclusive nor exhaustive but illustrate common analytic problems.

The first type is the serial correlation found in time-series data. This is relevant for longitudinal studies of the effects of pollution, which are increasingly used to study the effects of pollutants whose concentrations vary over time. The second type is correlation between different outcome measures. It is discussed in the context of structural equation modeling. The third type is that found within subject or site. It is discussed in the context of random-effects models, although other models are sometimes used.

LONGITUDINAL DATA ANALYSIS AND SERIAL CORRELATION

Longitudinal studies of the association between temporal variations in pollution and health outcomes have been useful in studying the health effects of outdoor air pollution. This design may also be informative in other areas of environmental epidemiology.

Longitudinal studies within a defined population have several attractive features. First, because they examine fluctuations within a sample, they are less subject than cross-sectional comparisons to several potential problems with confounding. For example, smoking, medical history, access to medical care, and socioeconomic factors are less likely to be serious confounders in a study in which the comparison is internal, i.e., the population serves as its own control. Although patterns of disease diagnosis may vary across regions or over long periods, these factors are unlikely to vary from day to day within an area, and any variation is unlikely to correlate with environmental pollution. Potential confounding in these longitudinal studies is limited to time-varying covariates, such as weather and seasonal factors. The potential strengths of this design, along with advances in statistical methods and software, have led to substantial growth in its application. For example, studies of daily mortality and air pollution have shown associations at concentrations found in many urban regions (Dockery and Pope, 1994). Whether these associations reflect a cause-effect relation is under active study.

Data from longitudinal studies also present analytic challenges, however. For example, if the value of the outcome variable under study is higher than average on a particular day, it is likely to be higher than average on the next day as well, even after conditioning on the covariates. This pattern, which affects almost all time series, is referred to as "serial correlation." In studies of disease occurrence, it arises from the persistence over several days of the conditions that alleviate, exacerbate, or depress illness (e.g., epidemics, weather, and allergy seasons) and also from

the natural persistence of disease. It may also arise from the slowness of change in variables that determine effective exposure or act as confounders. As a result, daily observations of most outcomes are not independent. Analyses of such data need to test for lack of independence and appropriately control for it when present.

METHODS FOR ANALYZING SERIALLY CORRELATED DATA

For normally (Gaussian) distributed outcomes, well-established methods of analysis can be used to take account of serial correlation. Work on autoregressive models dates from the 1940s; see for instance, Cochrane and Orcutt (1949).

The structure of the covariance is often parameterized in terms of autoregressive parameters, moving-average parameters, and combinations of the two. An autoregressive structure describes a model where the correlation between the residuals at time i and time $i - k$ declines monotonically as k increases. In a first-order autoregressive structure, for instance, the correlation between today and yesterday is assumed to be r, between today and 2 days earlier to be r^2, etc. A moving average has a correlation with a fixed lag and zero correlation with any further lags. Combinations can be chosen to fit the pattern of serial correlation observed in the data. In most cases, health and disease variables are likely to show autoregressive patterns because an abrupt termination of the correlation is not likely.

An alternative model, often called state dependence, refers to a Markov-type structure where the outcome on day i is dependent on the outcome on day $(i - 1)$ but, given the outcome at $(i - 1)$, not on any earlier outcome. For example, the prevalence of an illness with an average duration of a week (e.g., the common cold) will clearly depend on whether the subject had that illness the day before. Such models are described by Muenz and Rubenstein (1985) and were used in analysis of environmental data by Korn and Whittemore (1979). In contrast, incidence data are generally less subject to day-to-day correlation, though they can still be serially correlated (Schwartz et al., 1991), suggesting the covariance model described above. If there are covariates for which statistical modeling is imperfect (e.g., weather), the residuals of the model may also exhibit serial correlation. The presence of a lagged dependent variable in a model with serial correlation in the errors is unattractive because the correlation between the predictor variable and the error term means that usual least-squares regression estimates are biased and inconsistent. In these circumstances, the lagged dependent variable can be "instrumented." Instrumentation is the process of fitting a predictive model to a variable, using all possible predictors (except the hypothesis variable). Then the lagged

predicted value of the outcome is used as an independent variable instead of the actual lagged outcome variable.

In either case, long-term patterns in outcomes may be introduced by such slowly varying factors as season. Methods for filtering out such patterns include the use of seasonal dummy variables, seasonal autocorrelation, trigonometric filtering, and moving-average filters. The degree to which such filtering should be done when examining factors that may explain some of the seasonal variation in an outcome is a matter of epidemiologic judgment as to the biologic appropriateness of the alternative strategies.

The application of these methods in examining serial correlation in normally distributed continuous outcomes is illustrated in a study of air pollution reported by Pope et al. (1991). These investigators used an autoregressive model to examine day-to-day variations in peak expiratory flow, measured by mini-Wright peak-flow meters, in a panel of mildly asthmatic schoolchildren. The inhalable-particle concentration in outdoor air (PM_{10}) was significantly and inversely associated with peak expiratory flow. A followup study in nonsymptomatic children has found similar results (Pope and Dockery, 1992).

Poisson models for mortality counts, hospital admissions, or emergency-room visits may also exhibit serial correlation, as may daily diaries of binary outcomes, such as the presence or absence of coughing or wheeze; these outcomes can be modeled with logistic or probit regressions. Although ad hoc approaches have been used for serially correlated binary data, well-characterized statistical methods for dealing with serial correlated data in Poisson and logistic regressions have been developed (e.g., Gourieroux, 1984; Liang and Zeger, 1986; Zeger and Liang, 1986). These methods have been adopted to study effects of environmental agents (Schwartz et al., 1989; Braun-Fahrlander et al., 1992).

RANDOM-EFFECTS MODELS

The serial-correlation models discussed above are applicable in circumstances where the correlation between observations close together in time is not zero but decreases toward zero with increasing time between the observations. This pattern would be expected when, for example, the correlation is induced by external factors, such as weather or epidemics. A different pattern of correlation may arise from characteristics of persons. For example, if a child is taller than average at age i, the child is likely to be taller than average at some subsequent time i, and the correlation is not likely to go to zero even after a long interval. If a study does not need to explicitly include the trend of that factor in the model, the factor can be treated as a random subject effect. This type of correlation

may also affect some other end points of interest for environmental epidemiology, such as lung function. In analyzing data from studies of long-term trends in lung function, for example, this subject-mediated correlation needs to be considered. The period of measurement distinguishes between this correlation structure and the serial correlation described previously. If lung function were measured daily, there would undoubtedly be serial correlation in such data, in that each day's measurement is correlated with those of the preceding and following days. Annual measurements are separated enough in time for this short-term serial correlation to be diminished and for the correlation to be dominated by tracking of associations that do not chiefly reflect serial correlations. In most studies of environmental agents, individual effects are not of interest. Rather, we are interested in the effect of pollution on the entire population or, perhaps, the most sensitive segment of the population. The analytic strategy for data with repeated measures should recognize that the measurements on each individual are not independent. As a consequence, single intercepts are not of interest, and a large number of degrees of freedom would be used to estimate them. Estimating individual intercepts also is not consistent with the large-sample assumptions needed for estimation, because the number of parameters increases as fast as the sample size. In contrast, a random-effects model uses only a small, fixed number of parameters (sometimes just one) and thus preserves degrees of freedom for more-precise estimation of errors.

Measurements of individuals over time are not the only kind of data that exhibit such serial correlation. For example, persons who live in the same town tend to be more alike than persons randomly chosen from the population. People with similar socioeconomic and ethnic backgrounds often live in the same neighborhoods. As a result, outcome measures may be more similar between two subjects randomly selected from the same location than between two subjects randomly selected from the population as a whole. This is one source of the "design effect" in stratified clustered-survey designs. Unless all the causes of that similarity are controlled for in analysis of such data, the observations from within each site will be correlated, not independent, and the analysis should account for the correlation.

Two general types of correlation structures have been studied extensively. In one, the correlation between any two observations within a study unit is about the same. For example, there may be no reason to expect the correlation among subjects within a neighborhood to vary with index number or with location within the geographic area. Such models with random effects for sites, subjects, or other groupings have been extensively developed and used in environmental epidemiology and are well described by Laird and Ware (1982). In studies of environmental

agents by geographic region, use of random-effects models is critical if exposure data are ecologic. In ecologic studies, individuals' exposures are not direct individual measures, but estimates assigned to all members of subgroups on the basis of residence location.

Exposure is often estimated or measured at only a few geographic locations in air-pollution studies; exposures measured at one location are used as proxies for all individual exposures in studies in a surrounding area of hazardous-waste sites and other pollutant sources. If variables to indicate the location of each subject were used, differences in pollution exposure would be controlled so that no effect of exposure could be detected. In contrast, if the tendency for persons living in the same area to be similar is ignored, the standard errors of the regression coefficients are likely to be too small, which may lead to inflation of the level of statistical significance. The random-effects model represents a parsimonious approach to dealing with these design concerns while maintaining the ability to study exposures that are characterized geographically. Ware et al. (1986) illustrate the use of this technique in analyses of data on lung function and respiratory illness from children living in 6 US cities. Their approach incorporated a random-city effect. They found a significant association between mean covariate-adjusted rates of acute bronchitis and total suspended particle (TSP) concentrations across the cities. No association was found between TSP and pulmonary function. These findings were confirmed in data from additional groups of children in the same cities; these analyses were performed with Poisson regression (Dockery et al., 1989).

Similar correlations can affect data from studies that are not based on a clustered design. For example, Cook and Pocock (1983) reported that significant spatial correlation in a community affected study results.

When some added random variability is associated with unknown or unmodeled factors, a hierarchic formulation may be used to create a more-flexible model. The hierarchy assigns additional levels of random variability to the unknown parameters (such as underlying mortality rates or response probabilities). These unknown parameters are constructed according to a random model that allows potentially rich classes of statistical models to be considered. For example, methods for fitting random effects via hierarchic constructions have been discussed by Laird and Ware (1982), Racine-Poon (1985), Tsutakawa (1988), Vacek et al. (1989), Schall (1991), and Zeger and Karim (1991).

When the outcome measures are highly variable and the distribution of outcomes is well characterized, the empirical Bayes approach is useful for hierarchic modeling and for random-effects models (Reinsel, 1985). This approach is based on the concept that information about unspecified (or less than fully specified) random parameters may be imputed from

various portions of the model by using the Bayes rule. Conceptually, the approach imposes a distributional assumption on a set of parameters assumed to possess inherent variability. This allows one to effectively "borrow information" from the set as a whole, and then "pull back" the more-extreme and less-precise estimates of these parameters. This achieves a more-stable portrait of the pattern of variability as a whole. When there is information about errors in the exposure measures, measurement-error models are useful.

The empirical Bayes methodology has multiple uses in biomedical applications (Breslow, 1990) and may be of particular value in environmental epidemiology. For parametric models, the methods have been developed in some detail (Morris, 1983) and are known as "parametric empirical Bayes." Kass and Steffey (1989) refer to such a structure as a conditionally independent hierarchic model. The methods have been described for a variety of specific applications, including Poisson models (Albert, 1988; Gaver et al., 1990) where estimating mortality rates is of issue (Hui and Berger, 1983; Tsutakawa et al., 1985; Clayton and Kaldor, 1987; Desouza, 1991), particularly as regards geographic clustering or mapping of disease rates (Manton et al., 1989; Merril and Selvin, 1992); binomial-logistic models (Levin, 1986); and normal models (DuMouchel and Harris, 1983) of diseases that may be related to environmental factors.

MODELING COVARIANCE STRUCTURES

The primary effect of an agent is not always to change the expected value of a health outcome. Rather, the response to a pollutant may be heterogeneous, and the effects of interest may include modifying the response to other exposures, or the effect may be indirect through modification of outcomes other than that of primary interest. These issues are discussed below in order of increasing complexity.

Heterogeneity of Response

Heterogeneity of response to environmental agents is well established. For example, chamber studies of ozone exposure of exercising young adults identified a sensitive subgroup that had the largest short-term reductions in lung function in response to ozone (McDonnell et al., 1985). These differential responses were reproducible in subsequent challenges of the individual subjects. The degree of sensitivity to ozone did not seem to be associated with preexisting respiratory conditions, and markers predicting enhanced sensitivity have not been identified to date (McDonnell et al., 1985).

