英文原版
医学教材

医学流行病学
Medical Epidemiology

Raymond S. Greenberg
Stephen R. Daniels
W. Dana Flanders
John William Eley
John R. Boring, III

LANGE

third edition

人民卫生出版社
McGraw-Hill

a LANGE medical book

Medical Epidemiology

third edition

Raymond S. Greenberg, MD, PhD
President
Medical University of South Carolina
Charleston

Stephen R. Daniels, MD, PhD
Professor
University of Cincinnati School of Medicine
Cincinnati

W. Dana Flanders, MD, DSc
Professor
Rollins School of Public Health
Emory University
Atlanta

John William Eley, MD, MPH
Associate Professor
Rollins School of Public Health
Emory University
Atlanta

John R. Boring, III, PhD
Professor
Rollins School of Public Health
Emory University
Atlanta

人民卫生出版社
McGraw-Hill

人民卫生出版社

McGraw-Hill

A Division of The ***McGraw·Hill*** *Companies*

图书在版编目(CIP)数据

医学流行病学/Raymond S. Greenberg主编. -北京: 人民卫生出版社, 2002
ISBN 7-117-04709-7

I.医... II.格... III.流行病学 IV.R18

中国版本图书馆CIP数据核字(2001)第098261号

Medical Epidemiology, Third Edition

Notice

Medicine is an ever-changing science. As new research and clinical experience broaden our knowledge, changes in treatment and drug therapy are required. The authors and the publisher of this work have checked with sources believed to be reliable in their efforts to provide information that is complete and generally in accord with the standards accepted at the time of publication. However, in view of the possibility of human error or changes in medical sciences, neither the authors nor the publisher nor any other party who has been involved in the preparation or publication of this work warrants that the information contained herein is in every respect accurate or complete, and they disclaim all responsibility for any errors or omissions or for the results obtained from use of the information contained in this work. Readers are encouraged to confirm the information contained herein with other sources. For example and in particular, readers are advised to check the product information sheet included in the package of each drug they plan to administer to be certain that the information contained in this work is accurate and that changes have not been made in the recommended dose or in the contraindications for administration. This recommendation is of particular importance in connection with new or infrequently used drugs.

图字: 01-2001-4004

医学流行病学(英文版)

主　　编: Raymond S. Greenberg
出版发行: 人民卫生出版社(中继线 67616688)
地　　址: (100078) 北京市丰台区方庄芳群园3区3号楼
网　　址: http: //www.pmph.com
E - mail: pmph@pmph.com
印　　刷: 三河市潮河印刷厂
经　　销: 新华书店
开　　本: 787×1092　1/16　印张: 14.5
字　　数: 502千字
版　　次: 2002年2月第1版　2002年2月第1版第1次印刷
标准书号: ISBN 7-117-04709-7/R·4710
定　　价: 31.50元

In loving memory of
John Rutledge Boring, Jr., June Gibson Daniels,
Hubert Wayne Eley, Sr., William Dana Flanders, Jr.,
and Bernard George Greenberg

Authors

John R. Boring, III, PhD
Professor and Chair, Division of Epidemiology, Rollins School of Public Health, Emory University, Atlanta; formerly Epidemic Intelligence Service Officer, Centers for Disease Control, Atlanta

Stephen R. Daniels, MD, PhD
Professor of Pediatrics and Environmental Health; Associate Chair, Department of Pediatrics, University of Cincinnati School of Medicine, Cincinnati

John William Eley, MD, MPH
Associate Professor, Rollins School of Public Health, Department of Epidemiology; Director, Institute for Cancer Prevention & Control; Assistant Professor, Winship Cancer Center; Assistant Professor, Department of Medicine, Emory University School, Atlanta

W. Dana Flanders, MD, DSc
Professor, Rollins School of Public Health, Emory University, Atlanta

Raymond S. Greenberg, MD, PhD
President, Medical University of South Carolina, Charleston

Preface

The authors of *Medical Epidemiology* are pleased to introduce the third edition of this book. As the field of epidemiology has moved forward over the past few years, it has been necessary to update this book accordingly. In particular, new material has been added relevant to evidence-based medicine. This material should help the reader become acquainted with the role of systematic reviews in the medical literature and their application in informing clinical decision-making. This new edition also provides the most current information available on patterns of disease occurrence and new self-assessment questions at the end of each chapter.

