## CONTENTS

### CONTRIBUTORS

vii

### PREFACE

xi

### PART I  STUDY DESIGNS AND METHODS

1. Introduction to Epidemiologic Research Methods  
   Jerome A Fleming and Chung-Cheng Hsieh  
   3

2. Analysis of Categorized Data: Use of the Odds ratio as a Measure of Association  
   Stephen L. Hillis and Robert F. Woolson  
   35

3. Methods in Psychiatric Genetics  
   Stephen V. Faraone, Debby Tsuang, and Ming T. Tsuang  
   65

4. Reliability  
   Patrick E. Shrout  
   131

5. Validity: Definitions and Applications to Psychiatric Research  
   Jill M. Goldstein and John C. Simpson  
   149

6. Mental Health Services Research  
   Jack D. Burke, Jr  
   165

7. The Pharmacoepidemiology of Psychiatric Medications  
   Philip S. Wang, Alexander M. Walker, and Jerry Avorn  
   181

8. Peering into the Future of Psychiatric Epidemiology  
   Ezra Susser, Michaeleine Bresnahan, and Bruce Link  
   195

### PART II  ASSESSMENT

9. Studying the Natural History of Psychopathology  
   William W. Eaton  
   215

10. The Developmental Epidemiology of Psychiatric Disorders  
    Mary Cannon, Matti Huttunen, and Robin Murray  
    239
CONTENTS

11. Birth and Development of Psychiatric Interviews 257
   Lee N. Robins

12. Symptom Scales and Diagnostic Schedules in Adult Psychiatry 273
   Jane M. Murphy

13. DSM-IV and Psychiatric Epidemiology 333
   Michael B. First

14. The National Comorbidity Survey 343
   Ronald C. Kessler and Ellen Walters

PART III  EPIDEMIOLOGY OF MAJOR PSYCHIATRIC DISORDERS 363

15. Epidemiology of Psychosis with Special Reference to Schizophrenia 365
   Evelyn J. Bromet, Mary Amanda Dew and William W. Eaton

16. Epidemiology of Depressive and Anxiety Disorders 389
   Ewald Horwath, Rose S. Cohen, and Myrna M. Weissman

17. Epidemiology of Bipolar Disorder 427
   Mauricio Tohen and Jules Angst

18. The Epidemiology of First-Onset Mania 445
   T. Lloyd and P. B. Jones

19. The Epidemiology of Alcohol Use, Abuse, and Dependence 459
   Nancy L. Day and Gregory G. Homish

20. Epidemiology of Drug Dependence 479
    James C. Anthony, with John E. Helzer

21. Personality Disorders: Epidemiological Findings, Methods and Concepts 563
    Michael J. Lyons and Beth A. Jerskey

PART IV  EPIDEMIOLOGY OF SPECIAL POPULATIONS 601

22. Epidemiology and Geriatric Psychiatry 603
    Celia F. Hybels and Dan G. Blazer

23. The Epidemiology of Child and Adolescent Mental Disorders and Adolescent Mental Disorders 629
    Stephen L. Buka, Michael Monuteaux, and Felton Earls

24. Epidemiology of Mood and Anxiety Disorders in Children and Adolescents 657
    Kathleen Ries Merikangas and Shelli Avenevoli

Index 705
CONTRIBUTORS

Jules Angst, Research Department, Psychiatric University Hospital, P. O. Box 68, 8029 Zurich, Switzerland

James C. Anthony, Johns Hopkins University, 624 North Broadway, Room 893, Baltimore, MD 21205

Shelli Avenevoli, Mood and Anxiety Disorder Program, National Institute of Mental Health, 15K North Drive, MSC #2670, Bethesda, MD 20892

Jerry Avorn, Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women’s Hospital Medical School, Boston, MA 02115

Dan G. Blazer, Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC 27710

Michaeline Bresnaham, Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, and New York Psychiatric Institute, New York, NY 10032

Evelyn J. Bromet, State University of New York at Stony Brook, Stony Brook, NY 11794

Stephen L. Buka, Departments of Maternal and Child Health and Epidemiology, Harvard School of Public Health, Boston, MA 02115

Jack D. Burke, Jr., Department of Psychiatry, Harvard Medical School, The Cambridge Hospital, Cambridge, MA 02139

Mary Cannon, Division of Psychological Medicine, Institute of Psychiatry, London, UK SE5 8AF

Rose S. Cohen, College of Physicians and Surgeons of Columbia University, New York, NY 10032

Nancy L. Day, Western Psychiatric Institute and Clinic, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213-2593

Mary Amanda Dew, Johns Hopkins University, Baltimore, MD 21205

Felton Earls, Harvard School of Public Health, Boston, MA 02115

William W. Eaton, Department of Mental Hygiene, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205

Stephen V. Faraone, Department of Psychiatry, Massachusetts Mental Health Center, 750 Washington Street, Suite 255, South Eaton, MA 02375
Michael B. First, New York State Psychiatric Institute, New York, NY 10032

Jerome A. Fleming, Harvard Medical School, Department of Psychiatry at Massachusetts Mental Health Center, Harvard Institute of Psychiatric Epidemiology and Genetics, and Brockton/West Roxbury Veterans Administration Medical Center, Psychiatry Service, Brockton, MA 02301

Jill M. Goldstein, Harvard Medical School at Massachusetts Mental Health Center, Harvard Institute of Psychiatric Epidemiology and Genetics, Massachusetts General Hospital, Massachusetts Mental Health Center, Boston, MA 02115

John E. Helzer, Health Behavior Research Center, 54 West Twin Oaks Terrace, Suite 14, South Burlington, VT 05403

