MOLECULAR EPIDEMIOLOGY AND THE PREVENTION OF DISEASE

Ellen K. Silbergeld
Professor of Epidemiology and Toxicology, University of Maryland Medical School, Baltimore, MD, USA

Keywords: Epidemiology, Molecular Epidemiology, Epidemiologic Methods.

Contents

Summary

Recent advances in molecular biology and genetics offer new tools for understanding patterns of health and disease in human populations. The field of epidemiology has made major contributions to the identification of preventable causes of human disease, but understanding the relationships between risk factors and outcomes has been limited by lack of knowledge of the events that take place between exposure and disease. In particular, it has been difficult to explain why individuals and populations respond differently to similar risks, such as smoking or chemical exposures. Biomarkers are measurable signals of these events, related to exposure and response, as well as inherent susceptibility of individuals and populations. Their use and interpretation depends on our understanding of the underlying biology related the disposition of dose, metabolic transformation of absorbed dose, and cellular response to doses at critical cellular sites of action. New technologies will increase the amount of information available on these parameters, and require advances in our ability to interpret data.

1. Introduction

The prevention of disease and the promotion of health require the combined expertise and involvement of many clinical and scientific disciplines, integrated into the epidemiological perspective. Epidemiology is the tool of public health, as compared to clinical medicine, since it is the study of health and disease in populations. Epidemiological studies can be assessments of risks and health at a point in time (cross sectional) or over time (either forward, or prospective, or backward, or retrospective). Epidemiological studies are designed to test the possible associations between one or more exposures or risk factors and health or disease status, for example, the number of

cigarettes smoked per year and lung cancer, or dietary fat intake and blood pressure.

It is often difficult to discern whether or not apparent associations are real, or the result of other factors and events that may not be measured or even be measurable in a specific study. Even in complex epidemiological studies, many unrecognized variables can affect associations or obscure the ability to detect true associations. Very few diseases - primarily hereditary diseases such as Huntington's disease - are associated with only one risk factor. Most diseases are multifactorial in origin, and only a few of the relevant factors may be known at any time. Moreover, human populations are complex groups of individuals, who may vary in many respects (or variables) not wholly knowable by investigators. For example, nutrition is known to influence the response of children to mercury exposures, but accurate nutritional information is difficult to collect without careful monitoring of food selection (duplicate diet studies) or detailed questionnaires on actual food consumption.

2. Goals of Epidemiology

The goal of epidemiological research is to identify causes of disease and to evaluate the efficacy of interventions to diagnose, treat, cure or prevent disease (Beaglehole et al 1993). Investigations, monitoring, observations, and experiments accomplish these goals. Epidemiologists *investigate* sudden changes in health, such as outbreaks of infectious disease or poisonings, as well as more gradual changes in the incidence (new cases) or prevalence (existing cases) of disease. Epidemiologists *monitor* the health status of populations through surveillance of events such as birth and death statistics or through the collection of data on certain diseases, such as AIDS or cancer. Epidemiologists design and conduct *observational* studies on the prevalence and incidence of diseases in population groups (cohorts) in order to discern risk factors associated with increases in risks or severity of disease. Epidemiologists conduct *experiments*, or clinical trials, to test the efficacy of interventions, such as dietary modification, early medical diagnosis or therapeutic treatment, in eliminating or ameliorating diseases.

In observational studies, in which exposures and outcomes are assessed as they occur in populations, the challenge of discerning true associations is especially difficult. Precise information on exposure may be difficult to define; even within the same job category, individual workers may have different exposures owing to the exact nature of their tasks, the efficacy of worker protection equipment, and personal habits. Even experimental studies can be complicated by unmeasured or unmeasurable variables: in clinical trials, the willingness of patients to participate or continue in a test of a therapeutic may reflect variables that also affect health, such as education and economic status in trials of dietary modification in preventing heart disease.

3. Epidemiological Methods

To minimize these complications, epidemiological research has relied on rigorous methods to design studies and analyze data, and to avoid the introduction of unintended bias into research. In experimental studies, to separate the effects of treatment from other influences, epidemiologists generally utilize a rigorous procedure of

randomization to minimize the impact of undetected differences between persons treated and untreated in a clinical trial. In observational studies, where the goal is to understand factors influencing health status among populations without the opportunity to control exposures or to preselect subjects, epidemiologists generally use one of two types of study designs: a case: control design, in which persons with a particular disease are matched with persons who do not have the disease, and the frequency or levels of exposure to the hypothesized risk factor are measured in each group. If the exposure is more frequent, or greater, in the population with the disease (cases), as compared to the controls, then an association is deemed to exist between the exposure and the outcome. The other type of study is a cohort study, in which two groups are selected based upon exposure. They may be evaluated at one point in time (cross sectional study); they may be followed in time to determine the eventual outcomes of exposure (prospective study), or their medical history may be evaluated (retrospective study). If the incidence or prevalence of disease is significantly greater in the exposed group as compared to unexposed groups, then an association may exist between the exposure and the outcome. The value of these studies depends on the ability to clearly define risk factors and outcomes, and to account for or control (through design or statistical analysis) the potential impact of other factors in the appearance of disease among the studied populations.

