Alzheimer's Disease

Alzheimer's Disease

Life Course Perspectives on Risk Reduction

Amy R. Borenstein, Ph.D., M.P.H.
and
James A. Mortimer, Ph.D.

University of South Florida
Department of Epidemiology and Biostatistics
Department of Neurology
Tampa, FL, USA

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Academic Press is an imprint of Elsevier

Academic Press is an imprint of Elsevier
125 London Wall, EC2Y 5AS.
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, USA
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

ISBN: 978-0-12-804538-1

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress.

Publisher: Nikki Levy
Acquisition Editor: Emily Ekle
Editorial Project Manager: Timothy Bennett
Production Project Manager: Julie-Ann Stansfield
Designer: Victoria Pearson Esser

Typeset by MPS Limited, Chennai, India

Dedication

For my brave and loving parents, Lucia and Israel Borenstein, who encouraged me to do anything I wanted to do. I also dedicate this book to my loving and brilliant husband who is the best colleague I could hope to have.

—A.R.B.

For my intellectual and life partner, Amy Borenstein.

—J.A.M.

Preface

Alzheimer's disease (AD) is rapidly becoming a major public health challenge in the United States and around the world. The largest segment of the US population, the baby boomers, are just beginning to enter the age at high risk for AD. If effective methods for preventing this illness are not found and implemented, this demographic transition will lead to an epidemic of AD with huge economic, personal, and societal costs in the coming decades. Although the epidemiology of AD is still relatively young compared with the epidemiology of cardiovascular disease and cancer, knowledge about the risk factors for AD, and ways in which it might be prevented has exploded in the last 30 years.

The aim of this book is to provide students, researchers, clinicians, and the general public with a critical review of what is currently known about the distribution of AD in populations, its risk factors, and promising avenues for its early detection and ultimately its prevention. Given the enormous number of studies that have been published in this area, no review can claim to be comprehensive. We have endeavored to focus on epidemiologic investigations that provide the most valid and reliable data, while also including provocative viewpoints based on animal and human studies. Because we want the material in this book to be accessible to a more general audience, we begin with a primer on epidemiologic concepts to familiarize nonepidemiologists with the basic tools of the field, including the types of study designs that are used in epidemiology, the interpretation of study data, and methodological problems that may be encountered.

Although scientists continue to debate when AD begins, it is clear that genes are critical to its causation. It also is becoming clear that factors encountered in early- and midlife can either hasten or delay the appearance of clinical symptoms. Because the seeds of this illness are present at conception and risk is modified throughout life, we will take a life course approach in describing the risk and protective factors for the disease. The prevention of AD lies in our ability to either slow down the rate of accumulation of pathology through biologically modifying pharmaceutical drugs or by delaying onset of clinical symptoms through brain-healthy modifiable behaviors.

The book is divided into five main sections. In the first section, we begin with a description of the first recognized case of the disease as described by Dr Alois Alzheimer at the beginning of the twentieth century. This is followed by chapters on the clinical appearance, progression, diagnosis, and neuropathology of AD. The final chapter in this section considers the threshold model

that provides the context for interpreting the association of brain pathology and reserve with the clinical manifestation of dementia. In Section 2, we summarize the descriptive epidemiology of AD with regard to its prevalence, incidence, survival, and mortality. In Section 3, we introduce the concept of two types of risk factors: those for the underlying neuropathology of the disease and those for the expression of its clinical symptoms. The remainder of this third section is devoted to a critical summary of the large body of literature concerning contributions of genetic and environmental factors to the causation of this disease. In Section 4, we review biomarkers for this condition that can be observed during the preclinical phase of AD. Because the pathology is already advanced at the time a case can be diagnosed, effective prevention will likely require identification of the individuals at risk for this illness years or decades in advance of the initial symptoms. In Section 5, we address the science of preventing AD from an epidemiologic perspective, including how prevention might occur at the individual and population levels.