These laboratory findings are mirrored by field epidemiology studies

that have shown similar short-term reductions in lung function in response to ozone exposure in real-world situations. These data have come from studies of children in summer camps (Spektor et al., 1988; Lioy et al., 1985) and from a study of schoolchildren (Kinney et al., 1989). A reanalysis of the data from schools and the study of campers by Spektor et al. (1988) showed highly significant heterogeneity of response to ozone (Brunekreef et al., 1991). In contrast, short-term exposure to TSP was also associated with short-term reductions in lung function (Dockery et al., 1982; Dassen et al., 1986) but without evidence of heterogeneity of response to TSP (Brunekreef et al., 1991). The analytic method determined whether the variation in regression coefficients across subjects was greater than random, given the standard errors of subject-specific coefficients.

Similar results were noted in a study of respiratory symptoms. A panel study of asthmatics (Whittemore and Korn, 1980) found an association of exposure to TSP and ozone with increased respiratory symptoms, but no evidence of greater-than-random variability in the different TSP regression coefficients. In contrast, the ozone coefficients showed clear signs of heterogeneity of response.

Heterogeneity not only indicates sensitive subgroups, it affects estimation of the standard errors of regression coefficients, leading to improper hypothesis tests for pollution variables. Korn and Whittemore (1979) proposed a 2-stage method to address this issue. In the above-cited panel study of symptoms in asthmatics, they assumed that each subject's sequence of daily binary responses (with or without the symptom) followed a logistic model. However, instead of a common regression coefficient for air pollution, each subject was assumed to have a possibly unique regression coefficient. Although conceptually attractive, this approach requires sufficient data for asymptotic normality assumptions to hold. In fact, for consistency and asymptotic normality of the estimates, both the number of subjects and the number of days need to be large. Improvements to the Korn-Whittemore approach were given by Anderson and Aitkin (1985).

Population heterogeneity is an important statistical issue, even when there is no heterogeneity in response to the pollutant of interest. For logistic and Poisson models, a fixed relation between the mean and the variance of the distribution is generally assumed. There may be factors that alter the variation in the outcome and produce overdispersion or underdispersion. Either of these tendencies will result in incorrect standard-error estimates for regression coefficients, altering the probabilities of both type I and type II errors. McCullagh and Nelder (1983) discuss methods for estimating the overdispersion parameter in generalized linear models that include the logistic and Poisson regression settings.

Structural Equation Modeling

Environmental agents may have health effects that are evident only after a lag or lapse of time. The appropriate lag and averaging period for environmental agents are rarely known with certainty. Nevertheless, some studies have tried to characterize these times and to delineate temporal sequences of health effects. Biologically critical intervals of exposure may exist, particularly in the prenatal and neonatal periods. Further analytic problems may arise because some environmental agents may plausibly be related to more than one outcome. Concepts of pathogenesis may suggest that these outcomes may be components of the same broader biologic process. For example, increased airway responsiveness and asthma may increase the risk of permanent lung injury and of the disease referred to as chronic obstructive pulmonary disease.

These potential complexities in the relations between pollution and human health are often not explicitly acknowledged, in part because ordinary multiple-regression models deal with only one outcome at a time. The correlations among variables that are analyzed separately may be problematic. Such data may be better analyzed with methods for directly estimating systems of equations that reflect biologic understanding and can provide insights into causal pathways and underlying mechanisms.

The simplest example of how improper analysis of such correlated outcomes can lead to problems is called "simultaneous equation bias." Consider analyses of the effect of an air pollutant on 2 outcomes, e.g., coughing and a marker of inflammation. Typically these would be treated separately. If coughing and level of the marker are associated, analyses should consider this linkage. The association might take different forms. For example, elevations of the marker by other risk factors might increase the risk of coughing, but inflammatory responses to coughing might increase the level of the marker.

Bootstrapping

A statistical-computational innovation that has application in epidemiology is the use of resampling (Efron, 1982). This can be used to estimate variances of parameters that measure effects of environmental agents on human populations. The basic paradigm involves extensive computations to estimate certain underlying distributional qualities of the data when it is inappropriate or unwise to specify a distributional model (such as the normal distribution). One form of resampling that has been studied in some detail is the so-called bootstrap (Efron and Gong, 1983). The bootstrap procedure generates Monte Carlo pseudosamples from the observed data (Davison et al., 1986; Hall, 1987; Fisher and Hall, 1991). Re-

peated many times, each empirical pseudorandom variable provides information on the standard error or other parameters of interest. In many settings this approach can overcome analytic intractability or uncertain distributional assumptions by massive computer calculation (DiCiccio and Romano, 1988, 1990; Laird and Louis, 1989). Particular success has been achieved for various forms of regression models (Faraway, 1990; Huet et al., 1990), including generalized linear models (Mapleson, 1986; Rothe, 1989; Moulton and Zeger, 1991). Specific bootstrap applications have been developed for survival analyses and mortality estimation (Wahrendorf et al., 1987), longitudinal studies with repeated measurements (Moulton and Zeger, 1989), teratogenicity (Carr and Portier, 1992), and human genetic studies that include investigation of gene-environment interactions (Konigsberg et al., 1991).

ANALYSIS OF DATA WHEN THE SHAPE OF THE DOSE-RESPONSE RELATION IS UNKNOWN

Traditionally, toxicology has focused on finding the lowest-observed-effects levels (LOELs) and no-observed-effects levels (NOELs). Less emphasis has been placed on identifying the shape of the dose-response curve; typical experiments have compared 3 or 4 doses with a control regimen. In examining data from such studies, unsophisticated analytic techniques may not detect or may not properly define the true dose-response relation. Rothman (1986) cites an example of the relation between water-chlorination levels and brain cancer. Despite a monotonically increasing trend in outcome, an argument was made for a lack of association, against the evidence of simple visual inspection. Analysis was done by using pairwise t-tests against the lowest category, rather than a test for trend. The data were interpreted as suggesting a threshold between the third and fourth exposure categories. Nonparametric smoothing provides an estimate of the expected value of an outcome as a function of the exposure variable of interest, without making a priori judgments on the shape of the dose-response relation. This method can show indications of thresholds and nonlinearities in the dose-response curve. Alternating-conditional-expectations (ACE) regression models and generalized additive models are also valuable for examining dose-response relations. These models make fewer explicit assumptions about functional form than do linear models, for example, and hence their results are less likely to be biased. The price for this reduction in bias is a reduction in sensitivity to detect real effects.

Where data are not necessarily linear and the investigator has some notion about the shape of the underlying relation, parametric nonlinear models can be used. Examples are periodic regression (e.g., fitting sine,

parabolic, exponential, and power curves). Mathematical transformations are also often used.

Another alternative to linear expression is to allow the data to determine the shape to be fitted. One can assume that the expected value of Y varies continuously but not necessarily linearly with each x, then use a host of statistical methods, generally called nonparametric smoothing, to fit the expected value of Y for each x, assuming only continuity. For a detailed description of these methods, see Chambers et al. (1983) and Hastie and Tibshirani (1990). Examination of smooth plots can suggest the appropriate transforms for the linear regression, and also identify the existence of thresholds in the data, as well as the shape of other nonlinearities in the dose-response curve.

There are many different smoothing techniques, but all can be considered as generalizations of the following paradigm. If the expected value of Y is a continuous function of the independent variable x, then for a symmetric neighborhood around x_i the expected value of Y should be within a neighborhood around its expected value at x_i. If the neighborhood is small, the average of the expected values of Y at all the points in the neighborhood should be approximately the expected value at the center of the neighborhood, i.e., the expected value of Y at x_i. Hence, the average of the observed Ys in the neighborhood is an estimate of Y at point x_i that does not assume linearity. Since the error terms are random, averaging over multiple Ys allows the errors to cancel one another. The larger the neighborhood, the more error canceling there is, but the more opportunity there is for bias, if the relation is highly nonlinear within the neighborhood. This running-means approach is the basis for smoothing. More-sophisticated approaches use weighted averages, with weights that decline with distance from the central point x_i and deal with the variance-bias tradeoff and the problem at the ends of the distribution, where the neighborhoods are not symmetric. The general approach is called "kernel smoothing." Cubic smoothing spline estimation is another smoothing approach.

Recent simulation studies have shown that most of the modern smoothing approaches produce about the same curve. The critical parameter is the size of the neighborhood used. Some smoothing approaches, such as Supersmooth, use a cross validation method that can vary the size of the neighborhood.

The generalized additive model of Hastie and Tibshirani (1990) represents an alternative approach to nonparametric regression. This approach is equally valid for logistic and Poisson models and, indeed, for the entire family of generalized linear models. The generalized linear model (Nelder and Wedderburn, 1972) is a modeling approach that unifies a range of approaches, including ordinary linear models, logistic regression, Pois-

son models, and, with modification, the Cox proportional hazard model. An attractive feature of these models is that nonparametric chi-square tests can be computed for model improvement (reduction in deviance) compared with the assumption that the outcome depends linearly on each of the independent variables. This allows a direct test of evidence from nonlinearities. Linear and smoothed functions can be mixed, and for the linear functions, estimates of regression coefficients, standard errors, and chi-square tests of the significance of the association are available. For the variables represented by smooth functions, a test of the overall association can be obtained by comparing the improvement of the deviance when the term is added to the incremental degrees of freedom used up by the smoothed function.

Approximate standard-error bands are available in generalized additive models; however, their properties are not yet fully understood. Bootstrap estimates of the errors in prediction can also be produced. Generalized additive models can be used for hypothesis-testing and model selection. Alternatively, standard regression techniques can be used for model selection, and then the generalized additive model can be applied to the significant variables.

Nonparametric regression is particularly important in the study of multifactorial outcomes, where incorrect specification of the form of the relationship between the outcome and important covariates may result in an incorrect conclusion about relationships with a risk factor of interest. Efron and Tibshirani (1991) cite an example of a study of a procedure for treating cardiac abnormalities in infants. Use of the generalized additive model identified nonlinearities in the relationship between survival and age of child and to a lesser extent its weight. In their case, the estimated impact of treatment was not substantially affected.

ROBUST METHODS

To this point, this chapter has discussed methods for analyzing data that are not Gaussian by emphasizing data that clearly come from other distributions. Count data are often thought of as naturally Poisson-distributed, and the presence or absence of a condition as binomially distributed. Data may be roughly Gaussian, but there may be some deviation from that distribution that gives a relatively small number of observations an inordinate influence on some derived estimation, such as a regression coefficient. Such observations are often referred to as outliers; researchers often omit such observations from their analyses. This practice raises the issue of which observations, if any, should be deleted. Deletion is an extreme form of weighting—some observations are given no weight at all. The question can therefore be generalized: What weights

should be given to different observations to obtain robust results, that is, results that are not too sensitive to a few observations? Least-squares regression has attractive properties, such as being the minimal variance-unbiased estimate, if certain assumptions are met. If the data are not perfectly Gaussian, other estimators may be less variable. Since, as Mosteller and Tukey have pointed out, most real-world data are not really normally distributed, the issue of weighting arises in most situations.

A number of robust techniques, such as Mestimates and Lestimates, have been devised to give estimates that are more stable in the face of nonnormality in the data (Hampel et al., 1986). Some of these are available in commercial statistics packages. Application of these procedures may increase confidence that the results of analysis of epidemiologic data are unbiased (Efron and Tibshirani, 1991).

MODELING EXPOSURE

Statistical models can help to improve estimates of the exposures experienced by individuals. This section describes 2 methods that are helpful in improving exposure estimates.

KRIGING

Exposure data in environmental-epidemiologic studies are often sparse, irregularly sampled, and not based on measurements coincident with subject locations. For epidemiologic assessment, one needs to provide an exposure value for each study subject. For agents that vary in concentration across space, the methods generally used emphasize the importance of measured values near the subject, implicitly acknowledging the spatial continuity or spatial coherence of the exposure data. For example, to provide an exposure estimate for an unmeasured location, one could assign to that location the value from the nearest measured location or the average of the few nearest measured values. A method of interpolation that explicitly acknowledges and models the spatial similarity among measured samples, and hence may be more accurate, is kriging (Journel and Huijbregts, 1978; Cressie, 1991).