Although this book has been extensively revised, it remains faithful to the original intent of the authors—to provide a text that would serve the basic needs of medical students, as well as students of other health disciplines, such as public health, nursing, pharmacy, and dentistry. It is written concisely and can be used as a course textbook or as a stand-alone study guide.

OBJECTIVES

The aim of this book is to provide the reader with an overview of the principles and concepts of epidemiology. In so doing, it attempts to illustrate the complementary relationship between population-based science and the care of patients. Specific topic areas covered include:

- Measuring disease frequency
- Describing patterns of disease occurrence
- Investigating outbreaks of disease
- Assessing the utility of diagnostic tests
- Testing the effectiveness of treatments
- Identifying the causes of diseases
- Predicting the outcome of illness
- Decision-making about treatment strategies
- Summarizing evidence on clinical questions

Upon completion of this book, the reader should be able to calculate and interpret basic epidemiologic measures, recognize the various epidemiologic study designs with their respective advantages and limitations, understand the concepts of variability and bias, and characterize the means by which clinical evidence can be systematically summarized for decision-making.

APPROACH AND FEATURES

From the first edition of this book to the present version, the authors have taken the viewpoint that epidemiology should be both an understandable and interesting topic for students of the health professions. In order to introduce the topic in that manner, the following elements are emphasized:

- Conceptual topics are explained in nontechnical language.
- Liberal use is made of illustrations to facilitate comprehension and retention of material.
- The most current information available is presented on disease patterns and risk factors.
- The relationship between population-based science and patient care is demonstrated through patient profiles.
- A full range of clinical areas of application is shown, from infectious diseases to cancer to Alzheimer's disease to perinatal disorders.
- Critical formulas and equations are provided without undue emphasis on the mathematical applications of epidemiology.
- Questions are provided in standardized test format at the end of each chapter to help the student assess their knowledge and prepare for examinations.
- An updated glossary is provided to help the student master the vocabulary of epidemiology.
- Essential concepts are highlighted for emphasis and summarized at the end of each chapter.

If the response to the first two editions of this book is any indication, this approach is appealing to students of the health professions. The authors hope that the present edition continues to satisfy the demand for an engaging introductory text in epidemiology.

Charleston, South Carolina
November, 2000

Raymond S. Greenberg, MD, PhD

Acknowledgments

The third edition of this book was made possible through the dedicated efforts of many individuals, including Janet Foltin, Harriet Lebowitz, Arline Keithe, Charissa Baker, Laura Duprey, Minal Bopaiah, and Peter Boyle.

The comments and suggestions of several anonymous reviewers were helpful in refining the style and contents of this book. In addition, Dr. Beth Dawson and Dr. Paul Levy provided valuable advice in the initial developmental process. The didactic approach used in this book was developed largely through the experience of teaching epidemiology to medical students at Emory University and the University of Cincinnati. We have learned a great deal from these students, and we hope that their suggestions are adequately represented in the pages that follow.

The authors have had the good fortune to study under and to work with a number of outstanding epidemiologists. In writing an introductory text, we were inevitably drawn back to the teachers who first attracted us to the field. The influences of former mentors can be found throughout this book. Of particular note are the teachings of Dr. Philip Cole at the School of Public Health of the University of Alabama-Birmingham, Dr. Kenneth Rothman at Boston University, Dr. David Kleinbaum, formerly of the University of North Carolina and now of Emory University, and Dean Michel Ibrahim at the School of Public Health of the University of North Carolina.

The support and encouragement of our respective institutions were essential for the completion of this project. We thank Dr. Jeffrey Houpt, former Dean of the Emory University School of Medicine, for promoting strong educational and research linkages with the faculty of the Rollins School of Public Health. Dr. Charles Hatcher, Jr., Vice President for Health Affairs at Emory University, provided the resources and environment necessary for epidemiology and public health to develop at that institution. We also thank former Emory University President James T. Laney for his vision of the role of public health in serving human needs.

Ms. Essie Mills spent many long hours in the preparation of the manuscript. For her extraordinary tolerance in dealing with countless revisions, erratic work schedules, and urgent deadlines, the authors will remain forever in her debt. We are also grateful to Ms. Judy Holz for her help in revising the manuscript.

This book could not have been completed without the understanding and support of our wives and families. Time and again, precious hours at home were preempted by writing tasks, and we are grateful for the sacrifices that were made by our loved ones.