Amy R. Borenstein
James A. Mortimer

Acknowledgments

This book is the result of many years of collaboration and discussion about Alzheimer's disease and related neurologic diseases, epidemiology, statistics, genetics, and biology between the coauthors. It also is the culmination of many interactions with those who have been our teachers, colleagues, friends and students. Dr Borenstein first thanks her parents for supporting her through school and always encouraging her to reach her goals. She thanks her father for giving her the love of data from his profession as a Ph.D. economist with the United Nations. She also thanks Dr Ann Stromberg of Pitzer College, Claremont, CA, for introducing her to the field of epidemiology in 1976. Professor Emily White at the University of Washington School of Public Health and Community Medicine and Fred Hutchinson Cancer Research Center played a critical role in Dr Borenstein's academic development, serving as her mentor through her dissertation research. Professor White introduced Dr Borenstein to Alzheimer's disease in 1983 and proposed that Dr Borenstein pursue her dissertation in this area. She is particularly grateful to Dr Leon Thal (1945–2007), who awarded her a Neurosciences Education and Research Foundation Merit Award for most promising new investigator in Alzheimer's disease research in 1992. The greatest thanks are given to her coauthor, professional partner, and husband, Dr James A. Mortimer, for his constant professional advice and collaboration, his patience and love.

Dr Mortimer thanks his many mentors as well as collaborators that he has worked with over the past 40 years. His initial interests in neuroscience were sparked by the founders of this field, including Dr Rodolfo Llinas, as well as Dr Gardner Quarton, who served on his doctoral committee at the University of Michigan and provided direction in his early career. Later at the National Institutes of Health, he developed techniques for recording from single neurons in awake animals with Dr Edward Evarts, a prelude to his later interests in diseases of the nervous system. He thanks his collaborators in the Department of Neurology at the University of Minnesota, where he spent 20 years as a faculty member, in particular, Dr David Webster, who facilitated the development of a research center devoted to Alzheimer's disease and Parkinson's disease. Other important collaborators included Drs David Snowdon and Bill Markesbery at the University of Kentucky, his coprincipal investigators on the Nun Study, and Dr Margaret Gatz of the University of Southern California with whom he worked on the Swedish Twin Study of Dementia. His most important influence

over the past 20 years has been his wife, Dr Amy Borenstein, with whom he has shared this intellectual journey and without whom the ideas in this book would not have evolved.

Both of us would like to thank our friends and colleagues, most notably Drs Lon White, Peter Schofield, Lorene Nelson, and Lenore Launer for their many hours of insightful discussions and heated debates about the nature and causes of Alzheimer's disease and other neurodegenerative disorders.

Our gratitude goes to the many Alzheimer patients and their families, as well as people without the disease, who participated in the many research studies that are represented in this book. The altruism that goes into participating in observational studies and clinical epidemiologic trials represents the pinnacle of the subject-researcher connection. Despite their own uphill battles, they selflessly donated their time and effort to help others to move science forward and help solve the mysteries of our time.

Finally, we are incredibly fortunate to have our children, Rebecca Anna Graves and Kent Mortimer, and his wife and daughter, Diane and Eileen, without whose love and support this book would not have been possible.

Prologue

A Primer on Epidemiologic Concepts and Methods

Epidemiology is the basic science of public health that examines the distribution of diseases in human populations by person, place, and time (descriptive epidemiology), and the causes or etiology of these diseases (analytic epidemiology). Epidemiologic methods are used in both observational studies and clinical research investigating the effects of drugs or other interventions in the treatment and prevention of disease. In this primer, we explain the basic concepts of epidemiology for those who are not trained in public health.

While the ultimate goal of epidemiologic research is to identify potential causes of diseases and to use this information for prevention, its starting point is to describe the occurrence of disease by characteristics related to person, place, and time by answering the following questions: Who gets the disease? Where does the disease occur more or less frequently? What are the disease trends over time? These descriptive questions can be answered by looking at *rates* of disease among defined populations, across geographic locations, and over time. More detailed questions regarding risk or protective factors for the disease, such as having a particular exposure, can be addressed by analytic studies that compare individuals' histories who have the disease to those who do not have the disease. Alternatively, one can begin with individuals who are disease-free and follow those who are exposed to a risk or protective factor and those who are not exposed to the same risk factor over time to see whether those with the exposure develop the disease more often than those not exposed. In experimental epidemiology, a group of people are randomized to receive a drug or intervention of some kind (exposure) or not to receive it and are followed over time to test whether those who received the intervention develop the disease less frequently than those who did not receive the intervention.