Kriging is a weighted, moving-average interpolation algorithm. For each point to be estimated, nearby samples are assigned weights reflecting their relative importance, and then these values (the observed datum times the weight) are averaged for the value at the new location. More-traditional approaches assume an arbitrary weight structure; for example, observations within a given distance from the point of interest may be given equal weights. Alternatively, the chosen weight may be the mean inverse function of the separation distance or an inverse function of the

square of the separation distance. The kriging algorithm provides a set of weights that are optimal for minimizing prediction errors.

More formally, kriging is linear estimation that minimizes the mean square prediction error subject to unbiasedness conditions. Before using kriging, one typically conducts a structural analysis of the data to identify and remove trends and outliers and to estimate quantitatively the spatial structure or spatial covariance of the observed data (i.e., how the correlations process observations that are a function of distance between them). The estimated spatial-covariance function is then used in the kriging equations to define optimal weights for averaging observed values near the location where an exposure is needed.

The spatial-covariance function used in kriging, known as the variogram, is defined by its functional form and 3 parameters: the sill, the range, and the nugget. The nugget is the (residual) variance among point pairs separated by zero distance. It is analogous to a sampling or replicate variance. The sill is the asymptotic variance of point pairs at infinite separation distance. The range is the separation distance at which the variogram first reaches the sill. Typical functional forms used to describe the change in variance as a function of separation distance include linear, exponential, Gaussian, and spherical models.

Using the spatial-covariance function, a set of equations is solved to provide weights for linear interpolation. These weights can be used to provide interpolation estimates and the variances of these estimates, where the estimation variance is a function only of the number and location of samples and is independent of the observed values. The variances of the estimates can be used to review the sampling design and to plan future sampling to provide the most information.

Various modifications to the basic kriging procedure can be used to accommodate more-complex aspects of data. For example, rather than using a single variogram, if the data are anisotropic one may use a set of variograms. For data that are not normally distributed, one can use nonlinear kriging methods, such as log-kriging or disjunctive kriging (Journel and Huijbregts, 1978). A special case of nonlinear kriging used to estimate locations with a specific value of a variable is called indicator kriging (Journel, 1983).

To date, kriging has been little used in epidemiology. Wartenberg et al. (1991) conducted a simulation study that showed that, under a set of simple assumptions, kriging marginally outperformed some other methods of interpolation. They applied a modified form of indicator kriging to case-control data to estimate the probability of disease occurrence at the locations of particularly sparse exposure-data sets.

Related methods have been used in a variety of epidemiologic contexts. Glick (1979) has used spatial-correlation methods to analyze and

describe cancer-mortality patterns, and Wartenberg and Greenberg (1990) investigated the use of spatial correlations for the study of disease clusters. Cook and Pocock (1983) modeled the spatial correlation of errors in a multiple-regression analysis of mortality patterns. Diggle et al. (1990) used kernel-smoothing methods to derive expected values for assessment of the spatial distribution of cases of laryngeal cancer near a hazardous-waste incinerator. Additional applications include assessment of the spatial pattern of soil contamination near lead smelters (Simpson, 1985).

MODELING EXPOSURE WITH ADDITIONAL EXPOSURE DATA

Cost often limits the collection of detailed exposure data. A protocol for collection of additional information might allow for the development of a better predictor of exposure than could be derived from the least-costly data available for all subjects.

An example of such an improvement is the use of diaries to record activity patterns. Ostro et al. (1991) assessed the impact of air pollution on persons with asthma living in Denver. They constructed an estimate of exposure by using outdoor monitoring, diary data on time spent outdoors, and a crude estimate of indoor/outdoor ratios of air pollutants. A stronger association was found with this measure than when data from the outdoor monitor alone were used as the exposure measure.

More-complicated models are possible. For example, Hasabelnaby et al. (1989) used indoor fine-particle measurements in a subsample of homes to estimate passive smoking exposure. These measurements were regressed against questionnaire data on maternal and paternal smoking, amount of smoking in home, housing characteristics, etc. This yielded a predictive model, whose independent variables were available for all subjects. Using the predicted exposure for all individuals improved model fit over that found using only questionnaire data on passive smoke exposure.

An important caveat related to the use of these methods is that the exposure metric is altered by the decision to use a modeled personal exposure instead of measured outdoor exposure. The regression coefficients from these approaches cannot be applied directly to exposure data from, e.g., central monitoring sites, to forecast effects. Thus, while these methods generally increase the power to detect an effect in the epidemiology study, they may complicate risk assessment. Reliance on data from central monitoring sites, in contrast, simplifies the risk-assessment process.

ADVANCES IN STUDY DESIGN

The principal observational designs for assessing the effects of environmental agents include the cross-sectional, cohort, and case-control

study designs. Each of these designs has well-characterized strengths and limitations (Rothman, 1986). In addition, the ecologic design is used, as in recent studies of air pollution and mortality. Recently, variants of these designs have been developed that offer increased efficiency in assessing the effects of environmental factors.

Case-cohort sampling methods have been a major advance (Prentice, 1986). Methods have been developed for sampling within cohorts that provide unbiased estimates of effect while potentially enhancing feasibility and lowering costs. In the nested-case-control design, appropriate characteristics are used to match controls to incident cases. More intensive exposure characterization may be possible for the smaller number of cases and controls in comparison with the full cohort. In the case-cohort design, a sample of the total cohort is selected without regard to the characteristics of the cases. These designs are particularly appropriate if the costs of exposure assessment are substantial or if invasive sampling, e.g., phlebotomy, is needed. For example, the nested-case-control design might be used to investigate genetic determinants of susceptibility to an environmental agent or used in an industrial setting in which detailed exposure assessment is expensive.

Methods have also been proposed for strengthening the case-control design (Thomas et al., 1993). In a 2-stage design, a basic case-control study is conducted with collection of information on exposure and disease variables only; in the second stage, data on other factors and possibly additional exposure data are collected on random samples of the 4 groups: exposed cases, exposed controls, nonexposed cases, and nonexposed controls. This design has the potential advantage of reducing costs. In the case-crossover design, the subjects serve as their own controls; this design has been offered as an approach to examine the effects of acute exposures on disease risk.

EXPOSURE-MEASUREMENT ERROR

Epidemiologists have long recognized that errors may be inherent in the measurement of both exposure and outcome variables. Recent advances in the area of exposure-measurement error offer new approaches for evaluating the consequences of these errors and making adjustments that can take into account the effect of error on estimates of the effects of environmental agents. Thomas et al. (1993) have provided a comprehensive review of these new techniques. An understanding of the consequences of measurement error is particularly relevant to the quantitative estimation of the risk of disease associated with environmental agents for the purpose of policy development. Quantitative risk assessments may use exposure-response relationships derived from epidemiologic studies;

these data may be subject to measurement error that biases exposure-response relationships. Correction of these relationships for error may be appropriate for regulatory policy.

CONCLUSIONS

Analysis of data from epidemiologic studies often uses statistical models that make strong assumptions about the distributions of disease and exposure and about the relationship between them. The real world rarely offers pristine and perfect data or justifies strong assumptions about the form of the relationships among data items of interest. Often, environmental exposures may produce relative risks in the range of 1.1–1.3. However, because of widespread exposure to many environmental pollutants, such small relative risks can imply large attributable risks. Improvements in statistical analyses of both multiple exposures and multiple diseases or outcomes will enhance the role of environmental epidemiology in addressing small relative risks.

Studies of environmental agents are likely to focus increasingly on multifactorial outcomes for which the exposure of interest accounts for a relatively small proportion of the variation in outcome. Many of the chronic diseases of interest in environmental-epidemiology studies have both multiple stages and multiple causes. Improvements in statistical methods that have been introduced in recent years will enhance the assessment of the contribution of multiple factors to these multiple outcomes.

While much attention has focused on modeling the expected value of the outcome, attention must also be focused on modeling covariance structures of the outcome. Data from studies of environmental factors may have autocorrelations in their residuals. Ignoring those correlations can give inefficient estimates of the parameters of interest (such as the regression coefficient of pollution) and biased tests of hypothesis. Environmental-exposure data often are not normally distributed, and care needs to be taken to deal with such non-Gaussian data properly. With the small signal/noise ratios commonly examined today, the use of techniques appropriate for non-Gaussian distributions becomes critical.

There has been considerable advance in the last 30–40 years in techniques for analyzing time-series and cross-sectional data. Additional work needs to be done to improve the ability of cross-sectional analyses to pinpoint risk factors.

As chapters 3 and 4 indicate, many preliminary environmental-epidemiology studies rely on exposure and health-outcome databases that are inadequate. Often self-reported information forms the basis for a preliminary study. Although some of the databases need to be improved, re-

searchers also need to develop a greater sophistication and familiarity with newer methods. To the extent that gradients of exposure can be estimated from existing data sets, the ability to detect associations of exposure and response will be enhanced.

REFERENCES

Albert, J.H. 1988. Bayesian estimation of Poisson means under a hierarchical log-linear model. Pp. 519-531 in J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.F.M. Smith, eds. Bayesian Statistics 3. Oxford: Clarendon Press.

Anderson, D.A., and M. Aitkin. 1985. Variance component models with binary response: interviewer variability. J. Roy. Sta. B 47:203-210.

Braun-Fahrlander, C., U. Ackermann-Liebrich, J. Schwartz, and H.P. Gnehm. 1992. Air pollution and respiratory symptoms in preschool children. Am. Rev. Respir. Dis. 145:42-47.

Breslow, N. 1990. Biostatistians and Bayes (with discussion). Statistical Science 5:269-298.

Brunekreef, B., P.L. Kinney, J.H. Ware, D. Dockery, F.E. Speizer, J.D. Spengler, and B.G. Ferris. 1991. Sensitive subgroups and normal variation in pulmonary function response to air pollution episodes. Environ. Health Perspect. 90:189-193.

Carr, G.J., and C.J. Portier. 1992. Dose-response models in quantal response teratology. Biometrics 84.

Chambers, J.M., W.S. Cleveland, B. Kleiner, and P.A. Tukey. 1983. Graphical Methods for Data Analysis. Wadsworth Press.

Clayton, D., and J. Kaldor. 1987. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. Biometrics 43:671-681.

Cochrane, D., and G.H. Orcutt. 1949. Application of least squares regression to relationships containing auto correlated error kerms. J. Am. Stat. Assoc. 44:32-61.

Cook, D.G., and S.J. Pocock. 1983. Multiple regression in geographical mortality studies with allowance for spatially correlated errors. Biometrics 39:361-371.

Cressie, N.A.C. 1991. Statistics for Spatial Data. New York: Wiley.

Dassen, W., B. Brunekreef, G. Hoek, P. Hofschreuder, B. Staatsen, H. de Groot, E. Schouten, and K. Biersteker. 1986. Decline in children's pulmonary function during an air pollution episode. J. Air Pollut. Control Assoc. 36:1223-1227.

Davison, A.C., D.V. Hinkley, and E. Schechtman. 1986. Efficient bootstrap simulation. Biometrika 73:555-566.

Desouza, C.M. 1991. An empirical Bayes formulation of cohort models in cancer epidemiology. Stat. Med. 10:1241-1256.

DiCiccio, T.J., and J.P. Romano. 1988. A review of bootstrap confidence intervals. J. Roy. Sta. B 50:338-354.

DiCiccio, T.J., and J.P. Romano. 1990. Nonparametric confidence limits by resampling methods and least favorable families. Int. Stat. Rev. 58:59-76.

Diggle, P.J., A.C. Gattrall, and A.A. Lovett. 1990. Modelling the prevalence of cancer of the larynx in part of Lancashire: a new methodology for spatial epidemiology. Pp. 35-47 in R.W. Thomas, ed. Spatial Epidemiology. London Papers in Regional Science 21. London: Pion.

Dockery, D.W., and C.A. Pope. 1994. Acute respiratory effects of particulate air pollution. Ann. Rev. Pub. Health 15:107-132

Dockery, D.W., J.H. Ware, B.G. Ferris, Jr., F.E. Speizer, N.R. Cook, and S.M. Herman. 1982. Change in pulmonary function in children associated with air pollution episodes. J. Air Pollut. Control Assoc. 32:937-942.

Dockery, D.W., F.E. Speizer, D.O. Stram, J.H. Ware, J.D. Spengler, and B.G. Ferris, Jr. 1989. Effects of inhalable particles on respiratory health of children. Am. Rev. Respir. Dis. 139:587-594.