Charleston, South Carolina
November, 2000

Raymond S. Greenberg, MD, PhD

Contents

Introduction to Epidemiology 1

PATIENT PROFILE

A 29-year-old previously healthy man was referred to the University of California at Los Angeles (UCLA) Medical Center with a history of fever, fatigue, lymph node enlargement, and weight loss of almost 25 lb over the preceding 8 months. He had a temperature of 39.5°C, appeared physically wasted, and had swollen lymph nodes. Laboratory evaluation revealed a depressed level of peripheral blood lymphocytes. The patient suffered from simultaneous infections involving Candida albicans *in his upper digestive tract, cytomegalovirus in his urinary tract, and* Pneumocystis carinii *in his lungs. Although antibiotic therapy was administered, the patient remained severely ill.*

INTRODUCTION

Epidemiology is a fundamental medical science that focuses on the distribution and determinants of disease frequency in human populations. Specifically, epidemiologists examine patterns of illness in the population and then try to determine why certain groups or individuals develop a particular disease whereas others do not.

Knowledge about who is likely to develop a particular disease and under what circumstances they are likely to develop it is central to the daily practice of medicine and to efforts to improve the health of the public. To prevent an illness, health care providers must be able both to identify persons who, because of personal characteristics or their environment, are at high risk and to intervene to reduce that risk. This type of knowledge emerges in many cases from epidemiologic research.

This book serves as an introduction to epidemiologic methods and the ways in which these methods can be used to answer key medical and public health questions. This chapter begins by considering a single disease, as described in the Patient Profile. Focusing attention on one disease enables us to demonstrate the important contribution of epidemiology to current knowledge about this condition. Although the emphasis is on a single disease, it should be recognized that epidemiologic methods can be applied to a wide spectrum of conditions ranging from acute illnesses, such as outbreaks of food-borne infections, to long-term debilitating conditions, such as Alzheimer's disease.

The man in the Patient Profile was referred to the UCLA Medical Center in June 1981. At the time, there was no obvious explanation as to why a healthy young man would suddenly develop concurrent infections in three different organ systems involving three different microorganisms. More surprising was the nature of the infections that were present. Opportunistic infections, such as those caused by the parasite *P carinii,* are infectious illnesses that tend to occur only in persons with lowered resistance as may result from impaired immune responses. The young man described in the Patient Profile, however, did not have any obvious underlying causes of immune dysfunction. For example, he did not have cancer or severe malnutrition and he did not use immune-suppressing drugs. Why then was his body overwhelmed by the infections? This question was given a heightened sense of urgency by the severity of the patient's illness.

This patient was not the first to be referred to the UCLA Medical Center with this clinical presentation. Within the preceding 6 months three other previously healthy young men with recent histories of weight loss, fever, and lymph node enlargement had been examined. All had *P carinii* pneumonia and *C albicans* infections.

Why were four patients with similar symptoms appearing at about the same time in the same location? Suspecting that the illnesses in these four patients might be related, the UCLA physicians notified public health officials and prepared a descriptive report of their findings for publication.

Was this new appearance of a rare and life-threatening form of pneumonia confined to the UCLA Medical Center, or was it being observed by physicians elsewhere? If the experience at UCLA was unique, the entire episode might be regarded as a medical curiosity—unusual, but not a reason for great public health concern. On the other hand, if patients similar to those at UCLA were appearing in clinics or medical offices elsewhere, this episode could not be easily dismissed. Within a matter of weeks, public health authorities received reports of outbreaks of *P carinii* pneumonia among previously healthy young men in San Francisco and New York City.

In the United States, the federal agency that is responsible for monitoring unusual patterns of disease occurrence is the Centers for Disease Control and Prevention (CDC). Recognizing the potential for the widespread emergence of this new, unexplained, and debilitating condition, the CDC established a special task force to collect more detailed information on the affected persons. In addition, the CDC issued a formal request to report such patients to all state health departments. Between June and November 1981, 76 instances of *P carinii* pneumonia were identified in persons who did not have known predisposing illnesses and were not taking immune-suppressing medications. A few months later, the disease that afflicted these patients was named the acquired immune deficiency syndrome (AIDS).