Stephen L. Hillis, Department of Statistics and Actuarial Science, University of Iowa College of Liberal Arts, Iowa City, IA 52242

tute and Clinic, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213-2593

Ewald Horwath, College of Physicians and Surgeons of Columbia University, New York State Psychiatric Institute, New York, NY 10032

Chang-Cheng Hsieh, Division of Biostatistics and Epidemiology, University of Massachusetts Cancer Center, Worcester, MA 01605

Matti Huttunen, Department of Mental Health and Alcohol Research, National Public Health Institute, Helsinki, Finland

Celia F. Hybels, Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC 27710

Beth A. Jerskey, Department of Psychology, Boston University, Boston, MA 02115

Peter B. Jones, Department of Psychiatry, University of Cambridge, Adenbrooke’s Hospital, Cambridge, UK

Ronald C. Kessler, Department of Health Care Policy, Harvard Medical School, Boston, MA 02115

Bruce Link, Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, and New York State Psychiatric Institute, New York, NY 10032

Tuhina Lloyd, University of Nottingham, Duncan Macmillan House, Nottingham, UK

Michael J. Lyons, Center for Clinical Biopsychology, Department of Psychology, Boston University, and Harvard Institute of Psychiatric Epidemiology and Genetics, Boston, MA 02215

Kathleen Ries Merikangas, Mood and Anxiety Disorder Program, National Institute of Mental Health, 15K North Drive, MSC # 2670, Bethesda, MD 20892

Michael Monuteaux, Harvard School of Public Health, Boston, MA 02115
Jane M. Murphy, Department of Psychiatry, Harvard Medical School, Department of Epidemiology, Harvard School of Public Health, and Psychiatric Epidemiology Unit, Department of Psychiatry, Massachusetts General Hospital, Charlestown, MA 02109

Robin Murray, Division of Psychological Medicine, Institute of Psychiatry, London, UK SE5 8AF

Lee N. Robins, Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63110

Patrick E. Shrout, Department of Psychology, New York University, New York NY 10003

John C, Simpson, Harvard Medical School Department of Psychiatry, Harvard Institute of Psychiatric Epidemiology and Genetics, Massachusetts Mental Health Center, Boston, MA, and VA Boston Healthcare System, Mental Health Careline, Boston, MA 02115

Ezra Susser, Department of Epidemiology, Joseph L. Mailman School of Public Health, Columbia University, College of Physicians and Surgeons, and New York State Psychiatric Institute, New York, NY 10032

Mauricio Tohen, Lilly Research Laboratories, Indianapolis, IN, Department of Psychiatry, McLean Hospital, Harvard Medical School, Boston, MA 02184

Debby Tsuang, VAPSHCS (116) MIRECC, 1660 South Columbian Way, Seattle, WA 98108

Ming T. Tsuang, Department of Epidemiology, Harvard School of Public Health and Pediatric Psychopharmacology Unit, Psychiatry Service, Massachusetts General Hospital, Boston, MA 02115

Alexander M. Walker, Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115

Ellen Walters, Harvard Medical School, Boston, MA 02115

Philip S. Wang, Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women’s Hospital, Boston, MA 02115

Myrna M. Weissman, Department of Epidemiology in Psychiatry, College of Physicians and Surgeons of Columbia University, New York, NY 10032

Robert F. Woolson, Department of Biostatistics, The University of Iowa, College of Public Health, Iowa City, IA 52242
It has been seven years since the publication of our first edition of the *Textbook of Psychiatric Epidemiology*. The field has continued to expand and important new findings have been published.

The intent of the first edition was to produce a textbook for our students at the Harvard Program in Psychiatric Epidemiology and Biostatistics as well as for students from other training programs across the United States. We have received extremely positive feedback about the first edition from students and faculty from training sites across the United States. Our expectations were actually surpassed, as general psychiatrists, epidemiologists, and other mental health professionals have been very favorable of the textbook. The interest in our textbook, especially from Western Europe, has expanded our geographical boundaries.

The second edition includes an update of the chapters by the same distinguished faculty. We have extended our list of contributors to include our European experts who are contributing as co-authors or, in some cases, with chapters that were not included in the first edition. We have also added two chapters on the epidemiology of child mental disorders.

The textbook is prepared in four separate sections. The first focuses on study design and methods, the second on assessment, and the third on epidemiology of major psychiatric disorders. The fourth section focuses on the epidemiology of special populations, such as the elderly and children.

As in our first edition, our objective is to provide a comprehensive, easy to understand overview of research methods for the nonmethodologist. Our targeted audience is students of psychiatric epidemiology, psychiatric residents, general psychiatrists, and other mental health professionals.

We would like to acknowledge three individuals; Alexander Leighton, Gerald Klerman, and Brian MacMahon who were the foundation of the Harvard Program in Psychiatric Epidemiology.

*Ming T. Tsuang*
*Maurice Tohen*
CHAPTER 1

Introduction to Epidemiologic Research Methods

JEROME A. FLEMING and CHUNG-CHENG HSIEH

Harvard Medical School, Department of Psychiatry at Massachusetts Mental Health Center, Harvard Institute of Psychiatric Epidemiology and Genetics, and Brockton/West Roxbury Veterans Administration Medical Center, Psychiatry Service, Brockton, MA 02301 (J.A.F.). Correspondence to JAF: (116A) 940 Belmont Street (508) 583-4500 / fax 586-6791; Division of Biostatistics and Epidemiology, University of Massachusetts Cancer Research Center, Worcester, MA 01605 (C.C.H.).