DuMouchel, W.M., and J.E. Harris. 1983. Bayes methods for combining the results of cancer studies in humans and other species (with discussion). J. Am. Stat. Assoc. 77:293-313; Rejoinder:313-315.

Efron, B. 1982. The Jackknife, the Bootstrap, and Other Resampling Plans. Philadelphia: Society for Industrial and Applied Mathematics. 92 pp.

Efron, B., and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. Am. Statistician 37:36-48.

Efron, B., and R. Tibshirani. 1991. Statistical data analysis in the computer age. Science 263:390-395.

Faraway, J.J. 1990. Bootstrap selection of bandwidth and confidence bands for nonparametric regression. J. Stat. Comput. Simul. 37:37-44.

Fisher, N.I., and P. Hall. 1991. Bootstrap algorithms for small samples. J. Stat. Planning Inference 27:157-169.

Gaver, D. P., P.A. Jacons, and I.G. Muircheartaigh. 1990. Regression analysis of hierarchical Poisson-like event rate data: superpopulation model effects on predictions. Commun. Stat. Theo. Methods 19:3779-3797.

Glick, B. 1979. The spatial autocorrelation of cancer mortality. Soc. Sci. Med. [Med. Geogr.] 13D:123-130.

Gourieroux. 1984. Psuedo maximum likelihood. I. Theory. Econometrica.

Hall, P. 1987. On the bootstrap and continuity correction. J. Roy. Sta. B 49:82-89.

Hampel, F.R., E.M. Ronchitti, P.J. Rousseeuw, and W.A. Stahel. 1986. Robust Statistics: The Approach Based on Influence Functions. New York:Wiley.

Hasabelnaby, N.A., J.H. Ware, and W.A. Fuller. 1989. Indoor air pollution and pulmonary performance: investigating errors in exposure assessment. Stat. Med. 8:1109-1126; discussion 1137-1138.

Hastie, T.J., and R.J. Tibshirani. 1990. Generalized Additive Models. London: Chapman and Hall. 335 pp.

Huet, S., E. Jolivet, and A. Messian. 1990. Some simulations results about confidence intervals and bootstrap methods in nonlinear regression. Statistics 21:369-432.

Hui, S.L., and J.O. Berger. 1983. Empirical Bayes estimation of rates in longitudinal studies. J. Am. Stat. Assoc. 78:753-760.

Journel, A. 1983. Nonparametric estimation of spatial distributions. Math. Geol. 15:445-468.

Journel, A., and C.J. Huijbregts. 1978. Mining Geostatistics. London: Academic Press.

Kass, R.E., and D. Steffey. 1989. Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). J. Am. Stat. Assoc. 84:717-726.

Kinney, P.L., J.H. Ware, J.D. Spengler, D.W. Dockery, F.E. Speizer, and B.G. Ferris. 1989. Short-term pulmonary function change in association with ozone levels. Am. Rev. Respir. Dis. 139:56-61.

Konigsberg, L.W., J. Blangers, C.M. Krammerer, and G.E. Mott. 1991. Mixed model segregation analysis of LDC-C concentration with genotype-covariate interaction. Genet. Epidemiol. 8:69-80.

Korn, E.L., and A.S. Whittemore. 1979. Methods for analyzing panel studies of acute health effects of air pollution. Biometrics 35:795-802.

Laird, N.M., and T.A. Louis. 1989. Empirical Bayes confidence intervals for a series of related experiments. Biometrics 47:481-495.

Laird, N.M., and J.H. Ware. 1982. Random-effects models for longitudinal data. Biometrics 38:963-974.

Levin, B. 1986. Empirical Bayes estimation in heterogeneous matched binary samples with systematic aging effects. Pp. 179-194 in J. van Ryzin, ed. Adaptive Statistical Procedures and Related Topics. Hayward, CA: Institute of Mathematical Statistics.

Liang, K.Y., and S.L. Zeger. 1986. Longitudinal data analysis using generalized linear models. Biometrika 73:13-22.

Lioy, P.J., T.A. Vollmuth, and M. Lippmann. 1985. Persistence of peak flow decrement in children following ozone exposures exceeding the National Ambient Air Quality Standard. J. Air Pollut. Control Assoc. 35:1069-1071.

Manton, K.G., M.A. Woodbury, E. Stallard, W.B. Riggan, J.B. Creason, and A.C. Pellom. 1989. Empirical Bayes procedures for stabilizing maps of U.S. cancer mortality rates. J. Am. Stat. Assoc. 84:637-650.

Mapleson, W.W. 1986. The use of GLIM and the bootstrap in assessing a clinical trial of two drugs. Stat. Med. 5:363-374.

McCullagh, P., and J.A. Nelder. 1983. Generalized Linear Models. London: Chapman and Hall.

McDonnell, W.F., R.S. Chapman, M.W. Leigh, G.L. Strope, and A.M. Collier. 1985. Respiratory responses of vigorously exercising children to 0.12 ppm ozone exposure. Am. Rev. Respir. Dis. 132:875-879.

Merril, D. W., and S. Selvin. 1992. Analyzing geographic clustered response. Proceedings of the American Statistical Association, Section on Statistics and the Environment.

Morris, C.N. 1983. Parametric empirical Bayes inference: theory and applications. J. Am. Stat. Assoc. 78:47-55; Rejoinder:63-65.

Moulton, L.H., and S.L. Zeger. 1989. Analyzing repeated measures on generalized linear models via the bootstrap. Biometrics 45:381-394.

Moulton, L.H., and S.L. Zeger. 1991. Bootstraping generalized linear models. Comput. Stat. Data Anal. 11:53-63.

Muenz, L.R., and L.V. Rubenstein. 1985. Markov models for covariate dependence of binary sequences. Biometrics 41:91-101.

Nelder, J.A., and R.W. Wedderburn. 1972. Generalized linear models. J. Roy. Sta. A 135:370.

Ostro, B.D., M.J. Lipsett, M.B. Wiener, and J.C. Selner. 1991. Asthmatic responses to airborne acid aerosols. Am. J. Pub. Health 81:694-702.

Pope, C.A., and D.W. Dockery. 1992. Acute health effcts of PM_{10} pollution on symptomatic and asymptomatic children. Am. Rev. Respir. Dis. 145:1123-1128.

Pope, C.A., D.W. Dockery, J.D. Spengler, and M.E. Raizenne. 1991. Respiratory health and PM10 pollution: a daily time series analysis. Am. Rev. Respir. Dis. 144:668-674.

Prentice, R. L.. 1986. On the design of synthetic case-control studies. Biometrics 42:301-310.

Racine-Poon, A. 1985. A Bayesian approach to nonlinear random effects models. Biometrics 41:1015-1023.

Reinsel, G.C. 1985. Mean squared error properties of empirical Bayes estimators in a multivariate random effects general linear model. J. Am. Stat. Assoc. 80:642-650.

Rothe, G. 1989. Bootstrap for generalized linear models. Statistische Hefte 30:17-26.

Rothman, K.J. 1986. Modern Epidemiology. Boston: Little, Brown.

Schall, R. 1991. Estimation in generalized linear models with random effects. Biometrika 78:719-727.

Schwartz, J., D.W. Dockery, J.H. Ware, et al. 1989. Acute effects of acid aerosols on respiratory symptom reporting in children. Air Pollut. Control Assoc. Preprint 89-92.1.

Schwartz, J., D. Wypig, D. Dockery, J. Ware, S. Zeger, J. Spengler, and B.J. Ferris. 1991. Daily diaries of respiratory symptoms and air pollution: methodological issues and results. Environ. Health Perspect. 98:181-187.

Simpson, J.C. 1985. Estimation of spatial patterns and inventories of environmental contaminants using kriging. Pp. 203-242 in J.J. Breen and P.E. Robinson, eds. Environmental Applications of Chemometrics. ACS. Symp. Ser. 292.

Spektor, D.M., M. Lippmann, P.J. Lioy, G.D. Thurston, K. Citak, D.J. James, N. Bock, F.E. Speizer, and C. Hayes. 1988. Effects of ambient ozone on respiratory function in active, normal children. Am. Rev. Respir. Dis. 137:313-320.

Thomas D., D. Stram, and J. Dwyer. 1993. Exposure measurement error: influence on exposure-disease relationships and methods of correction. Ann. Rev. Publ. Health 14: 69-93.

Tsutakawa, R.K. 1988. Mixed model for analyzing geographic variabilty in mortality rates. J. Am. Stat. Assoc. 83:37-42.

Tsutakawa, R.K., G.L. Shoop, and C.J. Marienfeld. 1985. Empirical Bayes estimation of cancer mortality rates. Stat. Med. 4:201-212.

Vacek, P.M., R.M. Mickey, and D.Y. Bell. 1989. Application of a two-stage random effects model to longitudinal pulmonary function data from sarcoidosis patients. Stat. Med. 8:189-200.

Wahrendorf, J., H. Becher, and C.C. Brown. 1987. Bootstrap comparison of non-nested generalized linear models: applications in survival analysis and epidemiology. Appl. Stat. 36:72-81.

Ware, J.H., B.G. Ferris, D.W. Dockery, J.D. Spengler, D.O. Stram, and F.E. Speizer. 1986. Effects of ambient sulfur oxides and suspended particles on respiratory health of preadolescent children. Am. Rev. Respir. Dis. 133:834-842.

Wartenberg, D., and M. Greenberg. 1990. Detecting disease clusters: the importance of statistical power. Am. J. Epidemiol. 132 (Suppl.):156-166.

Wartenberg, D., C. Uchrin, and P. Coogan. 1991. Estimating exposure using kriging: a simulation study. Environ. Health Perspect. 94:75-82.

Whittemore, A.S., and E.L. Korn. 1980. Asthma and air pollution in the Los Angeles area. Am. J. Pub. Health 70:687-696.

Zeger, S.L., and M.R. Karim. 1991. Generalized linear models with random effects: a Gibbs sampling approach. J. Am. Stat. Assoc. 86:79-86.

Zeger, S.L., and K.Y. Liang. 1986. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 42:121-130.

7

Review of the Gray Literature from State Reports

THE COMMITTEE'S FIRST REPORT (NRC, 1991) reviewed published studies on the possible associations of exposure to hazardous wastes and human health. This chapter continues that assessment by reviewing selected studies on this subject from the gray literature, that is, studies that are available to the public but not published in the indexed scientific and technical literature. This chapter explains the process that the committee used to identify selected studies from the gray literature for assessment; presents a general review and assessment of those studies using criteria developed by the committee evaluates in depth some reports that specifically examined reproductive hazards of exposures to hazardous wastes and other materials and notes briefly some recent unpublished reports on human health and the environment that come from central Europe, where environmental contamination appears to be considerable. Despite some potentially serious problems, the gray literature can be an important source of information about environmental epidemiology.

INTRODUCTION

Many studies of health effects that may be associated with hazardous wastes are available on request but are never published. They fall into the category of "gray literature," which has been defined as literature that is not "white" (available and cataloged), and that is not "black" (not available, unknown, or not obtainable); gray-literature reports are usually produced in small quantities, intended for limited audiences, and not widely known (Schmidmaier, 1986). In the field of environmental epidemiology,

these gray-literature studies may include such items as state health-department reports, doctoral and master's theses, and reports produced by special-interest groups. In addition, the committee is aware that, with the dramatic changes in government in the former Union of Soviet Socialist Republics, large amounts of information on environmental health in that region are becoming available. Most of these reports are in the gray literature and even in that form have not appeared in the English-language literature.

REVIEW OF GRAY LITERATURE

Why is the gray literature gray? Many studies with substantial implications for environmental epidemiology are never published. A part of the problem is that in the United States and some other countries, the assessment of environmental health effects is often conducted in a context of litigation or potential regulatory action. This climate exerts strong pressures on any scientific assessment that may affect the selection of specific topics for study, protocol design, and methods of analysis, as well as public availability (e.g., sealing of records as a part of a negotiated settlement).