PERSON, PLACE, & TIME

The physicians at UCLA played a crucial role in establishing the presence of a new disease in their community. The first few affected patients identified with any outbreak of disease are referred to as **sentinel cases.** The story of the first few AIDS patients is particularly dramatic because of the severity of the illness and the extent and speed with which the disease spread to others. A sudden and great increase in the occurrence of a disease within a population is referred to as an **epidemic.** It quickly became apparent, however, that the emergence of AIDS was not confined to a few communities. A rapidly emerging outbreak of disease that affects a wide range of geographically distributed populations is described as a **pandemic.** In 1981, no one could have predicted that by 1998 almost 650,000 persons in the United States would be diagnosed with AIDS and almost 400,000 deaths from AIDS would be reported nationally. By 1996, AIDS was the eighth most common cause of death in the United States and the third most common cause of death for persons between the ages of 25 and 44 years. With the introduction of effective combination drug therapy, the death rate from AIDS has declined in the United States and in other industrialized nations. In developing nations a much more devastating picture is emerging; for instance, in sub-Saharan Africa, it is estimated that over 7% of young adults have AIDS.

Looking back to 1981, when AIDS was not yet recognized as a clinical entity, it is instructive to consider the features of the sentinel cases that suggested a possible connection. All the patients with AIDS who presented to the UCLA clinicians suffered from the same rare opportunistic infections. Had the infections involved more conventional human pathogens—or less severe symptoms—the entire episode might have gone unnoticed for some time.

Beyond their clinical similarities, the sentinel cases shared other features, as summarized in Table 1–1. All four patients were previously healthy homosexual men in their early 30s (personal characteristics) who resided in Los Angeles (place) and first became ill in the 9 months ending in June 1981 (time). These three dimensions—**person, place,** and **time**—are the features traditionally used to characterize patterns of disease occurrence, as discussed in Chapter 3.

Table 1–1. Characteristics of sentinel cases of AIDS in Los Angeles, 1981.

Characteristics of Sentinel Cases	Personal Attributes
Age	Early 30s
Gender	Male
Prior health	Good
Sexual preference	Homosexual
Place of occurrence	Los Angeles
Time of occurrence	October 19, 1980 to June 19, 1981

THE EPIDEMIOLOGIC APPROACH

Epidemiology is concerned with the distribution and determinants of disease frequency in human populations. Interest in frequency or occurrence of disease derives largely from a basic tenet of epidemiology, ie, disease does not develop at random. In essence, all persons are not equally likely to develop a particular disease. The level of risk for different individuals typically is a function of their personal characteristics and environment.

As applied to the outbreak of AIDS, for instance, it is highly unlikely that of the first four cases in Los Angeles each would have occurred in homosexual males if the disease was striking at random. The repeated occurrence of AIDS in homosexual men suggested that this segment of the population had an increased risk of developing the disease. Other high-risk groups for AIDS, including hemophiliacs and injecting drug users, were soon identified. On the surface, these three groups seemed to have little in common. On closer examination, however, it became evident that an increased risk of exposure to the blood of other persons was the factor they all shared.

Contemporary medical research is devoted largely to investigating the biologic elements of disease development. For example, in the study of AIDS, a microbiologist tends to focus on the infectious agent, human immunodeficiency virus (HIV). An immunologist might concentrate on the primary target of HIV infection, the $CD4^+$ T lymphocyte, which coordinates a number of immune functions. The epidemiologist, on the other hand, views a disease from both a biologic and a social perspective. It is not enough to know that HIV is transmitted primarily through contaminated blood. The epidemiologist must be able to understand the circumstances of HIV transmission among humans. Here, the influence of social factors is undeniable. The spread of AIDS in human populations cannot be fully appreciated without recognizing the role of certain behaviors, such as sexual practices or injecting drug use.

The desire to study social factors that impinge on health has definite implications for how epidemiologic research is conducted. In most instances, this research involves observations of phenomena that occur naturally within human populations. Such an approach is unique among the medical sciences. Two features distinguish the epidemiologic approach from other biomedical sciences: (1) the focus on human populations and (2) a heavy reliance on nonexperimental observations.

At first, the focus on human populations may not seem distinctive. Ultimately, all medical research is motivated by a desire to prevent or control human illnesses. The process leading to that goal, however, may take various routes. Laboratory scientists, for example, often rely on experiments that involve nonhuman animals, cells in tissue culture, or biochemical assays. Although these studies offer important advantages to the investigator, such as precise control over the experimental conditions, certain limitations must also be recognized. Obviously, a laboratory environment may not accurately reflect the actual conditions of exposure in the external world. Of equal importance is the recognition that animals of different species may have dissimilar responses to experimental manipulations. It cannot be assumed that biological effects detected in rodents will necessarily apply to humans.