INTRODUCTION

Epidemiology is the study of the distribution and determinants of disease frequency in humans (MacMahon and Pugh, 1970). Classic epidemiologic research designs developed to study chronic diseases are being used increasingly in investigations of psychiatric disorders. In turn, efforts to study psychiatric conditions have introduced new methodologic challenges for epidemiologists. Despite many advances in psychiatric classification in the last three decades, case definition, the sine qua non of many facets of epidemiologic research, remains an area of controversy in psychiatry. The complex manifestations and courses of psychiatric syndromes are often difficult to capture in basic epidemiologic study designs involving data collection at one or two points in time. In addition, risk factors for psychiatric conditions can be as difficult to conceptualize and assess as psychiatric outcomes.

Notwithstanding these methodologic challenges, epidemiology offers some of the best available research strategies for addressing critical questions in psychiatry concerning the nature, etiology, and prognosis of mental disorders. Psychiatric cases seen in treatment represent a small, highly self-selected segment of the full spectrum of psychopathology found in the general population. Epidemiologic study designs enable inferences to be made about the total population at risk, even when study subjects are drawn from treatment settings. Also, many putative determinants of mental disorders, such as gender, marital status, social class, and stress, cannot be randomly assigned to study groups for ethical or practical reasons.

Experimental methods used in medicine and psychology that rely on randomization therefore cannot be used to study these types of risk factors. In comparison, observational epidemiologic designs are fully appropriate.

In this chapter, we review some of the common approaches to quantifying the occurrence of psychiatric outcomes in a population and will present basic epidemiologic research designs used to identify the determinants of psychiatric conditions. Biases associated with observational epidemiologic study designs, and factors to consider in interpreting findings from these studies, are discussed. Attention is also given to the special problems faced in the application of these methods to the study of psychiatric conditions.

**EPIDEMIOLOGIC MEASURES OF OUTCOME OCCURRENCE IN POPULATION GROUPS**

The frequency of outcome occurrence in a population group can be measured several ways. The two principal approaches involve measures of proportions and measures of densities (rates). The distinction and relation between these two types of measurements have been discussed in detail in the context of psychiatric research (Kramer, 1957). They are described briefly here.

**Incidence Density (Force of Morbidity or Mortality)**

Incidence refers to new events (outcomes) occurring over time among members of the population who are candidates for such events. There are two commonly employed incidence measures: incidence density and cumulative incidence.

An incidence density quantifies the number of events occurring per unit of population per unit of time. It is not dimensionless because time is retained in the unit of measurement. In estimating incidence density, the population under study should exclude all individuals with the health outcome at the start of the period of observation. This candidate population is often referred to as the population at risk. In practice, when the number of cases in the population under study is very small, such as in studies of rare diseases in general population samples, the total population can be used for the population at risk. In small study cohorts, however, it is important to remove all current cases from the baseline sample before calculating incidence.

Incidence density can be assessed for an instantaneous time point by the slope of a curve measuring change in disease-free population size over time. This instantaneous rate of change is often referred to as the hazard rate or the force of morbidity. Incidence density is also often expressed as an average rate of change over a time interval. For example, if a group of 300 manic-depressive patients is followed for an average of 10 years with 12 deaths (the outcome event) occurring during the follow-up, the numerator of the average incidence density of death (the mortality rate) would be 12 deaths, and the denominator would be 300 patients times 10 years, or 3,000 person-years. After division, the mortality rate would be reported as 4 per 1,000 persons per year (or 4 per 1,000 person-years).

A density-type measure is usually referred to as a rate. However, in common usage, rates can also refer to proportions, such as unemployment rate, tax rate,
and prevalence rate. To avoid confusion, it is important to know the context in which rate is being used and to specify the method by which it has been calculated (Elandt-Johnson, 1975).

**Cumulative Incidence, Risk, and Survival**

Cumulative incidence, risk, and survival rates are estimates of the probability of the occurrence of an outcome event over a specified period of time. Cumulative incidence is usually used to describe the probability of outcome occurrence among a group or population. Risk is usually used to predict an individual's chance of such an event. Risk is also commonly expressed by its mathematical complement, the probability of surviving or the survival rate. Cumulative incidence, risk, and the survival rate are dimensionless measures.

Cumulative incidence can be either an observed probability or a theoretical quantity estimated from the incidence density function. The observed cumulative incidence is a simple proportion and is calculated as the number of health outcomes occurring over a time interval divided by the size of the population at risk. If the outcomes of all members of a candidate population are observed without any loss to follow-up from causes other than the event under study, cumulative incidence can be used as an estimate of individual risk for the time interval under study (e.g., five-year risk of dying) or, in a complementary fashion, as the survival rate.

In practice, however, loss to follow-up or censoring through subject dropouts or death by other causes is common. The interpretability of the observed cumulative incidence measure when such loss occurs is seriously compromised. For example, an observed five-year survival rate for a group of manic-depressive study subjects can be distorted by censoring, even if those who were lost had the same probability of surviving as the remaining study participants. This distortion will take place because outcomes occurring among subjects lost to follow-up are excluded from the numerator of the observed cumulative incidence calculation. Cases lost to follow-up are still retained in the denominator, however, which equals the total size of the candidate population at the start of the study with no adjustment for reduction in the size of the study cohort over the observation period. Consequently, observed cumulative incidence, and risk and survival estimates based on it, is only appropriate for studies in which there is negligible loss to follow-up over the course of the study. The types of studies for which these observed measures are best suited involve closed or fixed cohorts (that is, cohorts in which the disease course of each subject in the study is individually monitored over the period of observation) in which there is no loss to follow-up and the follow-up interval is short.