Earlier, the committee discussed some of the limitations that may confront public-health agencies in performing environmental-epidemiologic studies, including limited expertise, limited resources, lack of concern, secrecy of data, political pressure to conduct studies despite inadequate knowledge about the exposures or diseases in question, and the inherent limitations of uncontrolled "natural" experiments in which the populations are too small, the latent period too short, or the exposures or outcomes too poorly defined to yield useful results. In addition, there may be an issue of "publication bias" in which well-designed and well-conducted studies are not published, because of a lack of interest by journals or because of a bias against publishing negative findings in some instances and positive findings in others, especially when the "exposure" is not of great current interest among other scientists. Agencies may have little motivation for assembling studies into a format suitable for publication as staff move from fighting one fire to another. To understand better the question of why the gray literature is not published, the committee undertook to obtain a collection of such studies that had been produced by state health departments and others. Defined criteria were used to assess the quality, strengths, and weaknesses of each study and to estimate whether each report would be publishable in the peer-reviewed literature.

NGA-ATSDR STATE ENVIRONMENTAL HEALTH INFORMATION CLEARINGHOUSE

The National Governors' Association (NGA) under contract with the Agency for Toxic Substances and Disease Registry (ATSDR) initiated a

pilot project in May 1990 to set up a clearinghouse for state environmental-health studies. NGA established an advisory body to oversee the project, and members of this committee attended the initial meeting in July 1990. NGA and ATSDR decided at this meeting to obtain state-generated reports. The findings of the NGA pilot study were as follows:

- The State Environmental Health Information Clearinghouse was well received and very useful.
- The collection of clearinghouse information requires active solicitation.
- Operation of the clearinghouse by ATSDR would be mutually beneficial for the agency and the states.

Clearinghouse holdings were accessed to supplement the committee's database of studies. Because the clearinghouse was in its formative stages at the time of the committee's review, it could not serve as a primary source of studies for review, and it was necessary for the committee to request studies from the states directly as described below.

OBTAINING STATE STUDIES

While deciding which states to contact, the committee considered more than 20 state health departments. The committee formally contacted officials in states that were considered to have the most-appropriate data. These were California, Connecticut, Florida, Iowa, Massachusetts, Michigan, Minnesota, New Jersey, New York, Ohio, Texas, and Wisconsin in 1990 and again in 1991. Studies from additional states were obtained informally. Officials in the 18 states were asked to provide examples of studies meeting the following criteria.

- *Health outcomes*. The study investigates and reports on health end points, including biologic markers, physiologic alterations, or patterns of morbidity and mortality.
- *Exposure assessments*. The study is intended to identify risks of adverse health effects associated with specific exposures from hazardous-waste sites; these assessments are of potential use in epidemiologic investigation.
- *Population studied*. The study investigates health outcomes in a community possibly exposed to pollutants from hazardous-waste sites, rather than among workers.
- *Biologic plausibility*. The study investigates the association of biologically plausible (to state officials) health outcomes with potential exposure to hazardous wastes. (The committee is aware that concerned communities may request the investigation of associations that are not

biologically plausible, such as cancer clusters where the latency period is too long to link exposure to outcome.)

- *Explicit hypothesis or study question investigated*. The study was initiated to investigate a possible association between specified exposure(s) from a hazardous-waste site and specified health outcomes(s).
- *Methodology*. The study includes a comparison group, such as persons exposed at lower levels, an unexposed control group, or a baseline population or registry comparison.
- *Contributions to environmental epidemiology*. The study may contribute to the understanding of the health effects associated with hazardous wastes or may refine existing methods to introduce a new method for the epidemiologic investigation of hazardous-waste sites. Because many state-sponsored studies involve the investigation of cancer clusters, those approached were asked to use this criterion to screen such studies.
- *Stage of completion*. Preliminary or interim reports and descriptions of work in progress (for which the data collection and preliminary analysis had been completed) that met the above criteria were regarded as pertinent by the committee.

The limitations of our review must be noted. We necessarily focused on state reports, to the exclusion of other kinds of gray literature that may be quite different in important ways. Further, since only certain states participated and since the states themselves selected which reports to send, the reports reviewed by the committee would in no way represent a cross-section of all studies conducted by those states, nor would they represent studies conducted by all states. As noted above, studies submitted by the states and meeting the criteria above were supplemented with studies found in the NGA clearinghouse and those that were submitted informally. In general, we suspect that the reports reviewed here may be of better average quality than those not reviewed.

INITIAL EVALUATION

The committee evaluated the studies from the standpoint of characterizing the nature and quality of gray-literature reports. We did not attempt to evaluate results or conduct a meta-analysis of findings. Once the material was received, we selected a subset for further evaluation and asked the following questions: Is it an epidemiologic study? Is a community or residential population involved? Was there some kind of peer-review process or other evidence of quality assurance or accuracy checks? Did the investigators try to collect exposure data? Was the study done since 1980?

EPIDEMIOLOGIC CRITERIA FOR THE EVALUATION OF STUDIES FROM THE GRAY LITERATURE

To establish the role of the gray literature in environmental-health policy, the committee examined the literature submitted in much the same way that it would examine any other epidemiologic studies to assess quality, public-health significance, value to other researchers, and whether in the committee's judgment the (possibly revised) study would be publishable in a peer-reviewed journal. The following factors were examined.

Study Design

There is a relation between study power, sample size, prevalence of exposure, and expected rates of a given outcome in the study and control population. In general, studies of larger numbers of persons over longer periods with a big change in level of risk are more likely to yield positive results than are those involving smaller populations, shorter periods, and small changes in risk. The sample size to achieve a given study power is also related to whether exposure outcome is measured as a dichotomous or continuous variable, the variability in distribution of the exposure, the effects of confounders and errors on the measure of exposure, and the statistical methods used.

Exposure Assessment

Exposure measures are an important factor in the quality of an environmental-epidemiologic study. In some cases, exposure measurements already available can be used, but in other cases, special exposure assessments are required to fully assess possible effects on community residents. In still other cases, exposure measurements are not available and cannot be obtained, so some surrogate for exposure must be used. The validity and accuracy of these surrogate measures must be assessed. The time and duration of the exposure measurements, the methodologies used, and the relevance of these measures to the outcomes being evaluated are also important. Obtaining information for study participants who have had several different levels of exposure can provide an opportunity to look at important dose-response relations.

Health End Points

The health end points being evaluated should be appropriate with respect to the suspected exposure(s). Findings from a study examining a wide variety of end points may be less valuable than findings of a study

focusing on end points known to be associated with a particular chemical exposure both because the data and the analysis can be more-precisely targeted and because the problem of multiple comparisons is reduced. The method of ascertaining the end points is important and includes the accuracy of the source and validation of reported end points by other means. For example, the validity of self-reported medical diagnoses is improved if they are verified with medical records.

Study Results

Results of a study are stronger if the magnitude of the association between exposure and health outcome is greater; however, this must be balanced with the public-health significance of the outcome; i.e., a highly elevated relative risk of a minor and rare health problem may be less important than a less-elevated risk of a common and more-serious condition. The demonstration of a dose-response relation significantly increases the strength of a study. To determine whether a result is likely to reflect a true underlying relation, epidemiologists evaluate the congruence or consistency of findings with those of other, related studies and other scientific evidence.

Likelihood of Acceptance by a Peer-Reviewed Journal

The committee made a subjective evaluation of whether the studies would be definitely accepted, possibly accepted after substantial revision, or probably never accepted by a peer-reviewed journal. We determined potential reasons for rejection and made an assessment of the major weaknesses, if any, of the studies.

FINDINGS

Twenty-nine studies from the states of Colorado, Connecticut, Florida, Idaho, Louisiana, Maine, Massachusetts, Michigan, New Hampshire, New Jersey, North Carolina, Ohio, Tennessee, Texas, West Virginia, and Wisconsin were reviewed. California and New York were excluded because of the makeup of the committee.

Kinds of Designs

The most commonly conducted study was a comparison of disease incidence or prevalence in a study population with that in the general population. These studies were usually undertaken because of a reported clustering of some health outcome, such as leukemia, or because of com-

munity concerns about a source of exposure, such as a hazardous-waste site. Such self-selection tends to create a problem with response bias, in that persons who believe they are at risk might selectively recall exposures or outcomes better than those without such a concern. Some cross-sectional studies sought evidence on exposure, usually with a questionnaire accompanied by biologic monitoring (e.g., lead).

Ecologic correlational analyses were also undertaken in some instances. Some of these studies included data at the county level and compared disease rates in terms of available measures of exposure, such as drinking-water monitoring data. A few states conducted case-control studies to see whether individuals with and without some health condition had different environmental exposures.

Number of Subjects and Statistical Power

Studies mounted in response to community concerns often had low statistical power because of the small numbers of persons exposed. It was not clear whether estimates of statistical power had much influence on decisions about whether to conduct a study. Some studies were limited to the small numbers of people in the "concerned" or "exposed" community; others were expanded to a wider population to obtain a larger sample (or to use readily compiled data) but at the expense of including a greater proportion of people unlikely to be exposed. Investigators conducting the studies generally seemed aware of this problem and often provided lengthy explanations of the resulting limitations of the study design.

Exposure Assessment

In 25 of the 29 surveys reviewed for this chapter, the main exposure variable was a dichotomous estimate of exposure, usually whether or not the person lived in a defined area near a hazardous-waste site or industrial facility. Measurements from environmental sampling were usually sparse and only indirectly incorporated into the study design; e.g., only a few wells in a neighborhood were sampled.

In a few studies, biologic monitoring was conducted, generally accompanied by a questionnaire assessment of exposure to the substance of concern. The questionnaires often provided valuable information on exposures in a community. However, the biologic monitoring in the studies usually did not include any direct assessment of health end points.

The committee found the general lack of exposure assessment to be a significant weakness in these reports from the gray literature, and the characteristic most in need of improvement.

Adequacy of Measure of Health End Points

The range of health end points addressed in these studies was as broad as the spectrum of medical practice. They went from vaguely conceived complaints of neurobehavioral dysfunction to well-characterized syndromes of pulmonary hypersensitivity. A limitation of many of the studies was the reliance on available health records for ascertaining cases. State data can be efficient sources of information on health outcomes that can be nearly completely identified and aggregated by geographic area, but only a few health conditions can be adequately addressed in this way, e.g., cancer. While existing state data may provide a relatively inexpensive approach for responding to a community's health concerns, the limitations of these data sources must be recognized. These include availability only for relatively large geographic areas (unless special tabulations can be obtained) and lack of information on confounding factors. Few of these gray literature studies used clinical testing or medical examinations to assess health end points.

Study Results and Significance

A few studies demonstrated significant public-health problems that might not be appreciated by a review of the "white" literature alone, such as elevated levels of mercury in Chippewa Indians who consumed contaminated fish. While isolated positive findings must be viewed with skepticism, some are likely to reflect real risk. The results of this study and of 3 other biologic-monitoring studies examined provide information on human exposure from specific sources of environmental pollution. This information is very helpful for the evaluation of the human health significance of these pollution sources and for general scientific understanding of these types of exposure.

Approximately one-fourth of the studies report on methodologic approaches that might be useful for scientists considering studies of similar situations. For example, medical examination of residents living near hazardous-waste sites might be useful for investigators planning similar studies, even though the results from a particular study were inconclusive. As long as their existence is known to potential users, these studies need not be available in the peer-reviewed literature to help in evaluating a similar situation elsewhere; however, some centralized availability would be helpful.

Likelihood of Acceptability for Publication

Few of the state-sponsored studies would be acceptable for publication without substantial revision. However, in the judgment of the com-

mittee, at least one-third of these studies would be publishable if the authors invested the significant effort needed to extend the analyses and rewrite the reports. Other studies would never be publishable, because of the small numbers of persons, inadequate information on exposure, and potential confounders. This is not to say that the studies were of no value; many studies in the gray literature clearly have objectives other than the advancement of scientific knowledge, such as to allay public worries or to show official concern.

STATE STUDIES OF REPRODUCTIVE END POINTS

To focus attention on the problems and prospects for environmental epidemiology, the committee considered evidence from the gray literature on the possible association of environmental pollution with adverse reproductive outcomes. As a previous National Research Council report noted, the causes of the great majority of poor outcomes in human reproduction are not known (NRC, 1989). Moreover, increases in several of these adverse reproductive outcomes over time have been reported. These reports cannot easily be evaluated, however, because the completeness of the data for many reproductive outcomes, such as miscarriage and congenital anomalies, may have changed over time and is still not uniform in all regions of this country. The committee has found important data gaps in the field of reproductive epidemiology, and this review highlights the research opportunities.

