

THIS ISSUE:

**Design and Analysis of Longitudinal Studies**

**Robert D. Gibbons, PhD**  
*Guest Editor*



In 1993, several of the authors of this issue of *Psychiatric Annals* wrote an article titled “Some Conceptual and Statistical Issues in the Analysis of Longitudinal Psychiatric Data.” The article provided a timely discussion and illustration of newly developed mixed-effects regression models for the analysis of longitudinal psychiatric data. The article has now been cited more than 500 times. Not only has the article been used to inform the psychiatric research community, but it has been widely cited throughout the medical literature.

Well, we’re back, and in this issue, we focus on topics that have arisen in the analysis of longitudinal data over the past 15 years that we as statisticians believe are important for the progress of medical research in general, and psychiatric research in particular. Specifically, we focus on sample size determination for clustered and/or longitudinal studies; the role of the intent-to-treat principle in longitudinal studies and various alternatives; conceptual and experimental design issues related to missing data in longitudinal studies; advances in the analysis of longitudinal data that insulate inferences from the effects of missing data; and balancing treatment and comparison

groups in longitudinal studies. Each of these areas has been the focus of considerable statistical research, and advances made in these areas

---

***Multiple repeated measures  
 can greatly reduce the  
 sample size.***

---

are now ready for practice. It is our hope that the following articles provide a balanced view of these important statistical methodologies and issues with an emphasis on their practice in applied settings.

The article on sample size determination led by Dulal Bhaumik (see page 765) provides an in-depth discussion of the numbers of centers and the numbers of subjects that are required for longitudinal studies in which subjects are also clustered within research centers, clinics, hospitals, schools, counties, or other settings. Both the clustering of subjects within ecological strata, such as physician or clinic, and the clustering of repeated measurements within subjects have a profound impact on the required number of participants in both randomized clinical

trials and observational studies. This article makes an important distinction in sample size requirements between studies that involve patient-level randomization versus studies that randomize physicians or other clusters of individuals to treatments. As shown in the article, the total number of subjects that are required for cluster-level randomization is far greater than for studies that involve the randomization of subjects to treatments within research centers. On the other hand, multiple repeated measures can greatly reduce the needed sample size. These two issues are discussed in the context of a now historic trial of antipsychotics versus placebo in the treatment of schizophrenia.

The paper on the principle of intent-to-treat (ITT) led by Tom Ten Have (see page 772), focuses on the effects of non-adherence in longitudinal mental health clinical trials. Although ITT has been mandated as the primary analysis strategy for randomized trials by the Food and Drug Administration, often the ITT principle is augmented with as-treated methods entailing comparisons between treatment adherers and non-adherers in contrast with the ITT randomized treatment comparison. In the context of men-

tal health trials, the authors clarify and compare the definitions of intent-to-treat, as-treated, and per-protocol procedures. Because of the vulnerability of as-treated and per-protocol approaches to unmeasured confounders, they review a different approach based on the use of an instrumental variable, as an alternative to the as-treated method that is immune to unmeasured confounding under a set of often testable assumptions. Empirical comparisons of these methods used in several examples from the psychiatric and statistical literature reveal informative insights on the effectiveness of mental health interventions among different groups of patients defined in terms of intervention adherence.

The paper led by Phil Lavori (see page 784) addresses the complex issue of missing data in longitudinal clinical trials. In contrast to much of the literature on statistical approaches for the analysis of data with various forms of missing information, this article focuses on design and conceptual issues that lead to a reduction in missing data overall and a reduction in the effect of missing data on the conclusions of longitudinal clinical trials. This is an often overlooked yet critically important area that is directly applicable to any researcher contemplating conducting a longitudinal clinical trial. This paper provides critical insights into the ITT principle and routine methods such as “last observation carried forward” (LOCF) that have been widely used as a “remedy” for the effects of missing data on statistical inferences from longitudinal clinical trials.

The companion paper on missing data, led by Juned Siddique

(see page 793), focuses on the advantages of longitudinal studies and the complexities that are produced by the presence of missing data in longitudinal records, including sta-

---

***Estimates of treatment efficacy may be biased when the treated and comparison groups are not balanced.***

---

tistical approaches that avoid bias in results by taking the missing values into account in the statistical analysis. These more modern approaches, which include mixed-effects regression models, multiple imputation methods, and pattern-mixture models, are contrasted with ad-hoc approaches such as LOCF, end-point analysis, complete case analysis, mean substitution, regression substitution, and single imputation. The modern approaches are then compared and contrasted using data from the WECare study, a longitudinal randomized treatment study of low income women with depression.

The paper on balancing treatment comparisons, led by Sue Marcus (see page 805), explores the use of propensity score matching in the context of longitudinal trials. Estimates of treatment efficacy may be biased when the treated and comparison groups are not balanced. Longitudinal studies are particularly vulnerable to this type of bias in that treatment and comparison groups can become less comparable over the course of the study due to subject dropout, treatment switching, noncompliance, and missing data.

In this article, the authors show how techniques used in observational cross-sectional studies can be used to balance comparisons in longitudinal studies. They discuss when to adjust during the course of the study (eg, baseline, posttreatment, follow-up assessments, etc.), and present methods for creating propensity scores for longitudinal studies such that the comparisons are balanced and inference is not compromised by adjusting for variables that have been affected by treatment. To help fix ideas, three longitudinal studies in which propensity scores are used to balance treatment comparisons are used as illustrations: an intervention study aimed at preventing HIV transmission among runaway youths housed at shelters in New York City; a Montreal-based study that evaluated the effect of gang joining at age 14 on subsequent violence; and the observational follow-up to the randomized Multimodal Treatment Study of Children with Attention Deficit Hyperactivity Disorder (ADHD).

This issue of *Psychiatric Annals* takes the reader through an evolution of statistical steps that begin with design issues that mental health researchers can use to minimize the effects of missing data on the validity of statistical inferences, methods for the analysis of incomplete longitudinal data, solutions to problems of lack of treatment adherence that are often characteristic of longitudinal clinical trials, approaches to balancing treatment comparisons over time, and finally tools for the selection of sample sizes, numbers of research centers, and number of measurement occasions in longitudinal

## guest editorial

studies. As a field, we have come a very long way over the past 15 years in the development and refinement of conceptual and statistical approaches to the analysis of psychiatric data. We hope that this presentation leads to further advances in the statistical and methodological rigor of research in mental health.

As a final note, Groucho Marx said that he would never be a member of a club that would have him as a member. Although I have tried to

emulate Groucho in every possible way, I could not be more delighted to be a member of this distinguished group of statisticians. I am sure Professor Marx would feel the same.

*This issue is dedicated to the fond memory of Sam Greenhouse. Sam Greenhouse was Professor Emeritus at George Washington University. He established the first Biometry Division at the National Institute of Health, and was Chief of the Theoretical Statistics and Mathematics*

*Section of the National Institute of Mental Health. He received numerous prestigious awards throughout his remarkably productive career, including the American Statistical Association's Founders Award, and the Harvard Award for lifetime contributions to Psychiatric Epidemiology and Biostatistics. While a visiting professor at Stanford, his excellent teaching convinced a young mathematics student, Brad Efron, to pursue a career in statistics.*

## 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 Institutes (NCI), and MacArthur foundation. Recognition for his work includes a Young Scientist Award from the ONR, 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 Associations 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 180 peer-reviewed scientific papers and four books. His latest book, *Longitudinal Data Analysis*, with Don Hedeker, has recently been published by John Wiley and Sons.