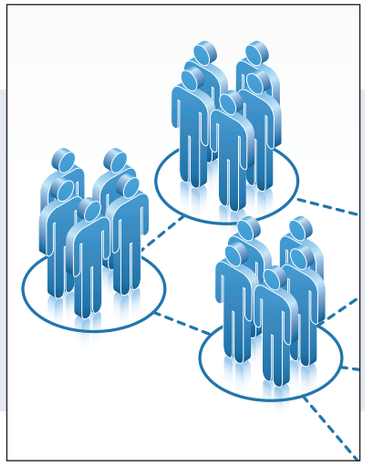

THIS ISSUE:

Design and Analysis of Longitudinal Studies, Part 2

Robert D. Gibbons, PhD
Guest Editor

This is the second issue on conceptual and statistical issues related to the design and analysis of psychiatric studies. In the first issue [(2009;38(12)], we provided a series of five papers related to 1) conceptual and experimental design issues related to missing data in longitudinal studies, 2) advances in the analysis of longitudinal data that insulate inferences from the effects of missing data, 3) the role of the intent-to-treat principle in longitudinal studies and various alternatives, 4) balancing treatment and comparison groups in longitudinal studies, and 5) sample-size determination for clustered and/or longitudinal studies.

In this issue, we provide an overview of many more general statistical topics that psychiatric researchers should be intimately familiar with. These topics include 1) the problems of determining association in two-by-two tables as related to assessing risk factors and diagnostic and prognostic tests, 2) the analysis of observational data in general and how nonexperimental data can be used to derive inference regarding treatment effects, 3) the role of analysis of covariance (ANCOVA) in analyzing experimental data, 4) statistical approaches to analysis of multiple outcomes in psychiatric investigations, and 5) a review of the concepts of "mediators and moderators" and how they are related to clinical decision making.

As in the first issue, the second issue benefits from the experience of many of the leading statisticians and biostatisticians who have ever been involved in psychiatric research. Here I provide a brief overview of these very significant (both statistically and clinically) contributions.

In the first article, Helena Kraemer leads us in a discussion and overview of the theory, application, and clinical relevance of determining association in two-by-two tables. The work covers the entire spectrum of statistical work in this area. Dr. Kraemer's focus is on the proper use of statistical methods to identify risk factors and the evaluation of the properties of diagnostic and prognostic tests. At the core of the problem is the availability of two binary variables that may be associated in a particular population. One variable may be a test, and the other may be a diagnosis. One variable may reflect the presence or absence of a factor such as early life stress, and the other may be a future event such as suicide. One variable may be a gene or gene profile, and the other may be success on a given pharmacologic treatment. Dr. Kraemer reviews all of the relevant statistical terms used in this area and shows how they are related. Not only are statistical issues covered in detail, but also design and sampling issues are thoroughly discussed.

In the second article, the critically important topic of analysis of observational or nonexperimental data is discussed. This article, led by Elizabeth Stuart, provides a detailed overview of how one uses nonexperimental data to estimate treatment effects when the absence of randomization can produce both measured and unmeasured sources of bias. The authors describe how to replicate a randomized experiment, at least with respect to the measured confounders, by making the treatment and comparison groups look as if they could have been randomly assigned to the groups, in the sense of having similar distributions of the confounders. The article describes a variety of statistical methodologies, including regression adjustment and several different approaches based on the propensity score. Once the samples are balanced, the next step is to estimate the average treatment effect. Finally, the authors describe various

approaches to assessing the degree to which the results are sensitive to potential unmeasured confounders.

In the third article, Armando Teixeira-Pinto and colleagues discuss the emerging area of multiple outcomes analysis. Unlike traditional multivariate methods such as multivariate analysis of variance (MANOVA), the authors consider more general approaches that relax distributional assumptions and allow for the pooling of different types of measurements (eg, a mixture of continuous and binary outcomes in a single analysis). Motivation for the paper comes from the WECare study, which investigates a sample of low-income minority women who were treated for major depressive disorder. An interesting bi-product of the analysis is a new approach to the treatment of missing data, which "borrows strength" from the outcome measurements that are available. When the outcomes are of different types (eg, a mixture of continuous and discrete), a latent variable is introduced to absorb their correlation. The authors show that in the presence of missing data, their approach does a much better job of recovering the underlying structure in the data than do a series of analyses that are conducted separately for each outcome.