Epidemiologists avoid these concerns by attempting to study people directly in their natural environments. With this approach, it is not necessary to make assumptions about similarity of effects either across species or across doses and routes of exposure. The epidemiologist actually observes the patterns of exposure and disease development as they naturally occur within human populations. Without such information, it would not be possible to reach a definitive conclusion about the extent of disease related to a particular agent.

As with any scientific method, the epidemiologic approach has inherent constraints. In observational research, which comprises much of epidemiology, the investigator merely watches the phenomena under study, ie, the epidemiologist has no control over the events that occur. It is often difficult, therefore, to separate the causal contributions of the exposure of interest from the causal contributions of other background influences in the population. Even direct measurement of the degree of exposure may not be possible in some settings, thereby forcing the epidemiologist to rely on indirect estimates.

The epidemiologist's perspective of the relationship between exposure to risk factors and the development of disease in human populations may appear rather crude in comparison to the exacting research performed at the molecular level. Indeed, epidemiology is not particularly useful for characterizing the precise biologic mechanisms of disease development. The epidemiologist frequently sees only how different levels of exposure across groups of the population affect the comparative likelihood that those groups will develop disease. Typically, the epidemiologist can identify the personal, social, and environmental circumstances under which a disease tends to occur, without being able to explain the exact processes that give rise to the disease.

Medical progress often is best advanced when the sciences that focus on subcellular and molecular basic research work in tandem with the population-oriented science of epidemiology. For example, as bench scientists were struggling to characterize the molecular properties of HIV, epidemiologists already determined that AIDS is a contagious disease that is spread through certain interpersonal behaviors. As the painstaking search continues for improved treatment, or even a cure or vaccine, public health professionals have recommended measures to prevent the spread of HIV by reducing the frequency of the following high-risk practices: (1) casual, unprotected sex and (2) sharing needles among injecting drug users.

THE APPLICATIONS OF EPIDEMIOLOGY

Epidemiologic methods can be used for a number of distinct purposes. In the following sections, these areas of application are specified, with corresponding illustrations drawn from the literature on AIDS.

Disease Surveillance

Perhaps the most basic question that can be asked about a disease is "What is the frequency with which the disease occurs?" To answer this question, it is necessary to know the number of persons who acquire the disease (cases) over a specified period of time and the size of the unaffected population. Measures of frequency of occurrence of a disease, described in Chapter 2, are used to characterize the patterns of the occurrence of the disease, described in Chapter 3, and the medical surveillance of the disease, discussed in Chapter 4. Typically, the criteria used to define the occurrence of a disease depend on current knowledge about the disease; such criteria may become more refined as the causes of a disease are delineated and new diagnostic tests are introduced. For example, in 1982, the CDC created an initial, relatively simple surveillance definition for AIDS:

> A disease, at least moderately indicative of a defect in cell-mediated immunity, occurring in a person with no known cause for diminished resistance to that disease.

A more specific definition became possible once the causative agent, HIV, was identified and tests for the detection of antibodies to the virus were developed. In 1987, the CDC surveillance definition was expanded to incorporate clinical conditions that are indicative of AIDS. A 1993 revision further expanded the surveillance definition to include three additional indicator conditions (pulmonary tuberculosis, recurrent pneumonia, or invasive cervical cancer) or the presence of a severely depressed CD4+ T-lymphocyte count.

Such changes in diagnostic criteria can have a profound effect on the apparent frequency of a disease. The expanded definition of AIDS introduced in 1987 led to an increase in the number of reported AIDS patients by about 50% during the next 2 years. The 1993 revision more than doubled the number of persons who met the surveillance definition. Most of the latter increase was attributable to persons made eligible on the basis of reduced CD4+ T-lymphocyte counts and HIV infection. Accordingly, analysis of trends in disease occurrence over time must account for the possible effects of any temporal changes in diagnostic criteria.

The identification of patients with a disease can occur through various mechanisms, most commonly by physician and laboratory reporting. In the United States, a number of diseases, including AIDS, must be reported to public health authorities. Monitoring the patterns of occurrence of a disease within a population is referred to as **surveillance.** There are many potential benefits from the collection of surveillance data. This type of information (1) can help to identify the new outbreak of an illness, such as AIDS, (2) can provide clues, by considering the population groups that are most affected by the illness, to possible causes of the condition, (3) can be used to suggest strategies to control or prevent the spread of disease, (4) can be used to measure the impact of disease prevention and control efforts, and, finally, (5) can provide information on the burden of illness, necessary for determining health and medical service needs.