When loss-to-follow-up occurs, or when incidence is estimated for a dynamic community population (i.e., where the disease experience of each individual is not monitored), a more appropriate measure of the probability of disease occurrence is derived from the observed incidence density function (Chiang, 1968). The estimate of the observed incidence density is not affected by the competing causes of subject removal (e.g., loss to follow-up from a candidate population since those who are lost will no longer be among the candidates for the occurrence of the next outcome event. For the prognosis of an individual patient in this study, the
complement of a five-year survival rate derived from the incidence density can be appropriately interpreted as the risk of dying in five years.

**Prevalence**

Simply put, a prevalence or prevalence rate is that proportion of persons in a population who have a particular health condition at a point or period in time. For example, the point prevalence of major depression in a community is the number of persons fulfilling diagnostic criteria for depression at a stated point in time divided by the number of persons in the community. As a proportion, prevalence is a dimensionless quantity; that is, it is not expressed in units of another characteristic, such as time.

Both newly onset cases and cases that begin prior to the study period contribute to prevalence. In a community population in which the numbers of entries and exits (from births, deaths, migrations, and so forth) are balanced and the disease rates are stable (a steady state), prevalence is proportional to the frequency of development of new cases of the condition (the incidence density) multiplied by the average duration of the condition. Exact relationships between prevalence, incidence, and duration have been presented by Freeman and Hutchison (1980, 1986).

Prevalence rates are frequently reported for population subgroups, such as age- or sex-specific rates. In these stratum-specific estimates, the numerator of the prevalence is formed by the number of cases within the population subgroup, and the denominator is the total size of the subgroup.

In psychiatric studies, “period” prevalence rates are also often reported. A period prevalence rate uses the same denominator as a point prevalence rate, but expands the numerator to include all cases present during a selected time period, such as one month, six months, one year, or a lifetime. Period prevalence has gained popularity in psychiatric epidemiology because of the complex, episodic course of many psychiatric conditions. Use of a period prevalence allows individuals with chronic psychiatric conditions who are temporarily in remission to be included in prevalence counts. Also, the diagnostic criteria for many psychiatric disorders requires the occurrence of clusters of symptoms over extended time intervals, such as one month (e.g., generalized anxiety) or one year (e.g., dysthmic disorder). A time period is therefore implicit in any prevalence measure involving these conditions, even a point prevalence.

Although period prevalence has several practical advantages, there are a number of limitations associated with this hybrid measure. In extended time periods, patients who remit early in the time interval without recurrence are likely to be missed in the period prevalence counts, especially when the information by disease status is gathered by recall (Aneshensel et al., 1987). In addition, empirical estimates of lifetime prevalence frequently exhibit a counter-intuitive age distribution. Over the age distribution of a population, lifetime prevalence should increase during age intervals associated with disease onset and remain constant at other ages. However, lifetime prevalences of many psychiatric disorders have been observed in several population surveys (Weissman and Myers, 1978; Robins et al., 1984) to decrease sharply in the older age groups. Several explanations have been offered for this artifact. In addition to recall bias, high case fatality rates (i.e.,
patients do not survive until older ages), increasing rates of psychopathology in recent cohorts, and changing diagnostic practices have been suggested as explanatory factors (Robins, 1985; Klerman, 1988).

**Measures of Association and Impact: Relative Risk, Odds Ratio, and Attributable Risk**

Epidemiologic studies yield statistical associations between a disease and exposure. We must interpret the meaning of these relationships, since an association may be artifactual, noncausal, or causal. An artifactual or spurious association may arise because of bias in the study. When an outcome is affected by multiple variables, in order to examine the influence of a single one, it is necessary to adjust for the effects of the others. A simple technique for isolating a specific effect due to one variable is to examine the outcome rates, at several levels of this variable, while holding the other variables constant. A more sophisticated approach involves the use of multiple regression analysis to measure the independent effect of the contribution of each of a series of variables on an outcome. Casuality is assumed when one factor is shown to contribute to the development of disease and its removal is shown to reduce the frequency of disease (Morton et al, 2001).

If there is an association between a study factor and a psychiatric disorder, the frequency with which the disorder occurs will differ in groups that vary on the study factor, such as groups who are exposed and not exposed to an environmental agent or a trait. Therefore, a measure of the association can be obtained by comparing the rates of disease occurrence in exposed and unexposed groups. Group comparisons can be expressed as a difference or as a ratio of rates. The magnitude of the difference or ratio is an indicator of the strength of association between the study factor and psychiatric outcome. In psychiatric epidemiology, ratios of disease rates are typically used to express the strength of the association. The ratio between two rates (or “rate ratio”) is often referred to as the relative risk. Since relative risk can also be a risk ratio and rates and risks are different measures of disease occurrence (see “Cumulative Incidence, Risk, and Survival,” mentioned earlier), it is important to know the context in which relative risk is used.

To illustrate, suppose an investigator is interested in comparing the mortality rates of adults with and without a psychotic disorder in a community of 120,000. In this population, 1,200 persons (1%) meet diagnostic criteria for a psychotic disorder, and 118,800 do not. Over a 1-year period, 312 deaths occur, including 15 individuals with a diagnosis of psychosis and 297 without. The rate (density) of dying for the group with psychosis (15 of 1,200 psychotics) and without this disorder (297 of 118,800) expressed as a mortality rate ratio (relative risk of dying) would be 5 (15 / 1,200 divided by 297 / 118,800).

