INTRODUCTION TO MODERN EPIDEMIOLOGY

Anders Ahlbom

Staffan Norell

demos logos

Foreword by

Kenneth Rothman

Introduction To
MODERN EPIDEMIOLOGY

Second Edition

ANDERS AHLBOM, Ph.D.
STAFFAN NORELL, M.D., Ph.D.
National Institute of Environmental Medicine
Stockholm, Sweden

Epidemiology Resources Inc.
1990

Fourth Printing

Library of Congress Cataloging-in-Publication Data:

Ahlbom, Anders.
[Grunderna i epidemiologi. English]
Introduction to modern epidemiology / Anders Ahlbom, Staffan Norell.—2nd ed.
p. cm.
Translation of: Grunderna i epidemiologi.
Includes bibliographical references.
Includes index.
ISBN 0-917227-06-9
1. Epidemiology. I. Norell, Staffan. II. Title.
RA651.A6613 1990
614.4—dc20
DNLM/DLCL 90-3243

Epidemiology Resources Inc.
One Newton Executive Park
Newton Lower Falls, MA 02162-1450

Translated from Swedish by Gunilla Ahlbom.

Cover Design: Sylvia Steiner

Printed in the United States of America

Introduction To

MODERN EPIDEMIOLOGY

Foreword

An old joke had it that an epidemiologist was a physician who could count. Nowadays epidemiology has changed; no longer the province of a few dedicated physicians, it has evolved into a distinct research discipline. Like many other disciplines, epidemiology demands of its aspiring students years of advanced training if state-of-the-art proficiency is to be achieved. During the past 20 years, the principles that underlie the conduct and interpretation of epidemiologic research have developed rapidly from a set of loosely understood common-sense rules of thumb into a body of logically related formal concepts that offer a coherent theory to guide the planning and interpretation of epidemiologic work. Whereas intuition and experience formerly dominated epidemiologic thinking, now a more clearly defined set of principles has been crystallized. These principles can steer the epidemiologist through the tangled problems of research with human populations. Common sense and experience may still be reliable resources for a researcher, but science progresses by formalizing understanding. In this regard the progress of epidemiology over the last two decades has been remarkable.

What is seen clearly to be common sense by experienced scientists may not be perceived in the same way by those new to a field. Students who inherit a set of formal ideas as the foundation for learning have the privilege of developing an advanced understanding of concepts without needing to repeat the mistakes of others. Unfortunately, many of the concepts of modern epidemiology have not, until recently, found their way into introductory books.

The publication in 1984 of this lucid monograph by Anders Ahlbom and Staffan Norell corrected that deficiency. Here, clearly put forth, are the core ideas that a beginner to the field will encounter and need to master. The presentation of newer concepts has been smoothly woven into the existing dogma, enabling a deeper level of understanding for the reader than was previously possible at the introductory level. For example, when I first studied epidemiology, my classmates and I were taught to distinguish between incidence and prevalence, but we never heard about the fundamental distinction between the two types of incidence measures, incidence rate and cumulative incidence. The use of stratification in data analysis to control confounding was presented to us then only in an advanced course. In this small volume, Ahlbom and Norell have put forth for the first time at the introductory level a clear description of these and related concepts.

Counting is still important in epidemiology, but understanding and interpreting the epidemiologic studies appearing in today's medical and public-health journals require a bit more. This new edition preserves the clarity and conciseness of the first edition. The biggest changes are in the chapters on study design, the basic principles of data analysis, and stratified data analysis. These topics have been expanded in keeping with the notion that a clear understanding of methods to control confounding in the analysis of epidemiologic data is essential even for an introductory grasp of epidemiologic research principles. The revised text remains true to its original purpose — a concise description of the core ideas underlying epidemiologic research. It will be a valuable springboard as well as a useful reference for those who desire to acquaint themselves with modern epidemiology.

Kenneth J. Rothman
Editor, **Epidemiology**
February, 1990

Preface

This introductory text covers material that was developed, for the most part, during the last two decades. To a large extent this development was inspired by the work of the Department of Epidemiology of the Harvard School of Public Health. Although many of the building blocks were at hand already, we believe that the conceptual framework became much better defined in the course of that work, and also that a consistent theoretical framework and methodological structure was created that had not been available before. In our view it seems justified to refer to the theoretical and methodological system so developed as a new paradigm replacing the "person, place and time" paradigm.

This textbook was first published in Swedish; in Sweden it has been used mainly in medical schools, but also in some other introductory classes. The book has been well received by reviewers and students who appreciated our attempt at a concise and precise format.