The course of the AIDS pandemic in the United States is depicted in Figure 1–1. To diminish the impact in changes of the surveillance definition over time, a single definition, the 1993 CDC version, was used throughout. From 1985 through 1992, there was an unrelenting rise in the number of newly reported cases. From 1993 through 1997, there was a gradual fall in the number of newly reported cases. It should be noted that the information in Figure 1–1 relates to **number** of newly diagnosed cases per year. Changes in the counts of new cases can be affected by a number of factors, including, among others, changes in the following:

(1) Frequency with which the disease occurs
(2) Definition of the disease
(3) Size of the population out of which the cases develop
(4) Completeness of the reporting of the cases.

With respect to point 2, the 1993 surveillance definition was used consistently for all years in Figure 1–1, minimizing any distorting influence of a change in definition of the disease over time. With regard to point 3, growth in the size of the population of the United States could not explain more than a trivial amount of the rise in the cases between 1985 and 1992. The national population grew at only about 1% per year, whereas the av-

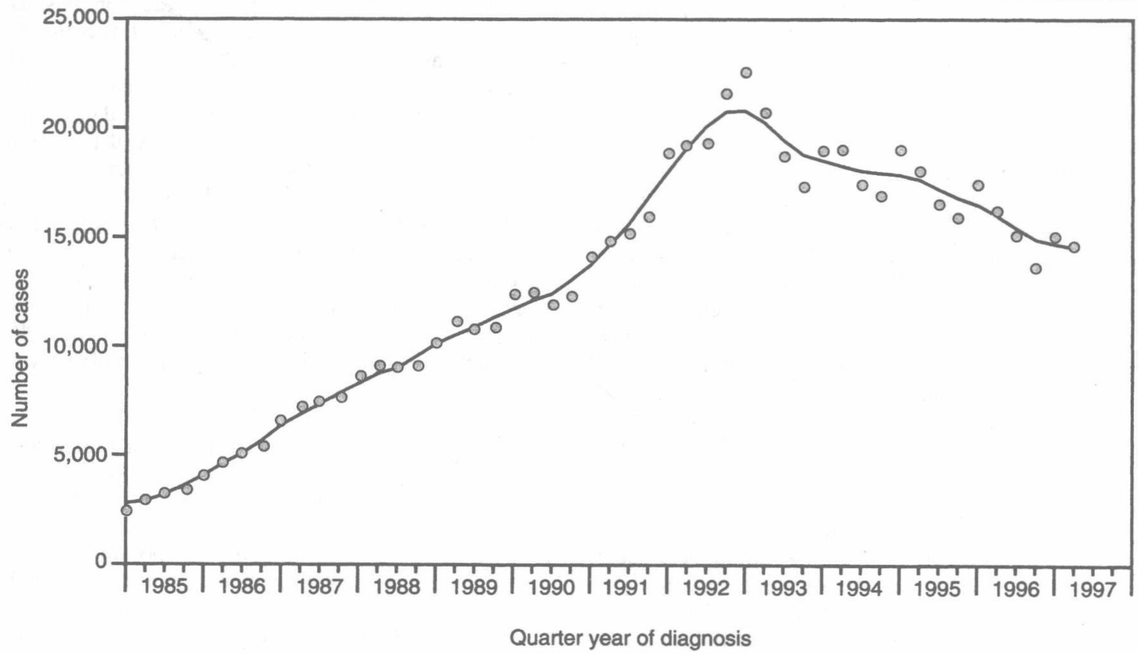

Figure 1–1. The number of cases diagnosed with AIDS in the United States using the 1993 CDC surveillance definition, after adjustments for reporting delays, 1985–1997 (Modified and reproduced from CDC: HIV/AIDS Surveillance Report 1997;**9**[No. 2]:1.)

erage annual increase in reported persons with AIDS exceeded 30%. Concerning point 4, the overall completeness of reporting of AIDS cases is estimated to be about 85% in the United States. Although there is some internal variation by geographic subregion and patient population, it is unlikely that these patterns could have given rise to more than a small part of the trend observed in Figure 1–1. Since items 2 through 4 do not appear to account for the marked changes in the annual numbers of reported AIDS cases, it is reasonable to conclude that the observed trend reflects a true change in the occurrence of the disease.