Here we review 4 reports from state agencies that have evaluated the association between reproductive effects and environmental exposures (CBDMP, 1989; White et al., 1989; Johanson, 1991; MDPH, 1983). Three of the 4 studies reported an increased rate of the reproductive effect studied. These investigations illustrate 4 common difficulties in studying reproductive effects. First, exposure was seldom measured directly, so relations between exposure and outcome must be inferred rather than determined. Two, some of the studies lacked adequate control groups, so effects of potential confounders could not be taken into account. Three, not all potential reproductive end points were assessed. Fourth, the studies often had low statistical power to detect an effect of even substantial size. Following the summary of the studies, these difficulties are discussed in more detail.

CALIFORNIA BIRTH DEFECTS MONITORING PROGRAM: INVESTIGATIONS OF SUSPECTED CLUSTERS OF BIRTH DEFECTS BY COUNTY, SEPTEMBER 1, 1989, CALIFORNIA DEPARTMENT OF HEALTH SERVICES

This report examines 113 investigations of possible clusters of various

categories of birth defects reported to the California Birth Defects Monitoring Program (CBDMP). The possible clusters were recorded between April 1981 and September 1989. Of the populations with these 113 reported clusters, 8 were found to have statistically significantly elevated risks. Birth defects were identified and evaluated appropriately with an active surveillance system based on a review of medical records through the first year of life. For cluster studies, exposure was not determined for individual cases, and place of residence was the proxy for exposure. For case-control studies, exposures were determined by interviews. Five of the clusters had been or were being investigated with case-control methods. Two were not followed up, because the rate of birth defects of interest went down in subsequent years.

One cluster of brain tumors was referred to the California Tumor Registry. In a separate study of 8 counties where there was concern about an elevated rate of limb reductions in the offspring of agricultural workers, the CBDMP found an excess in the general population. The program did not have sufficient resources to examine rates in offspring of agricultural workers separately. The numbers of cases are not given for the studies; thus, the power cannot be determined. Because the outcomes of interest are rare and small populations were covered by most of the investigations, the committee assumes that these 113 investigations may have generally had low power.

MICHIGAN DEPARTMENT OF PUBLIC HEALTH: EVALUATION OF CONGENITAL MALFORMATION RATES FOR MIDLAND AND OTHER SELECTED MICHIGAN COUNTIES COMPARED NATIONALLY AND STATEWIDE, MAY 4, 1983

The Michigan Department of Public Health examined birth defect rates in the Midland, Mich., area, the site of a substantial chemical industry, for the years 1970-1980 in response to concerns from the community about environmental and occupational exposures to certain chemicals. Information on birth defects was obtained from the Birth Defects Monitoring Program (BDMP) of the Centers for Disease Control (CDC) and from birth and fetal-death records at the Michigan Department of Public Health. This is a descriptive study that serves as a screen to identify areas for future research. It is divided into 2 parts: the birth-defect rates for Midland County, based on the CDC BDMP databases and the birth-defect rates for the state of Michigan, based on state birth and fetal-death records.

In the first part of the study, the investigators used the BDMP database to determine the number of observed cases, the number expected from rates in the general population, and the ratio of the observed to expected numbers. They used county of birth as a surrogate for exposure.

Investigators found that of the 37 congenital-malformation rates followed in the database from 1970 to 1980, only one rate, hip dislocation without central nervous system (CNS) complications, was found to be statistically significantly higher in Midland County than in the United States. They believed that the study of 37 different rates could produce 1 or 2 elevated rates because of chance. They concluded also that the rate of hip dislocation without CNS complications was not expected to be related to environmental or occupational agents. Furthermore, the increased rate was not peculiar to this county but was significantly high for 5 other counties in Michigan during the same period.

The authors discussed 4 important problems in their use of the CDC BDMP database. First, there were 3 changes in the coding scheme over the 11-year period examined: from 1970 to 1973, the data were classified with the first edition of the *International Classification of Diseases*, Adopted Code for Hospitals (HICDA I); from 1974 to 1978, the data were classified with the second edition; (HICDA II); and from 1979 onward, the data were classified with the *International Classification of Diseases*, Adapted, Clinical Modification (ICD-A-CM). Code categories were added, deleted, and modified by these changes, so consistency in the overall database was reduced. However, they did not report on whether the specific outcome found to be elevated was affected by the changes. The second principal problem with the CDC BDMP database is underestimation of the number of infants with major anomalies, because the data were derived from hospital medical-discharge data, which do not capture anomalies noted only at outpatient visits. The third problem is that, during the period studied, the facilities participating in the system changed. For example, a hospital may have participated for only 5 of the 10 years under study. Finally, the place of residence of the parents is not listed. Other data show substantial in-migration of women delivering babies; only about 68% of live births in Midland County are to residents of that county. Out-migration is a smaller problem; 92% of live births and fetal deaths with a selected anomaly born to Midland County residents are delivered in Midland County.

The second part of the study evaluates the data from the birth and fetal-death records from the Michigan Department of Public Health. The quality of data is uncertain because the only sources of data used are birth certificates and fetal death certificates. Thus, the anomalies would have to be apparent at birth and recorded as the immediate cause of death or as a significant medical condition contributing to death. Reporting is often incomplete and, when compared with the CDC BDMP database, some anomalies in the state of Michigan appear to occur 25% less often than those in Midland County. In this example, the birth and death records in Michigan include documentation for only 25% of the cases. Two of the

anomalies, oral cleft and clubfoot, are reported as consistently in the Michigan database as in the BDMP database from 1970 to 1980.

There were 13,689 births to Midland County residents in 1970-1980. Of these, 136 were recorded in the BDMP database as having one or more of the selected anomalies. Grouping all congenital abnormalities from the Michigan birth and death records produced significantly higher rates, i.e., an observed/expected rate ratio of 1.71 for the years 1970-1975 and of 1.37 for the years 1976-1981. Investigators found 4 anomalies (cleft lip with or without cleft palate, cleft palate without cleft lip, hypospadias, and hip dislocation without CNS defects) to have significantly higher rates in Midland County than in the state of Michigan for the years 1970-1975. Furthermore, only the rates for hip dislocation without CNS defects were significantly elevated from 1970 to 1981. Investigators grouped cleft lip and cleft palate into one category, oral cleft, and examined the rates for this group. They found a "run" of successive years of higher-rates-than-normal of oral cleft and commented that other counties experienced similar runs. Also, Midland County's rates of oral cleft fluctuated from highest to lowest from 1976 to 1981. Investigators concluded that a case-control study is needed to assess factors that might contribute to the higher rates of oral cleft for Midland County.

Grouping of possibly related abnormalities—for example, all those in a single organ, such as the heart—could raise the power of a study to detect a significant relation. Such biologically meaningful grouping has at least 2 advantages. First, it reduces the problems of imprecise and overlapping case definitions for specific abnormalities, such that different observers may diagnose and record the same defect differently. Second, such grouping recognizes the possibility that different abnormalities of one organ could stem from similar causes, although specific expressions of the defect could differ. On the other hand, if these abnormalities stem from different causes, pooling data risks a dilution of effect. Therefore, both "lumping" and "splitting" strategies need to be explored and considered in studies of birth defects.

Additional studies of possible environmental causes of reproductive abnormalities are warranted, given the data in this report. In addition, the development of more-sophisticated analyses of existing data should be explored, including techniques to group diagnostic entities in a meaningful manner.

LOUISIANA DEPARTMENT OF HEALTH AND HOSPITALS: FINAL REPORT OF ST. GABRIEL MISCARRIAGE INVESTIGATION, EAST BANK OF IBERVILLE PARISH, LOUISIANA, SEPTEMBER 27, 1989

Residents of St. Gabriel, LA, were concerned about an elevated miscarriage rate and a possible link to air contaminants. Researchers from

the Tulane University School of Public Health and Tropical Medicine (White et al., 1989) conducted a descriptive study for the Louisiana Department of Health and Hospitals in response to these concerns from the community. They collected information from 354 women who were 18-50 years old at the time of the study, had conceived between April 1, 1982, and April 1, 1987, and had miscarried or delivered between May 1, 1982, and December 31, 1987. During protocol development, the investigators decided not to select a comparison community. They based this decision on problems with recall and selection bias and other factors such as inbreeding, that might affect the rates of miscarriages in small community comparison groups but would be difficult to assess. They instead decided to compare the rates to a range of rates of miscarriages that were previously reported in well-designed and documented studies. A priori, they decided that miscarriage rates below 15% would be considered normal; rates from 15% to 24% would be cause for examining known risk factors; and rates greater than 24% would be cause for further investigation. The investigators used proximity to industrial sites during pregnancy as the measure of exposure. Stillbirth rates were compared with the rates for all of Louisiana.

The sample was obtained from volunteers through extensive community outreach and from nonvolunteers through a review of hospital records, emergency-room logs, and vital records. The study of volunteers was by telephone. The investigators determined that, of 877 telephone numbers in the area, one would expect about 232 eligible participants (women 18-50 years old who had been pregnant), but they found only 119 eligible, or 51% of the number expected. The researchers attributed the low number of observed eligible women to problems with recruitment, and they used hospital records to find the rest of the expected eligible women. No information was given on biases that may have been introduced by this low rate of response. If there was a problem in finding women, then there may have been a problem in determining miscarriages.

Reported miscarriages were divided into those that were documented (by medical records) and undocumented. The 354 eligible women had 372 live births, 7 stillbirths, and 69 miscarriages (54 documented and 15 undocumented). Researchers found a documented miscarriage rate of 12.7% ± 1.6% and a documented plus undocumented miscarriage rate of 15.7% ± 1.7%. After age adjustment, they found a miscarriage rate among whites of 17.4% ± 2.9% and among blacks of 13.2% ± 2.1%. The researchers commented that white women in the study might be more likely to have their miscarriages recorded because they were more likely to seek medical attention and to release their medical records. Whatever the reason, without a suitable control group it is impossible to tell whether these rates are elevated. First, ascertainment may have been either more or less

complete than in other studies. Second, the rates in a small area may differ from rates in other studies because of differences in ethnic group, economic status, social class, or other factors. There are also considerations of random variation in small samples. The stillbirth rates were based on very small numbers and not statistically different from those for the rest of the state. However, while rates for white women (6.6/1000) were similar to rates for white women in the whole state (6.4/1000), rates for black women (22.9/1000) were nearly double the state rate for black women (12.1/1000), $p = 0.08$.

To assess exposure, the researchers classified the 400 pregnancies that had mailing addresses according to their distance from chemical plants and other industrial sites in the area. Subjects were divided into 3 categories: less than 0.5 mile, 0.5-1 mile, and over 1 mile. The investigators found no statistically significant relation between miscarriage rates and proximity to industrial sites. They also divided the pregnancies into 2 calendar-year-of-conception groups, using dates of delivery and gestational ages. They did not find a difference in miscarriage rate between conceptions in 1984-1985 and those in 1982-1983.

The researchers concluded that miscarriage rates were not elevated and that further studies of the rates were not warranted.

UNIVERSITY OF TEXAS HEALTH SCIENCE CENTER AT HOUSTON, SCHOOL OF PUBLIC HEALTH: W. JOHANSON, AN ANALYSIS OF A HEALTH EFFECTS SURVEY CONDUCTED BY RESIDENTS LIVING NEAR A TOXIC-WASTE SITE, DECEMBER 1991 THESIS FOR MASTER OF PUBLIC HEALTH DEGREE, THE UNIVERSITY OF TEXAS HEALTH SCIENCE CENTER AT HOUSTON, SCHOOL OF PUBLIC HEALTH

For years, concern had been expressed that exposures from waste-disposal facilities near Brio, Tex., endangered the health of nearby residents. A citizen-generated survey compiled evidence of environmental pollution and suggestions of a broad array of health problems, including an increased rate of birth defects in children conceived during the period of waste-disposal operations. This survey was conducted from January to September 1990, using volunteers to canvass the neighborhood and to distribute questionnaires to households. One person in each home was asked to complete and return the questionnaire. The quality of responses varied widely. Some respondents were very thorough and included detailed information on visits to doctors and dates. Others included only vague descriptions of health problems, no physician names, or no dates, including no birth dates. No information was collected to examine participation rates in various areas of the community, so there was an unquantifiable potential for response bias. Reported symptoms and diag-

noses were not verified independently. A team from the University of Texas School of Public Health (UTSPH), led by a physician-epidemiologist, conducted an independent evaluation of these same data from Brio, Tex. The investigators divided the population into 3 zones of potential exposure based on proximity to the waste site and wind patterns in the area. Zone 1 was defined as adjacent to the waste site, Zone 2 was 1,460-3,000 ft (0.4-0.9 km) and downwind from the site, and Zone 3 was 2,100-4,100 ft (0.6-1.2 km) from the site and away from the prevailing winds.