Certain types of epidemiologic studies do not directly assess the disease experience in a population and compare estimated disease rates for individuals with and without exposure. Instead, the exposure histories of samples of individuals with (cases) and without (controls) the disease from the population are compared. This type of subject selection is commonly referred to as retrospective or case–control sampling. It is possible, nevertheless, to estimate the rate ratio of disease occurrence among cases and controls in these studies if certain conditions are met.
Suppose in the example above that available resources do not allow the investigator to determine the mental health status for each of the 120,000 residents in the community. With a retrospective (case–control) sampling approach, cases would be the complete or partial sample of the 312 subjects with the outcome event (death), and controls would be a sample of the 120,000 residents who were candidates of the outcome event. If the available resources allowed for sampling approximately two times the number of controls as cases, an investigator might decide to select 600 subjects as the controls. With a random sample of the population, the distribution of the exposure (the psychotic disorder) among these 600 controls would be proportional to the distribution in the original population. Therefore, 6 controls (1%) would be expected to have a psychotic disorder and 594 would not after an examination of their mental health status. Table 1 displays the cross-tabulation of the outcome and exposure status from this sampling design.

To estimate the relative risk of dying among those with and without a psychotic disorder in this case–control study, the odds of exposure among the cases (15/297) is contrasted to the odds of exposure among the controls (6/594). The result of the division of these two odds, known as an “odds ratio,” is 5. Note that the odds ratio computed from the case–control study in this example yields the same result as the mortality rate ratio among psychotics and nonpsychotics in the total population. An odds ratio is frequently computed as an estimate of relative risk or incidence rate ratios in case–control studies. The accuracy of this approximation depends on a number of factors, including the nature of the source population (i.e., whether it is a dynamic population with a “steady state” of in- and out-migration), the rarity of the outcome, the use of incident versus prevalent cases, and the length of the risk period between exposure and disease occurrence. The reader is referred to Chapter 3 (this volume) and to Kleinbaum et al. (1982) for a detailed description of the conditions under which an odds ratio equals or approximates a rate ratio or relative risk in the retrospective sampling schemes used in case–control studies. For the most common types of case-control studies involving incident cases, the odds ratio estimates the rate ratio exactly.

Another commonly employed epidemiologic measure is attributable risk (AR), which is also known as the etiologic fraction or population attributable risk percent (Kleinbaum et al., 1982). The AR describes the proportion (or percent) of new cases arising in a population that are attributable to the exposure under study. The AR depends on the prevalence of the exposure in the population and on the strength of the association between the exposure and outcome. The AR can be estimated by the following formula: AR = \( p_e(RR - 1) / (p_e(RR - 1) + 1) \) where \( p_e \) is the proportion of the source population that is exposed and \( RR \) is the relative risk estimate. The AR ranges in value from 0 (none of the outcome occurrence is

**TABLE 1. Results of a Case–Control Study of Relative Risk of Mortality Among Psychotics and Nonpsychotic Adults**

<table>
<thead>
<tr>
<th>Exposure Status</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>With psychotic disorder</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>Without psychotic disorder</td>
<td>297</td>
<td>594</td>
</tr>
<tr>
<td>Total</td>
<td>312</td>
<td>600</td>
</tr>
</tbody>
</table>
attributable to the exposure) to 1 (all occurrences take place in the presence of the exposure, i.e., the exposure is a “necessary” cause). The accuracy of this measure depends on the extent to which component measures used to calculate AR reflect current population characteristics. This index is useful for planning and policy purposes because it describes the potential impact of removing an exposure upon the frequency of disease occurrence.

Attributable risk can also be calculated specifically for individuals who have a positive history of exposure. This estimate, known as the attributable risk among the exposed (AR_e) or attributable risk percent, is computed as $\text{AR}_e = (\text{RR} - 1) / \text{RR}$. The AR_e can be interpreted as the probability that an exposed case developed the condition as a result of the exposure. As a hypothetical example, in a study where exposure is family history and the outcome is schizophrenia, an AR_e of 0.75 would indicate that 75% of the schizophrenic cases with a positive family history for this disorder developed their condition because of their familial loading.

OVERVIEW OF STUDY DESIGNS

Epidemiologic research in its most elementary form involves studying the relationship between a risk factor and a health outcome. The risk factor is often referred to as the exposure or treatment. To learn about its relationship to a health outcome, a comparative study is undertaken in which the experience of disease occurrence in a group of individuals with one characteristic (e.g., exposed) of the risk factor is contrasted with that of another group differing on the characteristic (not exposed).

Although the epidemiologic approach focuses on comparisons of the relative magnitudes of rates of disease occurrence between two groups, in practice a variety of research designs can be invoked. These designs can be distinguished by a number of characteristics. The most important distinctions involve the timing of data collection in relation to risk factors and disease occurrence, the separation between risk factor and disease occurrence in time, and the methods used in sampling study subjects (Miettinen, 1985a). In addition, studies vary in cost, feasibility, and quality of information gathered. The study designs listed below are in common use in psychiatric epidemiology, although there is some variation in the terminology employed at different research centers. We briefly describe each type of the study in turn and provide examples of studies examining psychiatric outcomes:

Experimental
Nonexperimental (observational)
  - Cross-sectional
  - Cohort
    - Prospective
    - Retrospective
  - Case–control
    - Case–crossover
    - Case–cohort
INTRODUCTION TO EPIDEMIOLOGIC RESEARCH METHODS

Hybrid studies
- Repeated cross-sectional
- Multistage
- Panel

In addition to these basic epidemiologic study designs, the reader is referred to Chapters 4 through 8 (this volume) for reviews and illustrations of other research designs currently employed in psychiatric epidemiology.

Experimental Studies

In an experimental study, the investigator controls the allocation of subjects to different comparison groups and also regulates the experimental conditions of each group. Study subjects are randomly assigned to comparison groups and followed up over time to record the outcome event of interest, such as the recurrence of a psychiatric illness or the occurrence of death. Clinical trials and intervention studies are the most common forms of experimental studies in human populations. To ensure the comparability between groups and obtain valid results, an experimental study employs three basic research strategies: randomization, placebo, and blinding.