These methods have been sufficient to provide public health and medical authorities with powerful information to prevent risks and institute effective interventions, even in the absence of complete understanding of how these risks or treatments actually work to promote health or prevent disease. For instance, in the nineteenth century, the English physician John Snow utilized careful observation of the geographic and temporal distribution of cholera cases in London to identify a source of water as a cause of disease; based upon his study, this source was disconnected and the cholera outbreak abated (Stolley and Lasky, 1995). In the early twentieth century, data were collected on rates of cancer among workers in the new synthetic organic chemical industries, and these studies were instrumental in identifying several polycyclic aromatic hydrocarbons (aniline, benzidine, coal tars) as causes of human cancer. This work inspired effective programs in occupational health in many countries, including the continuing research and evaluation activities of the International Agency for Research on Cancer (IARC) of the World Health Organization. Bradford Hill and Richard Doll noted increases in lung cancer in England over the first half of the twentieth century and correlated these increases with the growing popularity of smoking. As a consequence, many governments required warning labels and restrictions on advertising of tobacco products. By the end of the twentieth century, Mary Claire King had utilized studies of breast cancer in families to correctly predict that a single gene was important in predisposing women to very high risk of breast cancer. Her research guided the development of methods to identify women at high risk, and the eventual identification of specific genetic polymorphisms, which now forms the basis of biological screening programs. All of these conclusions were drawn before knowledge of the pathobiology of cholera, or understanding of the mechanisms by which chemicals and tobacco smoke constituents initiate carcinogenesis, or the cloning of human breast cancer genes.

The success of these public health interventions, based upon epidemiology, was due to careful design and observation, along with rigorous data analyses. Without the tools to

investigate the actual events occurring between encounters with risk factors (exposure) and the incidence of disease (outcome), epidemiologists generally designed and analyzed data on the basis of the schematic shown in Figure 1.

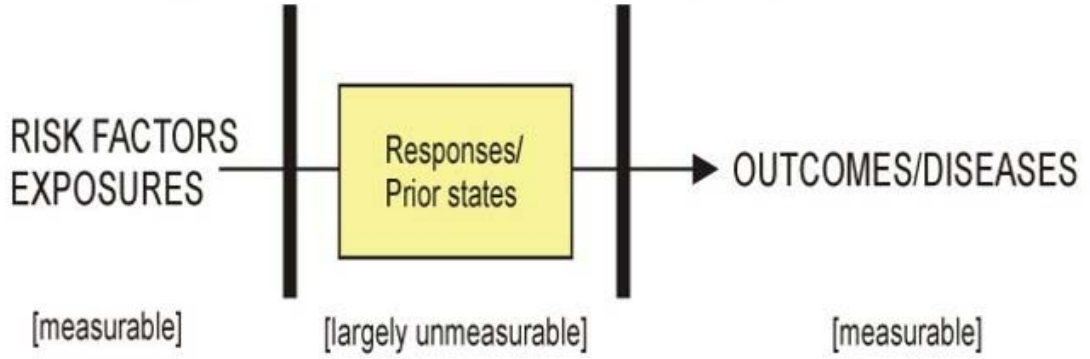

Figure 1: Traditional epidemiological model of the relationship between risk factors and diseases

-
-
-

Bibliography

Beaglehole R., Bonita R., and Kjellstrom T. (1993). *Basic Epidemiology*. Geneva: World Health Organization. [This is a highly accessible introduction to the methods and terminology of epidemiology, written for use internationally.]

Biggs P. J. et al. (1993). Does a genotoxic carcinogen contribute to human breast cancer? *Mutagenesis*, 8: 275–283. [This is one of the most compelling studies of the ways in which genetic analysis can help determine the origins of breast cancer.]

Coggon D. and Friesen M. D. (1997). Markers of internal dose: chemical agents. In R. Toniolo et al., eds. *Applications of Biomarkers in Cancer Epidemiology*. IARC Scientific Publication 142. Lyon: IARC. [This is an authoritative review of methods to evaluate human exposures to chemicals and other agents, for molecular epidemiology studies.]

Higginson J., Muir C. S., and Munoz N. (1992). *Human Cancer: Epidemiology and Environmental Causes.* Cambridge: Cambridge University Press. [This is the authoritative review of what is known of the causes of human cancer, based on international surveys of all the common types of cancer.]

Hulla J. E. et al. (1999). The role of genetic polymorphism and repair deficiencies in environmental disease. *Toxicol Sci*., 47, 135–143. [This paper presents the state of the art in our understanding of how genetic variation affects human response to chemical exposures.]

Ottman R. (1996). Gene-environment interaction: definitions and study designs. *Preventative Medicine*, 25, 764–770. [This presents a highly sophisticated analysis of the ways in which individual genetics can interact with environmental exposures, with models for developing studies and interpreting data.]

Perera F. (2000). Molecular epidemiology: on the path to prevention. *Journal of the National Cancer*

Institiute, 92, 602–612. [This is a compelling statement of the potential values of molecular epidemiology in identifying preventable causes of human disease, by one of the pioneers in the field.]

Silbergeld E. K. and Davis D. L. (1994). Role of biomarkers in identifying and understanding environmentally induced disease. *Clinical Chemistry*, 407, 1363–1367. [This paper reviews the strengths and limitations of using molecular markers to evaluate risks and outcomes.]

Stolley P. D. and Lasky T. (1995). *Investigating Disease Patterns: The Science of Epidemiology.* New York: Scientific American Press. [This is an exciting and vivid presentation of the history of epidemiology and the ways in which epidemiologists have contributed to preventing disease; from John Snow and the London cholera outbreaks to the identification of lead as a major risk to children's neurological development.]

Biographical Sketch

Ellen K. Silbergeld is Professor of Epidemiology and Toxicology at the University of Maryland Medical School. She received a Ph.D. in environmental engineering and postdoctoral training in environmental health sciences and neurotoxicology. She has held appointments at the National Institutes of Health, Environmental Defense, and Johns Hopkins. Her research has focused on mechanistic toxicology of metals and on molecular epidemiology related to reproductive disorders. In addition, she has served as a member of many national and international advisory committees and boards. She is author of over 250 scientific papers and reviews.