The UTSPH team evaluated 652 household-response forms for various health effects. Respondents reported 121 pregnancies, of which 25 (20.7%) ended in spontaneous abortion. They also evaluated the rates of reported birth defects and used data obtained from the CDC Congenital Malformations Surveillance Report (CDC, 1988) to estimate expected rates. Among the 96 live births, 18 (19%) were reported to have had congenital abnormalities. The investigators attempted to correct for recall bias by using a conservative risk estimate. They assumed that the number of cases recorded in 1990 for about one-third of the area was the annual incidence for the entire population that lived in the area over the period 1983-1989, while also assuming that the medical end points were valid as reported. With this assumption, they found a lower-bound relative risk of 2.4 for congenital birth defects and 3.8 for major CNS malformations. In 181 women in the study 19-50 years old, there were 126 pregnancies, for a fertility rate of 0.7 births per woman per year. This seems very high, even in the absence of a control group, and may indicate a serious problem in the data.

There may have been biases in the ascertainment of cases and confirmation of reported congenital malformations. First, prior to the health survey, extensive media coverage about the site included anecdotal reports about adverse health effects. This may have biased interview responses. Secondly, the survey was conducted by volunteers, and the response rate was low, leading to the possible biases. While there was some attempt to standardize the questioning procedure, there was no recording of home visits and outcomes or of attendance at training sessions. Volunteers may have been especially diligent in seeking positive responses, which would enhance the potential for recall bias. Third, the form was to be filled out by the interviewee rather than the interviewer, which could cause differences in interpretation of the questions and hence increase uncertainty in the answers. Lastly, no attempt was made to confirm the diagnoses by contacting physicians. Thus, there is no independent verification of reported cases. Other factors also inhibited the development of an independent assessment of this problem, including the protracted nature of the dispute, the inability to gather independent information, and the difficulty of obtaining validated measures of exposure.

ANALYSIS OF STUDIES REVIEWED

As with many public-health studies, none had a direct measure of the exposure that might have caused the observed outcome. The investigations addressed exposure assessment in different ways. The investigators in Texas used a surrogate measure of exposure that depended on distance to the toxic-waste site and the prevailing winds. In the Louisiana study, the researchers classified pregnancies by distance to industrial sites. In the Michigan and California studies, the place of residence was the surrogate for exposure, and rates were compared with background rates.

Another problem is the lack of non-exposed or less-exposed persons as controls. Comparing regional rates of outcomes with national rates can give misleading results because the national averages do not take into account regional characteristics (e.g., race and socioeconomic status) that may differ from characteristics of the general population. Many communities near hazardous-waste sites consist of minority groups, are economically disadvantaged, or both (Commission for Racial Justice, 1987; Bullard, 1990; EPA, 1992a,b). In the Michigan study, investigators found a higher rate of oral clefts in Midland County than in the state of Michigan. They commented that this would be expected because there was a higher proportion of whites in Midland County than in the state and because higher rates of oral cleft are reported in whites than in blacks. Similarly, the Louisiana study found different miscarriage rates among whites and blacks, indicating that race is a confounder. The Texas study also lacked a control group. For the California studies, availability of a population-based registry of birth defects allowed comparisons to be made among similar areas with identical means of case ascertainment and verification.

A third difficulty in examining potential reproductive hazards comes from the many reproductive and developmental health effects that could be studied (Mattison et al., 1989, 1990). Examination of multiple end points is necessary to avoid possibly missing some important effects. None of the investigations looked for changes in fertility rates.

One way to improve the power of a study is to group related defects. In the early stages of investigating possible links between outcome and exposure, one must consider all plausible outcomes to capture those that are most likely to occur from the exposure (Axelson and Söderkvist, 1991). However, a major problem is that grouping may be biologically implausible for certain birth defects. On the other hand, if the effects examined are from exposure to a hazardous-waste site, where there may be exposure to multiple agents, it may be appropriate to group all defects, since different agents may cause different adverse outcomes.

REPORTS FROM OTHER COUNTRIES

In accordance with its effort to examine the relevant literature on environmental epidemiology, the committee surveyed some preliminary findings that have been developed on environmental health in some countries other than the United States, focusing on the literature from China and eastern Europe. Given the difficulties in translation and access to information, the committee is not able to vouch for the quality of data from these countries. Instead, we have relied on reports developed by expert consultants to the World Bank, the Pan American Health Organization, and other international organizations. These studies do not appear in the peer-reviewed literature but are available from the sponsoring organizations. They represent the work of visiting teams of well-qualified researchers who have generated important analyses using accepted techniques. As is often the case, commonly available information tends to focus chiefly on patterns of environmental contamination. However, in several instances, analyses of geographic variations in disease have also been produced.

Examination of patterns of disease within countries can yield important clues about environmental causes of disease. Urban living tends to be correlated with greater smoking and alcohol drinking, as well as access to health services—factors that clearly influence health outcomes and their recording. A marked urban gradient is evident for many diseases, including lung cancer. In contrast, for outcomes that are dependent on lack of health services, such as maternal mortality, rates tend to be higher in rural areas (Bertaud and Young, 1991).

In some Chinese cities, dramatic environmental inequity exists, despite relatively homogeneous income levels and universal health services. In Tianjin in 1988, female lung-cancer mortality was 9 per 100,000, the highest in China; and in the more-urban sections of this area, the rate averaged 30 per 100,000, ranging from 10 to 88. Approximately 13% of the population lives in districts where the yearly rate of female lung cancer is higher than 40 per 100,000. This is more than 12 times the national average (Bertaud and Young, 1991). Smoking is unlikely to be a major determinant of these differences, as about 12% of females in China were smokers (Junshi et al., 1990), and smoking patterns are not likely to vary greatly within small areas.

The causes of these elevated rates of lung cancer in a largely nonsmoking population need to be sought but are likely to reside in environmental factors that are not yet identified. In some areas of China, researchers have linked elevated rates of lung cancer to the use of smoky domestic fuels and cooking oils (Mumford et al., 1987). Researchers from China and the US National Cancer Institute have recently confirmed that

the elevated risk of lung cancer in nonsmoking women can be linked to a variety of exposures to household smoke (McIntosh, 1992).

Studies from Rumania show a similar result, with a marked urban gradient for many diseases. Given the highly centralized nature of the former government, health statistics are believed to be reliable. Lung cancer in women exhibits interesting variations. Women in some areas of the country have rates of lung cancer that are 5 times the rates in other areas. In general, women have not smoked much, although those in cities smoke more than those in rural areas. Overall rates of cancer in Rumania are lower than those of developed countries.

According to one survey by the ministry of health of the former Soviet Union, people living in 103 cities breathe air that carries at least 5 times the allowed limits of many pollutants (Feshbach and Friendly, 1992). Studies of the impact of this pollution are just now emerging. Some reports from cross-sectional surveys of 68 cities where air pollution is regularly measured have found levels at least 10 times above permissible maximums, with rates of illness more than double those of the national average (Kondrusev, 1990). Almost three-fourths of the nation's surface water is polluted, and one-fourth is untreated.

REFERENCES

Axelson, O., and P. Söderkvist. 1991. Characteristics of disease and some exposure considerations. Appl. Occup. Environ. Hyg. 6:428-435.

Bertaud, A., and M. Young. 1991. Geographical Pattern of Environmental Health in Tianjin, China: The Development of an Environmental Health Data Base. INURD Working Paper 13, February, 1991. [Washington, DC]: Urban Development Division Sector Policy and Research, The World Bank.

Bullard, R.D. 1990. Dumping in Dixie: Race, Class, and Environmental Quality. Boulder, CO: Westview Press.

CBDMP (California Birth Defects Monitoring Program). 1989. Investigations of Suspected Clusters of Birth Defects by County. September 1, 1989. [California Department of Health Services.]

CDC Centers for Disease Control. 1988. Congenital Malformations Surveillance Report, January 1982-December 1985.

Commission for Racial Justice, United Church of Christ. 1987. Toxic Wastes and Race in the United States: A National Report on the Racial and Socio-Economic Characteristics of Communities with Hazardous Waste Sites. [New York]: Public Data Access.

EPA (U.S. Environmental Protection Agency). 1992a. Environmental Protection: Has It Been Fair? EPA Journal 18(March/April):entire issue.

EPA (U.S. Environmental Protection Agency). 1992b. Environmental Equity: Reducing Risk For All Communities. EPA230-R-92-008 (vol. 1), EPA230-R-92-008a (vol. 2). Washington, DC: Policy, Planning, and Evaluation, US Environmental Protection Agency.

Feshbach, M.O., and A., Jr. Friendly. 1992. Ecocide in the USSR: Health and Nature Under Siege. New York: Basic Books.

Johanson, W. 1991. An Analysis of a Health Effects Survey Conducted by Residents Living

Near a Toxic Waste Site. Thesis for Master of Public Health. Houston: The University of Texas Health Science Center at Houston School of Public Health.

Junshi, C., T.C. Campbell, L. Junyao, and R. Peto. 1990. Diet, Life-style, and Mortality in China: A Study of the Characteristics of 65 Chinese Counties. New York: Oxford University Press.

Kondrusev, A. 1990. Poison without illusions: who will guarantee the right to health? From interview with N. Gogolin, Pravda, November 14, 1990, 2d ed., p.6. Pp. 50-52 in Joint Publications Research Service, Soviet Union, Political Affairs. JPRS-UPA-90-066, December 4.

Mattison, D.R., J.W. Hanson, D.M Kochhar, and K.S. Rao. 1989. Criteria for identifying and listing substances known to cause developmental toxicity under California's proposition 65. Reprod. Toxicol. 3:3-12.

Mattison, D.R., P.K. Working, W.F. Blazak, C.L. Hughes, J.M. Killinger, D.L. Olive, and K.S. Rao. 1990. Criteria for identifying and listing substances known to cause reproductive toxicity under California's proposition 65. Reprod. Toxicol. 4:163-175.

McIntosh, H. 1992. Far East's cancer mortality patterns shift, reflecting cultural changes. J. Natl. Cancer Inst. 84:1069-1071.

MDPH (Michigan Department of Public Health). 1983. Evaluation of Congenital Malformation Rates for Midland and Other Selected Michigan Counties Compared Nationally and Statewide, 1970-1981. 33pp.

Mumford, J.L., X.Z. He, R.S. Chapman, S.R. Cao, D. B. Harris, X.M. Lit, Y.L. Xian, W.Z. Jiang, C.W. Xu, J.C. Chuang, W.E. Wilson, and M. Cooke. 1987. Lung cancer and indoor air pollution in Xuan Wei, China. Science 235:217-220.

NRC (National Research Council). 1989. Biologic Markers in Reproductive Toxicology. Washington, DC: National Academy Press.

NRC (National Research Council). 1991. Environmental Epidemiology. Vol. 1. Public Health and Hazardous Wastes. Washington, DC: National Academy Press.

Schmidmaier, D. 1986. Ask no questions and you'll be told no lies: or how we can remove people's fear of "gray literature." Libri 36:98-112.

White, L.E., F.J. Mather, J.R. Clarkson. 1989. Final Report, St. Gabriel Miscarriage Investigation, East Bank of Iberville Parish, Louisiana. September 27, 1989. Prepared by Tulane University School of Public Health and Tropical Medicine, New Orleans, Louisiana, for Louisiana Department of Health and Hospitals, in cooperation with Division of Birth Defects and Developmental Disabilities, Center for Environmental Health and Injury Control, Centers for Disease Control, Atlanta, GA., and Agency for Toxic Substances and Disease Registry, Public Health Service, Atlanta, GA. 55 pp.