**Randomization.** When an investigator randomly assigns subjects to different experimental conditions, differences between groups are determined by chance. If the randomization is carried out properly and the sample sizes are sufficiently large, the groups are likely to be similar in all regards other than the conditions under study. Consequently, if the experimental conditions have no effect the rates of disease occurrence are expected to be the same in the comparison groups.

Even with random allocation it is possible that the groups will be imbalanced with respect to extraneous factors that may influence the rates of disease occurrence, particularly if the sample sizes are small. Before analyzing the results of randomized experiments, it is generally recommended that investigators test whether the groups are balanced on all known or suspected determinants of the outcome under study. If an imbalance is detected, the investigator can use statistical methods to adjust for the effects of these factors on the distribution of disease occurrence across groups. For unknown determinants, it is usually assumed that randomization will achieve a balanced distribution on these factors in the long run over hypothetical repetitions of the same study. The confidence in this assumption increases if the number of study subjects is adequate.

**Placebo.** One complication of experimental studies is that extraneous aspects of the treatment procedure may influence the outcome under study. For example, psychiatric patients who are given a new medication may show improvements because they receive special attention from study staff monitoring the treatment trial. Participating in an experiment in and of itself can also influence outcomes, an artifact that is commonly known as the Hawthorne effect. To control for these unwanted effects, one comparison group is usually administered a placebo that, under optimal conditions, mimics the extraneous features of the experimental condition or treatment under study but does not otherwise influence the rates of
disease occurrence. Differences in disease rates between the placebo and experimental groups can be attributed to the effect of treatment per se rather than to the effect of other aspects of the procedure, activity, or environment associated with the treatment. Differences between placebo groups and groups that are not assigned to any experimental condition are also measured in some randomized trials, and these differences are referred to as placebo effects.

**Blinding.** For many of the reasons previously discussed, with placebo treatments it is important that participants in a randomized trial be unaware of the group to which they are assigned. It is equally important to withhold this information from the investigator and other professionals who manage the trial. Knowledge that an individual has been assigned to the experimental treatment may influence the handling, treatment, and measurements of participants in the randomized trial. Standardization of the study procedures are also easier to enforce when both the investigator and the patient are unaware of the group assignments. The process of “double blinding,” in which neither investigators nor study participants are given information about the group assignment, helps to ensure that group conditions are similar and that identical study procedures are followed with every study subject. Although double blinding is desirable in every randomized experiment, it is not always feasible, especially when the treatment produces other effects that are observable or require monitoring to protect participants, such as changes in blood pressure.

Even though experimental studies are considered a paradigm in many research fields, the randomized trial has several shortcomings for use in studying human populations. Ethical considerations dictate that experimental studies involving human subjects can only be used to study exposures (treatments or medications) that are likely to be beneficial. It is unethical to randomize human subjects to harmful exposures. Furthermore, constitutional characteristics such as inherited or congenital traits cannot be randomized. It is also not feasible to randomize groups into many other sociodemographic conditions that may influence mental health outcomes, such as marital status or religious denomination. Therefore, the effect of many putative risk factors for major psychiatric disorders cannot be evaluated by an experimental study. Also, if the follow-up period for the ascertainment of outcomes of an experimental procedure is long, the treatment assessed may be obsolete by the time the results are available (Elwood, 1988).

**A Randomized Clinical Trial in Psychiatry.** Random assignment of study participants to intervention and control groups is the procedure that will give the greatest confidence that the groups are comparable. If you have two groups of patients, and you apply a different treatment to each group (clinical trial), you can only attribute a difference in outcome to the differing treatment if that is the only factor that differs between the groups (Morton et al., 2001). This goal can only be achieved if group membership is determined randomly. There is usually a logical sequence to clinical trial analysis. It begins with a comparison of the intervention and control groups to demonstrate comparability, showing that randomization works. Finally, the main analysis is to test whether the hypothesized health effect resulted.

Gibbons et al. (1993) present results from a longitudinal analysis of a randomized clinical trial of two forms of psychotherapy using the NIMH Treatment of
Depression Collaborative Research Program Dataset. The objectives of this clinical trial were to evaluate and compare the effectiveness of cognitive behavior therapy (CBT) and interpersonal psychotherapy (IPT) in comparison to a standard reference treatment, imipramine plus clinical management group (IMI-CM). A placebo plus clinical management group was also enrolled to control for effects of standard treatment (PLA-CM). Subjects \( n = 250 \) were randomized into each of these four experimental conditions; 239 subjects entered treatment and 219 received measures after baseline. Depressive symptoms were assessed over 16 weeks with a modified Hamilton (1960) rating scale completed by clinical evaluators who were blind to treatment conditions. Contrasts between the experimental groups were made to test three main null hypotheses: (1) no difference between the two psychotherapies (IPT compared with CBT); (2) no difference between psychotherapy and standard treatment (IMI-CM); and (3) no difference between the standard treatment (IMI-CM) and the placebo (PLA-CM).

No significant differences were found between the two psychotherapies Hypothesis 1 or between psychotherapy and standard treatment Hypothesis 2, but rate of improvement for the standard treatment (IMI-CM) was significantly greater than for the placebo. This detailed report also describes methods taken to control for potential bias introduced by attrition after randomization, missing data, differences between collaborating research sites, and assumptions in statistical modeling.

**Nonexperimental (Observational) Studies**

In a nonexperimental study, the investigator has no control over the group designation of each study subject. The investigator generally selects subjects for the different exposure conditions from previously existing groups and then observes the resulting health outcomes. Hence, epidemiologic nonexperimental studies are sometimes called observational studies. The three most common epidemiologic observational studies designs are cross-sectional, cohort, and case–control studies.

Our discussion of observational designs begins with these classic methodologies.