We discussed the ideas behind the book with Kenneth Rothman and Nancy Dreyer, who told us that there was a niche for this kind of text in English as well as Swedish, and they suggested that the book be translated. We have enjoyed working on this project and we are grateful to Drs. Rothman and Dreyer for their encouragement and support during the process; without them this book would never have been published.

Anders Ahlbom
Staffan Norell
Stockholm
February, 1984

Preface to the Second Edition

After only five years it has become clear that this text needs revision, a result, at least in part, of the continuing rapid development of epidemiologic theory. That development together with a growing concern about public health continues to make epidemiology a highly interesting and challenging area.

The first six chapters of the book have undergone smaller, cosmetic changes, while the last chapters, on study design and data analysis, have seen major face lifts. In the study design chapters the changes reflect an increasing comprehension of the similarities rather than the differences between the cohort and the case-control study. In the data analysis chapters the development of personal computers and epidemiologic software has been important for the new text.

As with the previous edition we wish to thank Nancy Dreyer and Kenneth Rothman for encouragement and valuable criticism. We are also in debt to Sander Greenland for extensive and constructive criticism and to Gunilla Ahlbom for translating the Swedish edition into English.

Anders Ahlbom
Staffan Norell
Stockholm,
December 1989

Introduction To

MODERN EPIDEMIOLOGY

Table of Contents

1. What Is Epidemiology?

Epidemiology is the science of occurrence of diseases in human populations. Disease occurrence is measured and related to different characteristics of individuals or their environments. (The word epidemiology consists of the Greek words *epi* = among, *demos* = people, and *logos* = doctrine, and thus means the doctrine of what is among or happening to people.) Investigation of disease occurrence is not a new phenomenon. The development of epidemiologic theory and methods in recent decades, however, has opened new possibilities and stimulated interest within many fields of application.

For a long time the predominant interest in epidemiology was the area of infectious diseases. The occurrence of highly contagious infectious diseases varied in obvious ways, and often increased dramatically during so-called epidemics. It was found that individuals who had been in contact with sick people often became ill themselves and that those who recovered seldom got sick again. Such epidemiologic observations became the basis of theories about infectiousness and immunity—and suggested effective means to prevent diseases—even before microorganisms and antibodies were discovered. One well known example is the classic study on cholera in London that was conducted by John Snow in 1854 (Snow 1855).

Early epidemiologic observations were not limited to infectious diseases; other diseases also displayed variation in their occurrence. The distribution of different malnutritive diseases was studied early in this century and related to certain characteristics of food composition. Even before essential nutrients, such as certain vitamins, had been identified, theories about the causes of malnutritive diseases were formulated, preventive means undertaken, and sick people successfully treated. Studies on the distribution of pellagra undertaken by Goldberger between 1915 and 1926 (Terris 1964) form a good example of this process.

During the last few decades increasing attention has focused on the epidemiology of malignant diseases. Epidemiologic studies contributed decisively to understanding the role of cigarette smoking in the occurrence of lung cancer. Other studies have shown that there is an association between exposure to some types of ionizing radiation and certain forms of cancer. Many epidemiologic studies have demonstrated the connection between exposure to certain chemical substances and some kinds of malignant tumors. Although knowledge about the etiologic mechanisms for these diseases is still rudimentary, epidemiologic in-

vestigations occasionally have provided a sufficient lead for the implementation of preventive measures.

Another current field of application of great importance is cardiovascular disease. In this century myocardial infarction has become a leading cause of death in the industrialized world. One plausible explanation for this increase is the profound change in what has come to be known as "lifestyle." The role of various components of lifestyle is not clear, however. We still lack fundamental knowledge on the relative importance of factors such as stress, limited physical activity, smoking, high intake of calories and high proportion of saturated fats, and we do not know what the relation is between these characteristics and elevated blood pressure, serum cholesterol and triglycerides (blood fats). In recent years, a large number of epidemiologic studies have evaluated the role of these and other characteristics in causing myocardial infarction to clarify ways in which the disease can be prevented. With similar questions in mind, other vascular diseases such as stroke have also been studied.

In the past, attention has been focused mainly on diseases with short duration (e.g., acute infectious diseases), while in more recent years the focus has been increasingly on chronic diseases. Chronic diseases are of great importance because they represent longstanding suffering for many people and a considerable burden to the health-care system. An example is joint disease, such as rheumatoid arthritis. Even after corrections for differences in the age and sex distribution, substantial differences remain in the frequency of rheumatoid arthritis between different populations. Epidemiologists are now asking, "Which characteristics among the individuals (e.g., genetic) or their environment (e.g., exposure to infectious agents) explain these differences in morbidity from rheumatoid arthritis?" Another example is intestinal diseases such as ulcerative colitis and Crohn's disease. What affects the occurrence of these diseases and the risk of complications (e.g., colon cancer)? Other epidemiologic studies have focused on variations in the frequency of birth defects and the impact of factors such as smoking, alcohol consumption, drug use, and infections during pregnancy.