For surveillance purposes, the size of the source population from which cases arise usually is estimated from census data. The frequency of disease occurrence is then expressed as the number of new cases developing within a specified time among a standard number of unaffected individuals. For example, during 1997 over 60,000 cases of AIDS were reported in the United States; the U.S. population in 1997 was almost 270,000,000. Dividing the number of reported cases by the size of the population yields 0.00022 cases per person during that year. For ease of communication, epidemiologists typically express such frequencies of disease occurrence for a population of a specified size, say 100,000 persons. By multiplying 0.00022 by 100,000, the number 22 is obtained. That is to say, within a standard population of 100,000 persons in the United States, 22 persons would have been reported as developing AIDS during 1997. This measure of the rapidity of disease occurrence is referred to as an **incidence rate.** More information on incidence rates is presented in Chapter 2.

To characterize patterns of disease occurrence, incidence rates may be determined for groups defined by geographic area. For example, in Figure 1–2 annual incidence rates for AIDS are presented by place of residence in the United States. During 1997, the incidence rate for the District of Columbia was the highest observed, with 188.7 reported cases for every 100,000 residents. At the other extreme, South Dakota experienced the lowest annual incidence rate (1.5 cases per 100,000 residents). In other words, AIDS occurred in the District of Columbia over 125 times (188.7/1.5 = 126) more frequently than in South Dakota. Why are persons in the District of Columbia so frequently diagnosed with AIDS; conversely, why are persons in South Dakota so infrequently affected?

Answers to such questions typically do not derive from surveillance information alone. Surveillance data

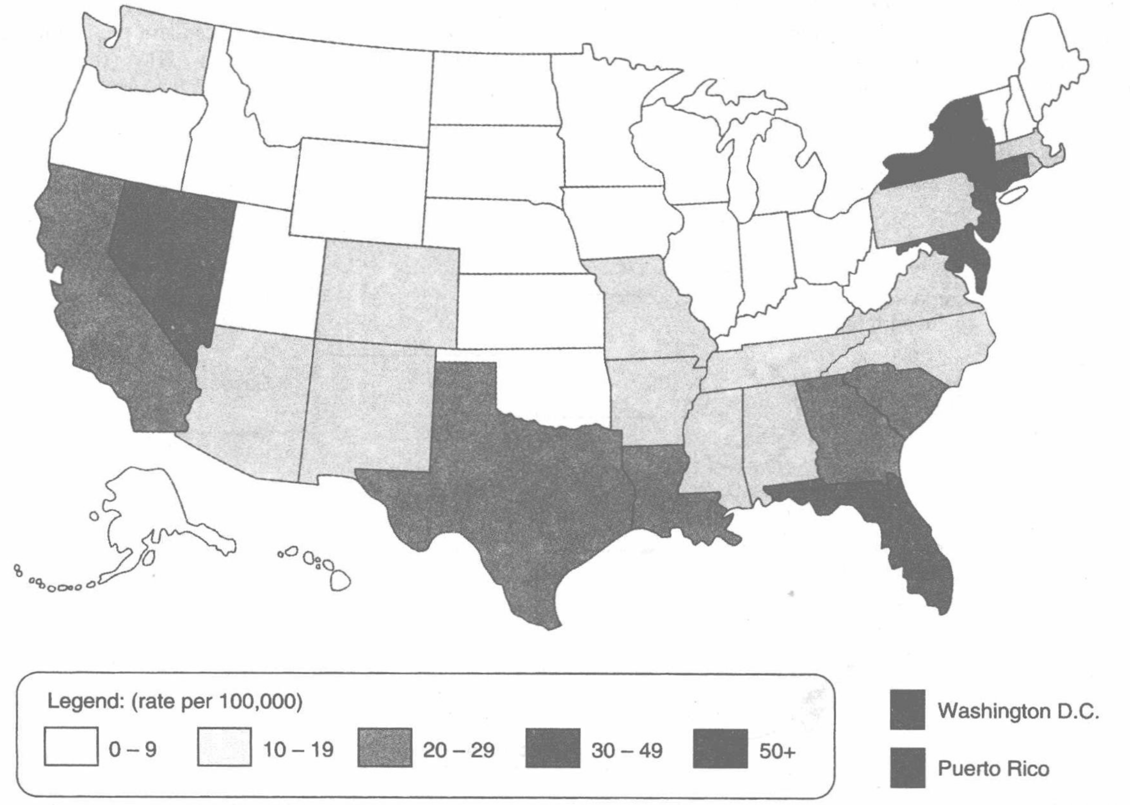

Figure 1–2. The incidence rates of AIDS per 100,000 person-years in the United States, 1997 (Modified and reproduced from CDC: HIV/AIDS Surveillance Report 1997;**9**[No. 2]:1.)

usually are limited to general characteristics of affected persons, such as their age, race, sex, and place of residence. Although variations in incidence rates according to these demographic features can lead to the identification of high-risk groups, explanations for these patterns generally require more in-depth investigation into personal characteristics, behaviors, and environments.