DESCRIPTIVE EPIDEMIOLOGY

Incidence

The most important measure for understanding causation is the incidence rate, the number of new disease events divided by the number of people who are at risk

in a defined population for a specified period of time. For example, in a study of Alzheimer's disease (AD), let's assume that a population of 1,000 people known to be free of AD at the start of the study (baseline) is carefully followed over the ensuing 10 years. If everyone in the study survived for the entire 10 years, the denominator for the incidence rate would be $1{,}000 \times 10 = 10{,}000$ person-years. However, some participants will not complete the 10 years of follow-up. Some may refuse to continue their participation because of sickness or another reason, others may be lost to follow-up because they leave the area in which the study is being conducted, and still others may die. The denominator of the incidence rate takes such losses into account. The numerator is a count of the number of people who did not have the disease of interest at the beginning of the study, but developed the disease over the 10 years. If the incidence rate is say, 2% per year, then over 10 years of complete participation, 200 people will get the disease (20 per year if all characteristics stayed the same for all years of the follow-up). Because the denominator will be smaller than 10,000, the true incidence rate will be a little larger than 2%. For example, if the denominator is 8,986, the incidence rate will be 200/8,986 or 2.23%. We can look at the incidence rate in any subgroup of the population we wish. For example, we can divide up or "stratify" the population of 1,000 people into men and women, different age groups, different exposure groups (e.g., people who eat fish three or more times a week and those who eat fish two or fewer times a week); or any other variable that we measure at baseline in the "cohort" we decide to follow. This allows us to look at the incidence rate by characteristics related to person. For example, we might find that people who eat fish three or more times per week have a lower incidence rate of the disease than those who eat fish two or fewer times per week.

To compare disease incidence according to place or geographic location, we could compare incidence rates by age groups in men and women in different countries or states. It is important to note that such comparisons will only be valid if the disease is defined in the same way in different places. To determine whether the frequency of the disease has been changing over time, we would need to measure the disease the same way over many years, and compare the incidence rates at the two (or more) time points, taking into account differences by other characteristics that may have changed in that time period, like the age structure in the population. It is important to recognize that there are reasons that we might see changing incidence rates over time that may not be related to real changes in the incidence rate (the speed of developing new disease). For example, the disease may become better recognized in the medical community or the stigma associated with the disease may lessen over time increasing the likelihood that doctors make a diagnosis. There also may be changes made in the diagnostic criteria for the disease that will lead to apparent increases or decreases in the incidence rates of a disease. Epidemiologists first consider whether such methodological changes could be responsible for differences in incidence rates by person, place and time. Only after such explanations can be

excluded or adjusted for is it possible to make a valid statement about differences related to etiology of disease.

Prevalence

The prevalence rate differs from the incidence rate in that it counts the number of *existing* cases of disease in a defined population *at a specified time*, rather than the number of new or incident cases, and divides this by the size of the population that is living and at risk (including those who are living with the disease). Another way of expressing the prevalence rate is that it measures the proportion of people in a defined population that has a disease at a specified point in time (Rothman, Greenland, & Lash, 1998). The prevalence rate provides a good measure of the burden of the disease. Going back to our example, if we followed 1,000 people for 10 years, and 20 people developed the disease each year, then by year 3, we should have approximately 60 people who are living with the disease, provided no one has died from the disease during that period. The prevalence rate (P) at year 3 would be 60/1,000 = 6%. The prevalence rate is heavily influenced by the duration of survival after disease onset and is generally equal to the incidence rate (I) multiplied by the mean duration (D) of the disease ($P = I \times D$). For AD, mean survival is about 7 years. Thus, a new case is expected to live on average 7 years. For some diseases in which the duration of disease is short, for example, pancreatic cancer, the incidence rate will be approximately equivalent to the prevalence rate. For other diseases such as multiple sclerosis with very long durations, the prevalence rate or burden of the disease in the population will be much higher than the incidence rate.

The prevalence rate generally is not of interest in studies of etiology. If we examine risk factors for prevalence, we will mix risk factors for incidence (true disease rate) with those for duration. The incidence rate is appropriate for studying risk factors related to the speed or rate of developing a disease.