Sometimes the starting point for an epidemiologic study is a certain characteristic or exposure rather than a disease. In studies of occupational hazards the starting point often is a characteristic of the occupational environment or work place; the effect of the exposure can be evaluated by measuring the health status or the frequency of disease occurrence in the occupational group and comparing it with a suitable reference group. For instance, how does exposure to asbestos in certain occupations affect the occurrence of diseases such as asbestosis, mesothelioma, and lung cancer? What is the relation of irregular working hours or occupational stress to health? What are the health hazards associated with mercury pollution from industries or use of insecticides in agriculture? In studies of

side-effects of drugs the starting point again is a certain characteristic (drug exposure) rather than a disease.

These examples of fields of application suggest a close connection between epidemiology and preventive medicine. Prevention programs are rarely implemented for an entire population; therefore, prevention programs may be planned to enable studies of the effect of the intervention on the disease frequency in the population by comparing disease rates among those receiving the preventive program with rates among those who do not. In this way there is usually an opportunity to evaluate preventive measures that have been undertaken, using *experimental epidemiology*.

In recent years the value of information about disease distribution for planning the delivery of health care has become more apparent. In several studies disease occurrence has been related to health-care need, demand, and supply. There is also an increasing interest in studying the effectiveness of the health-care system and/or of different treatments.

The common basis for these different applications of epidemiology is the study of disease occurrence and its relation to various characteristics of individuals or their environment.

Additional Reading

Breslow NE, Day NE: Statistical Methods in Cancer Research. Volume I—The Analysis of Case-control Studies. International Agency for Research on Cancer 1980. IARC Scientific Publications No. 32.

Breslow NE, Day NE: Statistical Methods in Cancer Research. Volume II—The Design and Analysis of Cohort Studies. International Agency for Research on Cancer 1987. IARC Scientific Publications No. 82.

Checkoway H, Neil EP, Crawford-Brown DJ: Research Methods in Occupational Epidemiology. Oxford University Press, Oxford 1989.

Greenland S (ed): Evolution of Epidemiologic Ideas: Annotated Readings on Concepts and Methods. Epidemiology Resources Inc. Chestnut Hill, 1987.

Kelsey JL, Thompson WD, Evans AS: Methods in Observational Epidemiology. Oxford University Press, Oxford 1986.

Kleinbaum DG, Kupper LL and Morgenstern H: Epidemiologic Research. Principles and Quantitative Methods. Wadsworth, Inc. 1982.

Lilienfeld AM, Lilienfeld DE: Foundations of Epidemiology. Second edition. Oxford University Press 1980.

MacMahon B, Pugh TF: Epidemiology Principles and Methods. Little, Brown and Company 1970.

Miettinen OS: Theoretical Epidemiology. John Wiley & Sons, New York 1985.
Monson RR: Occupational Epidemiology. CRC Press, Inc. 1980.

Rothman KJ: Modern Epidemiology. Little, Brown & Co, Boston 1986.

Schlesselman J: Case-control Studies. Oxford University Press 1982.

Snow J: On the mode of communication of cholera. Churchill 1855. *In*: Snow on Cholera. Commonwealth Fund 1936. Reprinted by Hafner Press 1965.

Terris M: Goldberger on Pellagra. Louisiana State University Press 1964.

2. Measures of Disease Occurrence

The previous chapter indicated that the objective of epidemiologic studies is to learn about the occurrence of diseases. Measures of disease occurrence are therefore central to all epidemiologic activity. Such measures can be formulated in a variety of ways.

Absolute numbers and numbers related to size of the population

To start with, measures of disease occurrence should generally be independent of the size of the population. To accomplish this, the number of cases of disease is related to the number of individuals in the population. For some administrative purposes the absolute number of cases may be relevant, but for most analytical purposes the size of the population that gives rise to the cases has to be taken into account.

Example: In a campaign aimed to encourage the use of life vests, some epidemiologic data were provided. Out of 125 people who drowned, only 11 used life vests, while 114 did not. These data were presumed to suggest an association between use of life vests and risk of drowning. The comparison, however, involves two absolute numbers of cases; the sizes of the two populations that gave rise to these drowning victims, those who wore life-vests and those who did not, have not been taken into account. The observed discrepancy might well reflect only that the size of the population of life-vest users is small compared with the size of the population of non-users.