**Cross-Sectional Studies.** In a cross-sectional study, the data on exposure and outcome are obtained at the same point in time, and both usually relate to the current period. The information is typically gathered through sample surveys of geographically defined populations. The current disease status of groups with and without the exposure, expressed as prevalence rates, are compared in analysis. By providing a “snapshot” of the current levels of illness in the total population and in different exposure groups, this design has been found to be useful for describing the health care needs of different populations.

Cross-sectional studies have enjoyed considerable popularity in psychiatric epidemiology for a number of reasons. A population survey allows investigators to gather information on all cases of disorder occurring in a defined area, including syndromes in an asymptomatic phase and conditions for which treatment is not routinely sought. Because current diagnostic procedures in psychiatry rely heavily upon the verbal report of symptoms, the interview methods used in most surveys can be used to obtain some of the basic information commonly used in formulating diagnoses. Also, prevalence rates obtained by cross-sectional surveys are widely
used in psychiatry because onset (incidence) is difficult to demarcate. The chronicity of many psychiatric disorders also facilitates prevalence estimation, which, as will be recalled, is proportional to the product of incidence times duration. Therefore, even though the incidence rates for most psychiatric disorders are believed to be very low, the number of prevalent cases detected in a cross-sectional survey of moderate size is often sufficient to obtain precise estimates of rates and measures of association.

For an illustration of a major cross-sectional study in psychiatric epidemiology, the reader is referred to Chapter 5 (this volume) on the Epidemiologic Catchment Area study.

**Cross-Sectional Survey Sampling.** A study sample that is representative of the target population is an essential feature of cross-sectional surveys. To achieve representativeness, subjects are selected as probability samples of the population using sample survey methods (Kish, 1965; Cochran, 1977). A variety of different sampling methods are in current use that vary in complexity. Before designing a cross-sectional survey, it is important to consult a statistician about the appropriate method to employ. The sampling method will influence the number of subjects required for the survey, and certain sampling designs will also require special data analytic procedures such as weighted data and variance adjustments. Although a comprehensive overview of sampling methods is beyond the scope of this text, we will mention some of the major approaches and highlight some of the major factors that influence selection of one method over another.

Before describing the sampling methods, some terminology must be defined. The target population is the group to which results are to be generalized. This may be inclusive of all individuals in a geographic area or may exclude certain groups, such as individuals above or below a certain age or institutionalized individuals. Elementary units are the elements or members of the target population to be studied. Individuals are usually the elementary units in epidemiologic studies, but examples of other elementary units include households, neighborhoods, or hospitals. A list of all of the units in the target population used to draw the sample is known as the sampling frame, and the entries (e.g., names or addresses) on the sampling frame are called enumeration or listing units. Examples of sampling frames include telephone directories, voter registration or tax lists, town censuses, and utility listings.

Before selecting a sampling scheme, the investigator should examine the available sampling frames. Ideally, there should be a one-to-one correspondence between the enumeration units on the sampling frame and the elementary units in the target population. In practice, this is rarely the case. Some frames only contain clusters or groupings of elementary units. For example, an investigator may wish to survey all individuals in a town, but only has access to a frame (e.g., utility listings) that enumerates households. Examples of other problems with sampling frames include missing elements (failing to provide coverage of all individuals in the target population), duplicate entries, and blanks or foreign elements (e.g., out-of-date lists that include individuals who have died or emigrated, or overly inclusive lists, containing individuals outside the target population or individuals whose primary residence is outside the geographic area under study). Before the sample is drawn,
the investigator should review and correct errors in the list. The list may need to be updated by contacting current residents in the survey area, a process referred to as enumeration.

There are several types of sampling plans used in cross-sectional surveys. Choice of a sampling plan depends on a number of issues, including the information contained in the sampling frame, the rarity of the characteristic under study, the desired precision of the prevalence estimates or prevalence ratios, the size of the area to be studied, and the cost of the study.

One of the most commonly cited sampling methods but infrequently employed in actual practice is simple random sampling. This method requires the availability of a complete listing of the population to use as a sampling frame. The usual method of drawing a simple random sample is to number each element on the sampling frame from 1 to \(N\), where \(N\) is the size of the target population, assuming that the frame is completely accurate. A set of \(n\) unique random numbers, where \(n\) is the desired number of elements to be contacted for the survey, is then obtained either from a random number generator on a computer or from a published table of random numbers. The frame is then searched for elements whose numbers correspond to each of the \(n\) random numbers. These elements are chosen to be the study sample. If random numbers are not available, a lottery method can also be used by preparing \(N\) cards or tokens representing enumeration elements on the frame and drawing the desired \(n\) number of tokens at random.

In simple random sampling, the probability that any individual element is chosen is the ratio of the sample size to the size of the population: \(n/N\). Although this sampling method is intuitively easy to understand, a complete listing of the population is not always available. In addition, it is possible that rare characteristics will not be represented in a simple random sample. This method is also very expensive for large study areas because interviewers will be required to travel throughout the survey region.

A modification of simple random sampling is known as stratified random sampling. In this method, the sampling frame is divided into different strata (such as age, sex, and ethnic-race groups), and simple random samples are drawn within each stratum. This approach ensures adequate representation of different groups under study. Under most conditions, it will also improve the precision of prevalence estimates. To carry out stratified random sampling, as with simple random sampling, a listing of the population is required. In addition, the characteristics to be used in stratification must also be available on the frame.