Mortality

The mortality rate uses as its numerator the number of deaths from the disease of interest in a specified time period divided by the number of people at risk of dying in the population. In our example, in the cohort of 1,000 people, let's assume that 3% died the first year. Because the closed cohort is growing older each year, we would expect the percent of deaths to increase a little in subsequent years and we could use actuarial data to estimate the expected death rates each year, which would depend on the cohort members' sex and age distributions. We can examine the mortality rate among people with AD and compare that to the mortality experience in the rest of the cohort to see whether AD increases the probability of death.

Unless we enumerate observed deaths within a defined study cohort in a rigorous manner, enumerating deaths from AD in large populations is extremely inaccurate. Most commonly, this would be done using death certificate data. For AD, these are highly unreliable due to variability in who is reporting and how much they know about the deceased person, as well as standards in reporting and how AD is assigned to a code on the certificate. Comparing mortality rates for AD between different countries is highly questionable, since the measurement of whether someone had AD or a related dementia varies from country to country, as does the practice and quality of maintaining death records.

ANALYTIC EPIDEMIOLOGY

Epidemiologic tools include two main types of analytic designs from which an estimate of risk can be derived. Using the scientific method, epidemiologists develop a hypothesis regarding potential causation and then analyze data to test this hypothesis. The presence of an association is determined by the ability to refute or not to refute the null hypothesis of no association between an exposure and an outcome. In epidemiology, we seek the same result over many studies (consistency) conducted with differing methodologies and in different populations, before we apply other causal criteria to infer a cause-and-effect relationship between a putative exposure and an outcome.

Randomized Clinical Trials

Scientists have long considered randomized clinical trials (RCTs) the gold standard of study designs, the design that "proves" whether an exposure causes an outcome. In such a trial, the exposure is manipulated by the investigator by randomly assigning a defined group of people to receive or not receive a drug, behavioral, dietary, or other intervention. The relative risk (RR) compares the incidence rate of disease among those randomized to the treatment or intervention to those randomized not to receive the treatment or intervention. The goal of randomization is for the two groups to resemble one another in all respects except for the intervention. Successful randomization depends on the size of the sample and proper implementation of the randomization scheme. RCTs can only be conducted using interventions that will putatively lower the incidence rate of the disease over the follow-up period. Risk factors for disease, such as smoking or exposure to chemicals, cannot ethically be randomly allocated to individuals. Although RCTs offer the best level of evidence for an exposure-outcome relationship, there are several methodological issues that can bias the results and many challenges that must be carefully considered.

Observational Cohort Studies

Observational cohort studies are generally considered to provide the next best level of scientific evidence to support causal inferences. It is important to note

that null findings in a randomized trial do not prove that a risk factor does not cause a disease. Null findings may be the result of low statistical power from a small sample size or an inadequately long duration of the trial. As we shall see in this book, AD evolves over several decades. Therefore, relatively brief RCTs may not be the optimum design to examine risk factors that exert their effects over years or decades. Randomized controlled trials and observational cohort studies for AD *answer different questions*, rather than provide different levels of evidence.

When potentially harmful exposures are examined and when potentially protective exposures that have their effects over many years are of interest, the observational cohort study is the study design of choice. Epidemiologists typically assemble a cohort of individuals that can be defined geographically, ethnically, or by specific characteristics, such as religion or occupation, examine each person for the presence or absence of the disease of interest at the start of the study, and then follow the population carefully to document incident cases of disease.