Incidence and prevalence

Measures of disease occurrence can describe either the pool of existing cases, or the occurrence of new cases. Measures of **prevalence** describe what proportion of the population has the disease in question at one specific point in time. Measures of **incidence**, on the other hand, describe the frequency of occurrence of new cases during a time period. It is useful to think of each individual as being in one of two "states": diseased or disease-free. In this framework the prevalence measure describes the proportion of the population that is in the diseased state at a specific time. The incidence measure describes the rate of flow from the disease-free state to the diseased state.

The magnitude of disease prevalence obviously depends on incidence, since a greater rate of occurrence of new cases will tend to increase the number of existing cases; but it also depends on the duration of the disease. Thus, a change in prevalence may be an effect either of a change in incidence or a change in duration of the disease. The duration in turn depends upon the time it takes to get well or the survival time with the disease.

In epidemiologic studies where the aim is to explore causal theories or to evaluate effects of preventive means, the interest is focused on the rate of flow of cases from the disease-free state to the diseased state. The relevant measure of disease occurrence, therefore, is incidence. Measures of prevalence may be relevant in connection with the planning of health services or in assessing the need for medical care in a population. Occasionally, the choice between incidence or prevalence is made for pragmatic reasons. For example, studies of chronic diseases such as rheumatoid arthritis or diabetes, in which the point of transition from non-diseased to diseased occurs gradually, and the definition of disease is arbitrary, generally employ prevalence measures, whereas studies on cancer or myocardial infarction generally use incidence measures.

Three measures of disease occurrence

Three specific measures of disease occurrence will be presented. The first is a prevalence measure and the other two are incidence measures.

Prevalence:

The prevalence measure is called simply the "prevalence" (P) (other terms in use include prevalence rate and prevalence proportion). The prevalence is defined as:

$$P = \frac{\text{number of individuals having the disease at a specific time}}{\text{number of individuals in the population at that point in time}}$$

The prevalence corresponds to the proportion of the population that has the disease at a certain point in time. Like all proportions it is dimensionless and can never take values less than 0 or greater than 1.

Example: A sample including 1,038 women age 70–74 years was selected from the population of Stockholm (Allander 1970). After examination, 70 were classified as having the diagnosis of rheumatoid arthritis. The prevalence of rheumatoid arthritis was

$$P = \frac{70}{1{,}038} = 0.07 \text{ for women age 70–74.}$$

Cumulative Incidence:

The next measure to be defined is called the "cumulative incidence" (CI) (other terms are cumulative incidence rate and incidence proportion). The definition is

$$\text{CI} = \frac{\begin{array}{c}\text{number of individuals who get the disease}\\ \text{during a certain period}\end{array}}{\begin{array}{c}\text{number of individuals in the population}\\ \text{at the beginning of the period}\end{array}}$$

Both numerator and denominator include only those individuals who at the beginning of the period are free from the disease and therefore at risk to get it. The cumulative incidence is, therefore, the proportion of individuals in the disease-free state at the beginning of the period that move to the disease state during the period. That is, the numerator is a subset of the denominator. Simply stated, the cumulative incidence is the proportion of healthy individuals who get the disease during a certain period. Alternatively, it can be viewed as the average risk for the individuals in the population to get the disease during that period. Being a proportion, the cumulative incidence is dimensionless and can only take numeric values in the range from 0 to 1. (Occasionally the cumulative incidence is defined in a theoretical way with a slightly different numerator, namely the estimated number of individuals who would have developed the disease if nobody had died during the interval from other diseases.)

In some studies different subgroups of the study population are considered at risk of getting the disease during different periods of time and individuals are considered at risk for periods of varying lengths. This variation in the risk period derives from the fact that different individuals enter the study at different points in time or that some migrate during the observation period. In such situations the cumulative incidence may not be directly calculable from the data. The length of the observation period directly affects the cumulative incidence: the longer the period, the greater the cumulative incidence. An extreme example would be the study of total mortality among newborn infants over, say, the ensuing 115 years. The cumulative incidence would always be 100% although the timing of the deaths could vary considerably between populations. The length of the period at risk must therefore always be reported along with the cumulative incidence and taken into account in interpreting any reported value of cumulative incidence. The mortality from other, competing, causes of death also influences the cumulative incidence as defined here; some of those dying from other causes would have been expected to develop the disease under study had they not died.

Example: The Swedish census from 1960 showed there were 3,076 males age 20–64 who were employed as plastic workers. According to the Swedish Cancer Environment Registry, 11 of those workers developed brain tumors during the