When a list of the population is not available, two commonly employed sampling methods are systematic sampling and cluster sampling. Systematic sampling is one of the most widely used methods in practice and has the advantage of being easily taught to individuals who have little knowledge of survey methods. It can also be used for samples that accrue over time, such as patient enrollments. In this method, sample members are drawn at fixed intervals, as, for example, every fifth household or every seventh name on a class enrollment list. The sampling interval, \(k\), can be calculated by dividing the projected total population size (\(N\)) by the desired sample size (\(n\)). For example, if it is estimated that there will be 100 houses in a community and a sample of 25 is desired, the sampling interval is 100/25 or 4, and interviewers can be instructed to go to every fourth household.
Despite its simplicity, an investigator should consult with a statistician before using this method, because it may yield biased, imprecise prevalence estimates. If the population \((N)\) and sample size \((n)\) are reasonably large and the elements randomly ordered, the estimates can be assumed to be unbiased with variances approximating simple random sampling.

Cluster sampling is the most complex survey sampling procedure of the four methods described here. As previously described, a cluster is a listing element that may contain more than one elementary unit. Examples of clusters of individuals include hospitals, classrooms, and households. Geographic areas, such as states, counties, cities, or blocks, also represent clusters in many sampling schemes. In cluster sampling, a probability sample of clusters is drawn. In a single-stage cluster sample design, information is then gathered on all elements in each sampled cluster. Alternatively, multistage sampling may take place, in which probability samples of elements are drawn at each stage until a sample of the desired elementary units is obtained. To illustrate the multistage cluster sampling process, consider the following example of a five-stage design for a probability sample of adults in the United States: In stage 1, a random sample of counties is drawn; stage 2 consists of a random sample of towns within each selected county; in stage 3, a random sample of blocks is drawn from each selected town; stage 4 consists of a random sample of households in each sampled town; and the process concludes with a random selection of one adult from each household (stage 5).

There are several advantages to this approach. First, the investigator does not need a list of all of the elementary units (e.g., all adults in the United States) in order to sample. Second, data collection is concentrated in small areas, decreasing the fieldwork costs. These potential benefits have to be weighed against two principal disadvantages. First, there is frequently a loss in precision of the population estimates obtained by cluster sampling, which is reflected by larger standard errors, broader confidence intervals, or a decreased statistical power to detect differences between groups in the sample compared with simple random sampling. This loss in precision is commonly measured by a design effect, which is the ratio of variances obtained under cluster sampling versus simple random sampling. Another related disadvantage of cluster sampling is that special statistical software for complex survey samples may be needed in order to obtain correct variance estimates.

**Measures of Disease and Exposure Status in a Cross-Sectional Survey.** Study participants in a cross-sectional survey are not enrolled on the basis of their exposure or disease status. All information regarding these factors is obtained during the investigation and is usually limited to survey interview information. There are three common methods of conducting surveys: mail surveys, telephone surveys, and face-to-face interviews. [These methods vary in expense and quality. Dr. Robins reviews the relative merits of these approaches in Chapter 11 (this volume).]

We limit our discussion of measurement in cross-sectional studies to one problem concerning the time frame for information obtained in survey interviews. Cross-sectional surveys are conventionally viewed as assessing both disease and exposure data at the current point in time. Cause and effect cannot be distinguished for true cross-sectional data of this type. However, many cross-sectional
surveys also attempt to obtain some information about events predating the current point in time. This historical information is usually based on the respondent's recollection and may be subject to considerable error (Neugebauer, 1981). Severe or salient events that are not embarrassing to report, such as death of a parent, birth of a child, or marriage, may be recalled with reasonable accuracy (Funch and Marshall, 1984; Kessler and Wethington, 1991). However, past emotions or behaviors are difficult to recall accurately, and historical reports of psychiatric symptoms may be biased by the current mental health of the respondent (Aneshensel et al., 1987; Schrader et al., 1990). A researcher should exercise considerable caution in attempting to assess life history information through respondent recall. Time lines, visual cues such as medication charts, or organization of questions around concrete events or by social contexts (e.g., home, work, school) may be used as memory aides (Kessler and Wethington, 1991).

**Cohort Studies.** Cohort studies in epidemiology have two essential features. First, study subjects are defined by characteristics present before disease occurrence, and these individuals form the study cohort. This is in contrast to case–control studies, where subjects are selected according to their disease status, and to cross-sectional studies, where subjects are selected by neither disease nor exposure status, but, instead, are selected to be representative of a target population.

The second characteristic of a cohort study is that real time is allowed to elapse before disease status is ascertained. Cohort members are followed through time to determine the frequency of new outcomes or events in each group. Measures of exposures and outcomes thus are both gathered at the time of their occurrence. This type of study design thereby offers the greatest potential of the epidemiologic observational studies to separate cause and effect. However, if the time elapsing between exposure and disease onset is long, and if the exposure levels vary over time, this type of study can be extremely costly and difficult to undertake.

There are two types of cohort studies that differ primarily in regard to the timing of study in relation to the occurrence of exposure and disease outcomes. The experience of a cohort can be studied prospectively or retrospectively, as is described in the following section.

**Prospective Cohort Studies.** In prospective cohort studies, groups of initially disease-free people are classified in terms of their exposure and are then followed forward in time. It should be noted that “disease free” is a relative term. For disorders with a poorly defined onset, such as psychiatric disorders, it may be difficult to guarantee that all members of the cohort are truly disease free at the outset of the investigation. [This issue is explored in greater detail in Chapter 9 (this volume) on studying the natural history of psychopathology.] Also, in practice, some retrospective information on exposure history may also be collected at baseline in addition to assessing current exposure levels.

Prospective cohort studies can be further subdivided into two study types based on whether the cohort is selected with or without regard to exposure status. Selection without regard to exposure status is frequently undertaken by following a study cohort sampled in a cross-sectional population survey over time. Three major cross-sectional study samples that have formed longitudinal cohorts in psychiatric