In studies of dementia or AD, potential cohort participants are usually first screened with a cognitive test and are asked to participate in a full neurologic and neuropsychological examination if they meet specified screening criteria or if they are otherwise deemed to be at high risk for dementia or AD. Because a cohort study aims to identify new or incident cases of dementia or AD, only those judged not to be demented at initial evaluation (baseline) are included in the cohort to be followed. At the baseline examination, these nondemented participants are questioned using a highly structured questionnaire about a large number of risk factors, such as their family and medical histories, their educational background, diet, exercise, and other personal habits including smoking and alcohol consumption. Typically every 2–3 years, the cohort is seen again, and the cognitive screening test is readministered, with individuals meeting study criteria for going on to the clinical evaluation being tested in depth for the presence of incident dementia. This is called a two-phase, or multiphase case-finding procedure. Each potential case is discussed by the study consensus diagnostic committee in which physicians and neuropsychologists fill out diagnostic criteria for dementia and for dementia subtype, for example, AD, vascular dementia, or dementia due to another cause. The study may continue for many years or may end, usually depending on funding. The investigators can then examine risk factors from the baseline or from any examination during the study follow-up period as they relate to incident cases of the prespecified outcomes. The RR is typically calculated, by dividing the incidence rate of dementia or AD among individuals exposed to a factor by the incidence rate of dementia or AD among individuals who were not exposed. If this RR exceeds 1.0 and the probability of this happening by chance is smaller than 5% ($p < 0.05$) (which is equivalent to saying that the 95% confidence interval or 95% CI does not include 1.0), the risk factor is said to be "statistically significant" with respect to its association with increased risk for the outcome. If the RR is less than 1.0 and $p < 0.05$, the risk factor is said to be inversely associated with or protective for

the outcome. If the *p*-value is >0.05 (the 95% CI includes 1.0), the factor is said not to be significantly associated with the risk for disease. There are multiple forces that can falsely elevate or decrease the RR estimate, which we will deal with briefly in the sections on Validity, Bias and Confounding.

One type of RR that we will use in the book is the hazard ratio (HR). Simply put, the HR is the true ratio of two incidence rates, in which the estimated measure of effect (the RR) takes time into account. Thus the HR not only tells us the magnitude of the RR (how strong the RR is comparing exposed to unexposed individuals) but also tells us that people with the factor will get the disease earlier (risk factor) or later (protective factor) in time than those without the factor.

The attributable risk, or the risk difference, is the rate of disease that would be eliminated if we could eliminate a causal exposure. This measure subtracts the incidence rate among unexposed individuals from the incidence rate among exposed individuals. The population attributable fraction estimates the proportion of the disease that is due to the exposure if the exposure is causally related to the incidence of disease. This proportion depends on the prevalence of the exposure in the population of interest.

Case-Control Studies

While cohort studies (and experimental clinical trials) can measure the speed of acquiring the disease of interest, case-control studies cannot. The only exception is when a case-control study is nested within a cohort study. Case-control studies are considered the next best analytic design with regard to study design quality. In a case-control study, all or a representative sample of cases of the disease are found, preferably in a known population, and controls (people without the disease) are selected from the same or a similar population as the cases. It is very important to use the same criteria to select cases and controls, the single exception being presence of the outcome. For example, if the disease criteria require that cases be free of vascular disease, controls must also be selected from among those free of vascular disease. Otherwise, vascular disease may be found to be a protective risk factor for the disease. Controls may be matched to cases on characteristics like age, sex, or education to increase the power to examine other factors, or they may be selected at random from the same population as the cases. If matched, the factors used to match cannot be examined for their association with the disease. Once the cases and controls have been identified, they may be interviewed about their exposures before the disease began in the cases, and DNA and other biospecimens may be obtained. To identify risk factors, the two groups are compared. Since incidence rates are not available, the odds ratio (OR) is used. The OR is the odds that cases (people with the disease) have been exposed to the risk factor divided by the odds that controls (persons without the disease) have been exposed to this factor, that is, the OR compares the proportion of cases who have been exposed to the factor with the

proportion of controls who have been exposed. The OR is interpreted the same way as the RR or HR. If the OR exceeds 1.0 and $p < 0.05$ (the 95% CI does not include 1.0), the factor significantly increases the risk for disease; if the OR is less than 1.0 and $p < 0.05$, the factor decreases the risk for disease.

Case-control studies are the preferred study design for diseases that are sufficiently rare (less than 5% of the population at risk) that a cohort study would not easily generate a sufficient number of cases for meaningful analysis (the concept of statistical power). Because their memory for past events and exposures is impaired, we cannot interview AD cases after they present with dementia. For this reason, case-control studies of AD have traditionally utilized proxy informants, including spouses, children, or other knowledgeable informants, to gather past exposure histories by interview. This intrinsically lowers the quality of data that can be used in the estimation of risk. Because proxy informants respond less reliably than cases would if they were able to, interviewing only case proxy informants and not control proxy informants creates a bias that is mitigated somewhat if control proxy informants are also interviewed (Nelson et al., 1990). A validation substudy can be implemented within the case-control study such that controls are interviewed about themselves and control proxy informants are interviewed about the controls and their responses compared to check for data quality.