In the fourth analysis, Helena Kraemer and I provide a detailed discussion of the ideas of "mediators and moderators" as they relate to psychiatric research. A moderator is a variable that helps to identify whom or under what conditions treatment has an effect on a particular outcome of interest. To show that a variable is a moderator, a) it must precede treatment, which in turn must precede the outcome; b) it must be independent of treatment; and c) if the population is stratified on the moderator, the effect size of the treatment on the outcome will be different in the different strata. For example, a gene may moderate the effect of treatment if the effect size of the treatment on the outcome is different for individuals with and without the gene. By contrast, a mediator is a variable that explains how or why the treatment has an effect on the outcome. For a variable to be a mediator, a) it must occur after treatment but before the outcome, b) the treatment and the mediator must be correlated, and c) the effect of treatment on the outcome can be explained either completely or partially by the mediator. A simple example of a mediator is compliance with treatment, which may be easier for one treatment versus the other and therefore explains at least part of the effect of the treatment on the outcome of interest. The authors review methods for the detection of moderators and mediators, costs and risks of ignoring moderators and mediators, testing hypotheses regarding moderators and mediators, and why randomized clinical trials so often mislead clinical decision making.

As a field, we have come a long way in the past 20 years in the development refinement of conceptual and statistical approaches to the analysis of psychiatric data.

In the final article, Tom Belin and colleagues provide an in-depth discussion of analysis of covariance (ANCOVA). Although perhaps most familiar to psychiatric researchers in general, the authors expand our view of ANCOVA by showing how it can be used to achieve a variety of scientific and statistical objectives. Among these, the authors highlight: 1) increasing precision in randomized experiments, 2) adjusting for sources of bias in observational studies, 3) enhancing our understanding of the nature of treatment effects in randomized studies, 4) the study of similarity and differences of regressions across levels of a classification or grouping variable, and 5) its use as a tool for the analysis of missing data.

This issue of *Psychiatric Annals* builds on the foundation provided in the previous issue by continuing the journey of taking the reader through an evolution of statistical thinking that is designed to improve our ability to extract meaning from quantitative information.

As a field, we have come a very long way over the past 20 years in the development and refinement of conceptual and statistical approaches to the analysis of psychiatric data. We hope that this issue leads to further advances in the statistical and methodological rigor of research in mental health. I am honored to have had the opportunity to continue my work with this distinguished group of statisticians.

doi: 10.3928/00485713-20090625-04

about the guest editor

Robert D. Gibbons, PhD, is Director of the Center for Health Statistics, Professor of Biostatistics and Psychiatry, University of Illinois at Chicago.

Dr. Gibbons received his doctorate in statistics and psychometrics from the University of Chicago in 1981. He has spent his entire career at the University of Illinois at Chicago (1981-present) where he directs the Center for Health Statistics, a consortium of 15 statisticians working in both theoretical and applied areas of environmetrics, chemometrics, biometrics, and psychometrics. Support for his research includes numerous grants and contracts from the National Institutes of Health (NIH), National Institute of Mental Health (NIMH), Office of Naval Research (ONR), National Cancer Institute (NCI), and the MacArthur Foundation. Recognition for his work includes a Young Scientist Award from the ONR, the Research Scientist Award from NIH, the Harvard Award for lifetime contributions to psychiatric epidemiology and biostatistics, the Lucaks Award for contributions to environmental statistics in the 20th century, and two Youden prizes (2001 and 2006) from the American Statistical Association, for statistical contributions to the field of chemistry. Dr. Gibbons is a Fellow of the American Statistical Association and a member of the Institute of Medicine of the National Academy of Sciences. He has authored more than 200 peer-reviewed scientific papers and five books.

Dr. Gibbons has dedicated this issue to the fond memory of Joe Fleiss

Professor Joseph L. Fleiss was the long-time head of the Division of Biostatistics (1975-1992) at Columbia University. Professor Fleiss was a leader in mental health statistics, particularly in the areas of assessment of reliability of diagnostic categories and the measures, models, and control of errors in classification. His book, *Statistical Methods for Rates and Proportions,* remains a landmark text on the topic of analysis of categorical data. Now in its third edition, the book has been in print for more than 35 years.

In an influential 1974 paper co-authored with Dr. Robert Spitzer, Professor Fleiss demonstrated that the second edition of the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (DSM-II) was an unreliable diagnostic tool. They found that different practitioners using the DSM-II were rarely in agreement when diagnosing patients with similar problems. The article by Fleiss and Spitzer led to major improvements in diagnostic criteria for psychiatric disorders, which were incorporated in future editions of DSM.

Among his many honors, Professor Fleiss was given the Mortimer Spiegelman Award by the Statistics Section of the American Public Health Association in 1973. In 1997, Harvard University recognized Professor Fleiss with a Lifetime Achievement Award for his contributions to psychiatric epidemiology and biostatistics.