8

Major Conclusions and Recommendations

BECAUSE THE CHARGE TO THE committee focused on an examination of the gray literature to determine its possible contribution to our knowledge about links between environmental exposures and chronic diseases, we summarize here in some detail our conclusions regarding the gray literature in the United States. The remainder of this chapter focuses on the other main element of the charge—improvement of methods for performing epidemiologic studies—and presents major conclusions and recommendations on methodologic and related issues.

The review of the gray literature, in which the committee examined studies that states selected for our review, probably caused us to look at the best of this genre of epidemiologic research. Nonetheless, major limitations in the utility of the data exist. A major limitation of much of the gray literature is that many of the studies reported are initiated in response to a political problem but with inadequate funding and unavoidable limitations that seriously undermine the credibility and value of the final product. Although these political pressures can never completely be avoided (and perhaps should not be), the committee believes that it would be useful to establish guidelines to determine the circumstances under which epidemiologic studies should be undertaken. In some cases, there may be appropriate responses to public concerns other than initiating an epidemiologic study, for example, using existing evidence to address the concern. Fewer, but better, studies could advance the public welfare. External scientific advisory committees can also help to promote better scientific approaches for these studies.

Many of the studies in the gray literature could have been published

in the peer-reviewed literature with additional effort. Researchers working in state agencies need both incentives and support to publish study findings. Many studies reviewed by the committee would be of great interest to researchers in other areas. The committee encourages activities, such as the Agency for Toxic Substances and Disease Registry (ATSDR) State Environmental Health Information Clearinghouse, to make these gray-literature reports available to others.

Public-health agencies need a combination of speed and low cost in their research efforts. Thus they use readily available data for populations and geographic regions, and many gray-literature studies used an ecologic design. The disadvantages of this approach are not fully and widely understood; in the literature reviewed for this report, many studies had significant shortcomings. More research on methods, improvements in the methods already available, and better training are needed for scientists using this approach. It would be beneficial to develop guidelines for typical health-department studies to assist practitioners. These might be versions of existing texts used by academic epidemiologists.

Another common limitation in these studies was their inappropriate use of exposure and health-outcome databases that do not provide adequate information for this type of health study. While some of the databases need to be and could be improved, researchers need greater sophistication about their limitations.

Yet another problem is the lack of nonexposed or less-exposed persons as controls for many studies. Often, regional rates were compared with national rates, which can give misleading results because the national rates do not take into account confounders (e.g., race or socioeconomic status) in the study population. Appropriate comparison groups must be used to correct for possible confounders.

The weakest aspects of most of these studies was their use of imprecise measures of exposure or small study populations, so type II errors may often occur (failing to reject the null hypothesis when it should be rejected). There were often unavoidable limitations in the size of the exposed population. Creative approaches (e.g., meta-analysis) may help to overcome this problem. Also, it may be possible to group related health outcomes in ways that increase statistical power. Many adverse reproductive outcomes are uncommon, so that even a large relative risk may be difficult to detect.

CONCLUSIONS AND RECOMMENDATIONS CONCERNING THE GRAY LITERATURE

We conclude that most studies and reports in the gray literature have serious limitations, such as lack of adequate exposure information, that

seriously undermine their credibility and value. However, the gray literature may contain studies and reports that point to directions for further research or that contain the only information on a topic. Used with caution, the gray literature may make a contribution to the study of effects of toxic substances. *We recommend that the studies and reports be collected, electronically listed, and made available in a central repository at ATSDR so that investigators can have access to them. The database should include gray-literature studies, whether positive or negative effects have been reported, and a list of studies in progress and completed.*

Because some of the gray literature in the United States provides useful information to communities, public-health officials, and researchers about specific exposures in local areas and about general health problems that may be associated with exposures from hazardous-waste sites, these studies need to be made more available to those interested in this field.

The committee recommends that ATSDR, perhaps with the National Library of Medicine or other agency, explore the feasibility of establishing an on-line database of completed state health-department studies in environmental epidemiology.

Failure to publish nonsignificant results poses a major obstacle to meaningful meta-analysis. A possible method for minimizing this bias is the prospective registration of research. Approaches to minimizing publication bias, including registration of studies, should be actively investigated and developed.

The committee recommends that ATSDR examine the merits of developing an electronic directory of on-going environmental-epidemiology studies that includes sufficient statistical results, if available, with the publication status of the study.

Although the limitations of studies in the gray literature may not be completely avoided, the committee believes that it would be useful to establish guidelines to improve the credibility and utility of the studies. We suggest the following as a starting point.

External scientific committees, advisory to state health departments or other sponsoring agencies, should help to promote better scientific approaches for studies.

Some of the gray-literature reports that the committee reviewed could have been published in the peer-reviewed literature, and hence made more readily available elsewhere, if the authors had invested the necessary effort.

Researchers working in state agencies should be given incentives to publish studies and to invest the time and effort needed to make these reports suitable for publication.

Many of the gray-literature studies used an ecologic design. For public-health agencies, there is an advantage to using readily available data quickly and inexpensively. The limitations of this approach may not be widely understood; many of the studies we reviewed had significant

shortcomings in the definition of the question to be addressed, study design, and its implementation.

More research and method development on the use of ecologic data are needed, as well as training for scientists using this approach.

Another common limitation in these studies was the use of exposure and health-outcome databases that do not provide adequate information for this type of health study.

Databases need to be improved, and researchers need to develop a greater degree of sophistication about their limitations.

CONCLUSIONS AND RECOMMENDATIONS ON METHODOLOGIC ISSUES

Epidemiologic research is often expensive and time-consuming, especially where longitudinal studies of large populations are involved.

Efforts to update previous prospective cohort studies should be encouraged, as these can provide cost-effective means to evaluate chronic diseases that could be linked with exposures that were previously characterized. Similarly, new analyses of old data should be generally supported, as they may yield important clues about the etiology of chronic diseases.

Common to all epidemiologic designs is the need to consider at the outset the chance of producing a positive finding. Statistical power, or the probability of finding a real effect, should guide any decision about whether to undertake the research. The higher the expected relative risk (RR), the smaller the population that needs to be surveyed. Conversely, the larger the population studied, the smaller the RR that can be detected.

In general, larger samples are needed when exposure measures are not continuous, when the effects of confounders and errors of measurement cannot be taken into account, and when the adverse outcome is a rare event. With many environmental exposures, such as those in air and domestic water, the attributable risk can be quite large. Even a small RR can be important if it applies to hundreds of thousands of people.

Cohort studies or even case-control studies large enough to detect the effects of usually low-level environmental exposures may be prohibitively costly in many cases. Ecologic studies, although they lack the detail available in studies that define individual exposures and are not often helpful in quantifying relations between exposure and disease, provide relatively inexpensive access to large populations. New methods for estimating exposure in ecologic studies may address some of the common concerns with ecologic analysis and improve its usefulness in environmental epidemiology.

Efforts to improve on existing methods for ecologic analysis and to make greater use of existing data sets should be encouraged. Recent advances in the

use of ecologic data have greatly improved the value of such information. Efforts should continue to refine these analyses.

Special emphasis should be given to establishing training programs in environmental-exposure assessment. Whatever data gaps exist in this area can be filled only when sufficient personnel and resources are used to conduct the needed assessments. We recognize that personal monitors provide sophisticated information about exposure. However, in designing studies with fixed budgets, researchers must weigh the benefits of improved exposure measures against the costs of reduced study power. In many circumstances, exposure assessment can be adequately estimated without personal monitors. Better use of activity logs and the gathering of more information about environmental conditions, such as sources of indoor heating and cooling, can also refine exposure assessments. The development of these and other techniques for advancing exposure assessment should be supported.

Health outcomes of interest to environmental epidemiologists range from syndromes or constellations of clinical measures of physiologic or neurobehavioral function to well-characterized diseases, such as cancer. Changes in patterns of these health outcomes can be effects of environmental factors. To ascertain whether changes in chronic diseases have occurred, researchers need to determine the expected, or baseline, rates of those diseases. This is difficult, especially for diseases other than cancer, because case definitions for various diseases and syndromes are not uniform or well defined, registries and codings are not uniform and consistent and may not even exist, other demographic variables can influence the outcome, and time trends may not be reliable, because of changes in case ascertainment.

We recommend that agencies concerned with the prevention of disease and the promotion of health make a deliberate effort to identify chronic diseases and syndromes that appear to be increasing or to have an important impact on public health. Concern should not be limited to fatal diseases. All forms of morbidity, ranging from ocular effects to impaired reproductive health, should be considered in epidemiologic studies. Symptoms, such as headaches and respiratory irritation, may be disabling and are appropriate for study.

Based on those strategies, registries should be established that can be connected with existing data systems, such as hospital records or health-administration information. Thus, the Health Care Financing Administration includes data on renal dialysis, hospitalization for asthma, and chelation therapy. Linking this information with better characterization of the cases might allow researchers to determine time trends in underlying diseases.

Where time trends reveal recent shifts in disease patterns, researchers should explore possible etiologic explanations. Health-information databases should contain variables that would facilitate linkage to exposure information and other

relevant factors to improve analyses of local patterns and relations among possible causal factors. Furthermore, the databases should include demographic information so that populations that are more susceptible to different environmental exposures, or more exposed, can be identified.

Notwithstanding the initial enthusiasm for their potential, biologic markers need to be validated as to their specificity for exposure and health outcome. Major gaps exist, as many of the markers that have been identified and characterized in the laboratory have not been specifically tied to chemical, biologic, or physical exposures. Validation of these will strengthen the ability of environmental epidemiology to discern relations between exposure and disease.

Almost no data systems focus exclusively on the environment-health relation. The committee identified existing data systems that could be improved in this respect. Human-health surveys—such as the National Health and Nutrition Examination (NHANES), the National Health Interview Survey (NHIS), and the National Vital Statistics Program—all have followup capabilities and should collect followup information on both exposure and health.

The federal government should coordinate the evaluation of existing data systems and data collection to develop modifications that might enhance the linkage and usefulness of data. One simple addition that could enhance the utility of NHANES and NHIS would be to obtain information on residential histories. Such an effort should involve experts from multiple agencies and from outside government. Emphasis should be placed on better documentation of the degree to which data systems represent or characterize a larger universe, such as a population or regional environment. Common definitions and specifiers of disease or health status, of geographic location, and of chemicals and other relevant data items would also enhance usability of these systems.

Entities outside government have a legitimate need for access to data collected by government agencies and for linking data from different data systems.

Use of data by outside researchers should be encouraged, under appropriate procedures to guarantee confidentiality. Because linkage of data often requires the use of confidential identifiers, agencies should where possible generate linked data tapes, stripped of identifiers.

Academic researchers often fail to recognize the multiplicity of questions and research designs that can be addressed with data systems. There is a tendency for academic researchers to discount data collected from regulatory-agency data systems. There is a tendency for policy-makers to consider the design or funding of data systems as though other data systems do not exist or as though regulation is the only purpose of a data system. The potential impact of the environment on health is too great to be approached piecemeal.

The federal government should establish mechanisms to track the health and

illness experience of populations for which data on exposures and other baseline measurements are available. For example, the National Death Index is an essential and widely used resource for the public-health community and should be maintained.

Wide application of statistical techniques to adjust for autocorrelation and covariance structures will enhance the ability of environmental epidemiology to identify exposure-related diseases. Statistical modeling (e.g., logistic regression and Poisson regression) is an effective tool, particularly for the multifactorial outcomes generally examined in environmental epidemiology. In carrying out such statistical modeling, one must recognize that relations between environmental toxicants and health outcomes may be nonlinear. Statistical models that implicitly assume linearity may miss important associations or important features of associations. Modern statistical-modeling techniques allow the investigation of nonlinear models.

Many preliminary environmental-epidemiologic studies rely on inadequate databases for exposures and health outcomes. Although some of the databases need to be improved, researchers also need to develop a greater degree of sophistication about their limitations and to refine their ability to analyze existing information.

To the extent that the estimates of gradients of exposure are improved, the ability to detect associations of exposure and response is also improved. Statistical models for improving exposure assessment can help with this process.

Although not treated in this report, meta-analysis is a potentially useful tool for the analysis of environmental-epidemiologic data. However, attention will need to be given to improving data collection so that results are amenable to meta-analysis and to the refinement of methods for performing meta-analyses.

Index

C

D

E

F

G

H

I

N

O

P

R

S