Cross-Sectional Studies

This study design uses a defined population or group of people to measure exposures and outcomes at the same time. While exposures may also be measured from the subjects' past mimicking a case-control or case-cohort design, the traditional cross-sectional study measures both at the same time. Therefore, no incidence rates are available and prevalence is usually the only measure available of the occurrence of the disease. Prevalence studies are cross-sectional studies in that they provide a snapshot of the population at a certain time. In this sense, this study design is considered descriptive rather than analytic. However, because we can calculate ORs and prevalence ratios (the proportion of exposed with the disease divided by the proportion of unexposed with the disease, used when the disease is not rare or when we are not sure of the temporal sequence of the exposure and the disease), cross-sectional studies can give us valid data regarding *possible* causation, particularly for exposures that do not change with time, such as genes.

Because an exposure must precede the onset of the disease in time to be considered a cause of the disease and cross-sectional studies do not allow for the temporal sequence of the exposure-disease association to be determined, they provide a lower level of scientific evidence than case-control studies. It is important to note that with regard to AD cohort or case-control studies also may have difficulty in establishing temporal sequence between exposure and outcome. Even though we try to establish that the exposures occurred before

clinical disease onset, if the disease is characterized by pathological changes that precede clinical symptoms, the exposure may represent a marker of the underlying disease pathology and may not be causally associated with the disease. This issue is considered in Chapters 21 and 22 of this book.

Ecologic Studies

Studies that are based on average levels of an exposure in populations rather than specific levels in individuals and in which the outcome of interest is prevalence or incidence of a disease in those populations are called ecologic studies. For example, mean dietary fat intake in different countries can be studied with regard to AD prevalence or incidence. While the mean level of exposure in a population may be associated with the incidence or prevalence of a disease, this does not necessarily imply that the exposure in individuals within the population is associated with the disease. Studies comparing mean exposures in populations can be subject to the ecologic fallacy, whereby patterns observed at the aggregate level are not seen at the individual level. Therefore, ecologic data cannot be used to support a statement of association. Because of this caveat, we will not consider this study design in this book.

Case Reports and Case Series

These types of reports provide the weakest level of scientific evidence, and fall into the descriptive study design category. A case report is a description of clinical and/or histologic observations for a single patient, and a case series describes such observations for a small group of patients. Because these reports do not include a comparison or control group, such data cannot be used to support the existence of an association between an exposure and an outcome. Often, when a disease is first seen by a physician, it is described in the literature as a case report. A small group of cases may then be assembled to see if there are commonalities among the cases. This occurred with the first case series report of Kaposi's Sarcoma in eight homosexual men in New York City, which heralded recognition of the AIDS syndrome (Hymes et al., 1981). Case reports are often used to generate a hypothesis to be tested later using an analytic study design, such as a case-control or a cohort study. Because case series and reports do not provide reliable data with regard to causation, they will not be considered in this book.

CASE DEFINITION

What is critically important about case reports and case series as well as analytic designs is the definition of a case. In epidemiology, explicit case definition criteria are of paramount importance to the uniform measurement of the outcome. If the outcome is misclassified (there are people in the case group without the disease or the control group is contaminated with some cases), the observed measure of effect from analytic studies may be falsely underestimated.

ASSOCIATION VERSUS CAUSATION

Epidemiologic methods can sometimes directly address causation if the exposure of interest is suited for an experimental study (randomized controlled trial or RCT). This is because at baseline, neither group has been exposed to the intervention and no one has had the outcome. When one group is experimentally exposed and shows a difference in outcome over time, assuming all other methodological issues are addressed, we can infer that the exposure occurred before the outcome (temporal sequence) and is likely a cause, at least in some of the participants in the trial.

In cohort and case-control studies, once an association has been found, there are three main questions that need to be addressed. Is the association *precise*, that is, could the study findings be due to chance alone? Second, is the association *valid*, that is, do the study findings reflect the true relation between the exposure and the outcome? Third, is the association *causal*, that is, is there sufficient evidence to infer a causal association between the exposure and the outcome? We will address each of these briefly.