Searching for Causes

To study personal and environmental characteristics, epidemiologists often rely on interviews, review of records, and laboratory examinations. Through such sources of information, a profile of characteristics that accompany the disease can be generated. Associations between these characteristics and the occurrence of disease can arise by coincidence, by noncausal linkages to other features, or by cause-and-effect relationships. Of course, the epidemiologist is primarily interested in the last category, ie, determinants of disease development, also known as **risk factors.** Identification of risk factors can result in a better understanding of the pathways leading to disease acquisition and, consequently, better strategies for prevention.

Again, returning to the AIDS example, early epidemiologic studies played an important role in determining the cause of this disease. Within the first 5 months after recognition of this syndrome, the CDC had received reports on 70 patients with AIDS in four urban centers. Of these individuals, 50 homosexual male patients with AIDS were interviewed; also interviewed were 120 unaffected homosexual male comparison subjects. Persons who are affected with a disease are referred to by epidemiologists as **cases,** and unaffected comparison persons are called **controls.** Comparison of the responses from cases and controls revealed that the AIDS patients had a higher number of sexual partners. This type of investigation is referred to as a **case–control study;** the basic design of such a study is illustrated in Figure 1–3.

In essence, this study is an attempt to look backward in time to identify characteristics that may have contributed to the development of the disease. The increased number of sexual partners—as well as a greater frequency of syphilis among cases—suggested that AIDS resulted from a sexually transmitted infectious agent, later discovered to be the HIV virus. Case–control studies are described in Chapter 9.

Comparison of historical exposures reported by cases and controls can provide suggestive evidence of a cause-and-effect relationship. This type of information, however, may be distorted or **biased** by the fact that the ability of cases and controls to recall earlier exposures differs. Such bias could be avoided by using a **cohort study** design in which exposure is assessed among unaffected persons, and subjects are then observed for subsequent development of illness. To collect such data, a cohort of 2507 homosexual men without antibodies to HIV (seronegative) was questioned about their sexual practices and then followed for development of antibodies to HIV (seroconversion). Within 6 months, 95 men (3.8%) seroconverted,

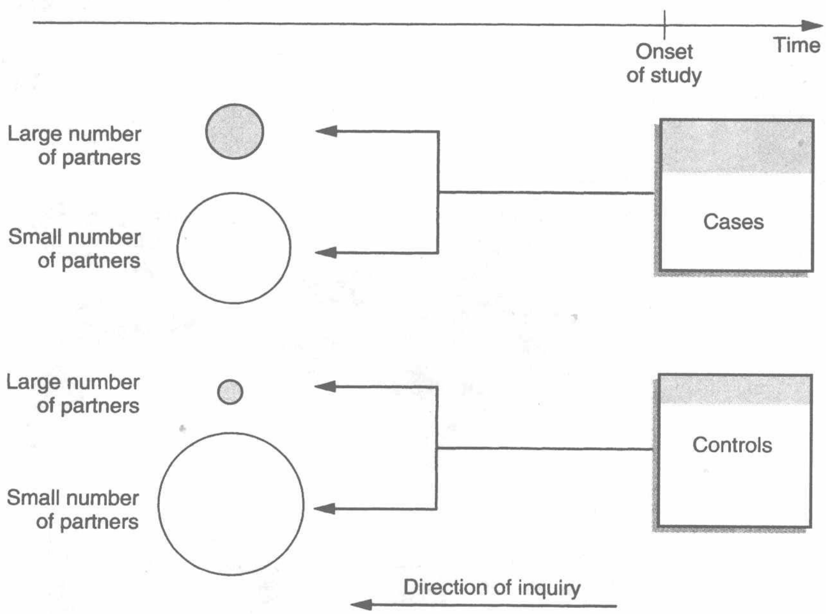

Figure 1–3. Schematic diagram of a case–control study of the association between the number of male sexual partners of homosexual men and the risk of AIDS. Shaded areas represent subjects with a large number of sexual partners and unshaded areas represent subjects with a small number of sexual partners.