PRECISION

Since each study uses a new sample of people from the population, there are natural errors associated with the selection process. In epidemiology it is important to minimize such errors. A "Type I" error occurs when one concludes that there is an association between the exposure and the disease, whereas in fact there is no association. These errors can occur because of poor study design or analysis or be due to chance. A "Type II" error occurs when the study fails to find an association, but in fact there is one in the natural world. This type of error can be due to bias or to other methodological shortcomings, such as the need to use proxy informants or designing the study with too few subjects. Precision refers to the size of the random error that can occur in the estimate of the OR or RR. Low precision is reflected by wide CIs around the estimate of risk. Precision of the estimate can be increased by increasing sample size. Samples that are too small may lead to wide CIs that include 1.0, leading to a Type II error (failure to reject the null hypothesis of no association when an association is actually present). It is important to add that while we use an α-value (size of permissible Type I error) of $p \leq 0.05$ as our default setting to declare a "statistically significant finding" (corresponding to a 95% CI that is statistically significant if it excludes the null value of 1.0 and not statistically significant if it includes the null value of 1.0), we should not dismiss p-values that are a bit higher than 0.05. For example, if the p-value for a result is 0.06 or 0.07, this means that the inclusion of a few more people in the sample might have made the difference between being statistically significant or not statistically significant. Thus, we might give such a p-value more consideration than if the p-value is clearly larger.

VALIDITY

Other errors in epidemiology fall under the umbrella term "validity," which refers to the extent to which the findings from a study reflect the true underlying situation in nature. These errors can be broadly grouped under two subheadings: internal validity and external validity. Internal validity refers to the degree to which the two groups being studied—the cases and the controls or the exposed cohort and the unexposed cohort—are like one another in all ways except the disease (case-control study) or the exposure (cohort study) permitting a *fair* comparison to be made. In a case-control study, the cases and controls should be selected from the same underlying population such that if the control became a case, s/he would be diagnosed in the same way with the disease as the cases in the study. External validity refers to the generalizability of the studied groups to the outside world. A study must be internally valid first, and then can be judged for its ability to extrapolate to the general population (external validity). An example of an internally valid study is the Nun Study (Snowdon et al., 1996), a study of 678 American Roman Catholic nuns who were members of the School Sisters of Notre Dame and were ages 75–102 at the beginning of the study. One inclusion criterion into the study was that the nuns were required to agree to donate their brain to the study when they died. This made the study unique, as it was the first to be able to examine neuropathologic lesions in the brain in an unbiased population of individuals who either became demented during the follow-up or remained cognitively intact. Before the Nun Study, autopsy case series were highly biased because only certain individuals who die are selected for autopsy, with dementia patients being over-represented. While the Nun Study has high internal validity, some scientists have questioned whether the study's findings can be generalized to other populations (external validity), for example, to men, or to individuals of different race/ethnicities living in the community.

In order for a study to be internally valid, the two groups being compared will ideally be similar to one another with regard to all characteristics of person, place and time except for the disease (in a case-control study) or the exposure (in a cohort study). This is why the randomized controlled trial has traditionally been considered the gold standard of epidemiologic study design. Successful randomization will result in intervention groups that match each other on both measured (and unmeasured, such as all genes) variables. In an observational study, we may have to adjust for differences between the groups on measured variables that differ between the groups and are also related to the outcome of interest (confounding). In order for the groups to be similar to one another, epidemiologists pay a great deal of attention to the selection of participants. In a case-control study, for example, a group of cases that come from a clinic or hospital may have some biases that are not immediately measurable. If the clinic or hospital is known for its work in an area, for example, it may have an Alzheimer's Disease Research Center, then it may attract highly educated

caregivers who want their loved one who is developing dementia to be clinically evaluated by expert doctors. When investigators conduct a case-control study using these cases, they may select controls that come from the same geographic area as the cases but are less educated. The cases may be found to have more education than controls, which in fact is the reverse of the truth. This finding would have been generated by referral bias and led to the wrong conclusion about the association between education and AD. Such biases are known in epidemiology as *selection bias*. Epidemiologists are trained to choose comparison groups that minimize selection bias such that fair comparisons can be made between them.

Another important source of bias in epidemiologic studies is called *observation bias*, which has to do with the way in which variables are defined and measured in the study. When a variable is measured with a large degree of random error, the RR or OR is reduced toward the null value of 1.0. This increases the chance of making a Type II error and erroneously concluding that an association is not present when in reality it is. Conversely, the effect measure (the RR or OR) can be erroneously overestimated if cases and controls are compared with regard to an exposure, and proxies for cases either over-recall the exposure, proxies for controls under-recall the exposure, or both occur simultaneously. This will result in a RR or OR that is overestimated relative to the underlying truth and may cause the investigator to conclude that an association exists when in fact it does not exist in the underlying population (a Type I error). Biases can work in different ways, and although cohort studies are thought to be superior to case-control studies, there are biases that apply to cohort studies that do not apply to case-control studies. For example, in a cohort study, individuals who at baseline do not have AD but present with clinical symptoms in the first few years of follow-up will likely already have some memory loss. When these individuals are interviewed for the study, they may selectively under-recall their past exposure histories. Individuals who are not destined to get AD will not under-recall such exposures. The net effect of this error is that the RR for positive risk factors will be attenuated toward the null value of 1.0, making a true effect more difficult to detect. In this book, we will have many opportunities to discuss different forms of bias and how the study results may be distorted by them.

One of the most important forms of bias is *confounding*. This occurs when a third variable associated with both the exposure and the outcome of interest explains some or all of the association that is seen between the exposure and the disease. We might be interested in the effect of a higher or lower education on the risk for AD. If we were just to examine the association between higher education and AD, we might find that there is a strong inverse association (people with higher education levels have lower risk for AD) but when we look at age, we find that the younger people in the sample have a higher education level than the older people in the sample. Since age is related to education level and also is related to the risk for AD, a subsequent adjustment for age would give

more "credit" to older people for low education and less "credit" to younger people, therefore leveling the playing field so that a fair comparison can be made between the two education groups. We might still find an inverse association between education and risk for AD, but it will likely be weaker than we observed before we adjusted for age (the "crude" RR). In this example, age is a confounder of the education–AD association and it is necessary to adjust for it to estimate an internally valid RR. Most of the time, we have many confounders that must be dealt with simultaneously and statistical modeling is used to for such adjustment.

Related to the topic of confounding is another concept we will deal within the book where a variable may seem on the surface to be a confounder, but is instead an "intermediate" or a "mediating" variable. This means that it may meet the criteria for confounding (it is associated with the exposure and is an independent risk factor for the outcome), but it is in the causal pathway between the exposure and the outcome. In this case, the exposure is causally related to the intermediate variable, which in turn causes the outcome. An example of this might be that physical exercise causes an increase in whole brain volume as measured by magnetic resonance imaging, which in turn determines if a person has clinical dementia or not. Brain volume in this example is not a confounder that needs to be adjusted, but is instead a stepping stone in the causal pathway between physical exercise and the probability of developing dementia.

A third variable in addition to the exposure and the outcome can be a confounder, an intermediate variable or an effect-modifier. *Effect-modification* refers to the magnitude of an association between an exposure and an outcome being different for different levels of a third variable. For example, if a person has a genetic predisposition to AD, for example, carries an Apolipoprotein E-ε4 allele, and has a low education level, it might be hypothesized that their risk for AD would be much higher than if they had the genetic susceptibility gene alone or they had a low education level alone. Do the two risk factors biologically interact to produce a much higher risk? Effect-modification is defined as a process in which an exposure acts with another factor to increase or decrease the risk for disease. There are two kinds of effect-modification: multiplicative (sometimes called synergistic) and additive. The first type is where the combined effects of the exposure and the third variable exceed the *product* of each effect alone. The second type is where the combined effects of the exposure and the third variable exceed the *addition* of each effect alone. We will have an opportunity to discuss both types of effect-modification in the ensuing chapters.

As we enter the age of personalized medicine, scientists are concerned with defining subgroups of individuals who have multiple genetic risks for AD (G–G interactions), or who have genetic risks and environmental risks (G–E) that interact with one another. As AD is a highly complex disease, there are likely to be multiple genetic and environmental risk factors that act together to modify its risk.