Professor Everitt argues convincingly that for psychiatry to advance systematically as a scientific discipline, psychiatric researchers will need to rely more and more on the "scientific approach" and on statistical techniques. His clear and stimulating presentation has performed a great service to the field of psychiatry by introducing a large audience of statisticians to the problems and challenges of statistics in psychiatry.

ADDITIONAL REFERENCES

CHEN, W. C., HILL, B. M., GREENHOUSE, J. B. and FAYOS, J. V. (1985). Bayesian analysis of survival curves for cancer patients following treatment. In Bayesian Statistics II (J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds.) 299–328. North-Holland, Amsterdam.

FAREWELL, V. T. (1982). The use of mixture models for the analysis of survival data with long-term survivors. Biometrics 38 1041–1046.

GREENHOUSE, J. B. and WOLFE, R. A. (1984). A competing risks derivation of a mixture model for the analysis of survival data. Comm. Statist. A—Theory Methods 13 3133–3154.

KUPFER, D. J. (1984). Neurophysiological 'markers'—EEG sleep measures. Psychiatric Res. 18 467–475.

KUPFER, D. J. and FOSTER, F. G. (1978). EEG sleep and depression. In Sleep Disorders: Diagnosis and Treatment (R. L. Williams and I. Karacon, eds.) 163–209. Wiley, New York.

MULLEN, P. E., LINSELL, C. R. and PARKER, D. (1986). Influence of sleep disruption and calorie restriction on biological markers for depression. Lancet 1051–1054.

Comment

Craig D. Turnbull

Everitt notes that psychiatry is a relatively young discipline as compared to other branches of medicine since its theories regarding the etiology and treatment of psychiatric disorders are in their infancy. He reports that psychiatrists have become increasingly aware that to build and to advance their theories requires well designed quantitative studies in combination with the use of appropriate statistical tools in order to properly evaluate the results of such studies.

While I concur with these assessments, I wish to provide additional material regarding the historical background and various efforts to develop a nosology (or classification) of mental disorders. This discussion will contain the major portion of my comments since the nosology of a science serves as the grist (i.e., the diagnosis of a case) which yields the data which psychiatrists wish to analyze. Although I will only provide a sketch of the developments which have led to the current nosology (DSM-III), I wish to note that considerable advances have been made in the fields of psychiatry and psychiatric epidemiology since 1800.

Psychiatry has indeed struggled to attain the "dignity of science" by submitting its observations to measurement and quantification. In fact, Grob (1985) has noted that after 1800, several currents converged to create a type of social inquiry whose methodological distinctiveness was a commitment to quantitative research. Underlying this urge to quantify was the assumption that such a methodology could explain social phenomena. He also reported that early and mid-nineteenth century commentators were preoccupied with the development of elaborate classification systems and an almost obsessive concern with the collection of statistical data. That the field of epidemiology emerged in such an environment was not surprising.

Grob observed that American psychiatrists were among the staunchest proponents regarding the collection of statistical data. The annual reports of mental hospitals as well as the American Journal of Insanity, which was first published in 1844, included statistics on the demographic and geographical characteristics of mentally ill patients as well as the results of therapy. In addition, the federal census of 1840 provided some aggregate data on the mentally ill population.

Although nineteenth century psychiatrists were avid data collectors, their approach to statistical data consisted of the following interests:

1. They used statistics to demonstrate high "cure" rates.
2. They thought that the collection of data would help to uncover laws governing health and disease.
3. They used statistics for purposes of policy advocacy.
4. They viewed statistical data as a means to establish the legitimacy of public mental hospitals and to build support among state officials and the public.

Mid-nineteenth century psychiatrists thought there was a direct relationship between the rising incidence of mental illness and the advance of civilization. Such

Craig D. Turnbull is Associate Professor of Biostatistics, School of Public Health, University of North Carolina, Chapel Hill, North Carolina 27514.

beliefs did not arise from the application of epidemiologic methodology, but reflected instead their social and religious views.

Prevailing nineteenth century psychiatric nosologies were inadequate. The link between organs and behavior was unknown. Psychiatrists attempted to identify the presence of disorders by observing external signs and symptoms. Also, nineteenth century psychiatrists recognized that their statistics dealt not with incidence, but rather with hospital admissions.

Grob also reported that the impetus to create an epidemiology of mental disorders came largely from outside of psychiatry or medicine. Toward the end of the nineteenth century, new social science disciplines had come into existence. Many of the individuals associated with these new disciplines were concerned not only with developing a scientific understanding of individual and social behavior, but they also wished to apply such knowledge to social problems. In their search for empirical data, the social scientists drew upon the tradition of statistical analysis that had emerged. Ultimately a concensus developed regarding the utility of a federal census. The social scientists did not view the census merely as an instrument to collect data. It also represented their opinion that statistical knowledge could serve as a foundation for social policy.

Although the census of 1840 attempted to enumerate the insane, its deficiencies were overwhelming. The special census of 1904 focused attention on the ethnic and racial characteristics of the institutionalized mentally ill. The 1910 census was even broader in scope than its predecessor. By this time, the relationship between the Census Bureau and the social sciences had been institutionalized. The 1910 census reflected a growing sophistication in analyzing statistical data.

For instance, Hill (U. S. Bureau of the Census, 1914) insisted that the claim insanity was increasing rapidly was dubious. He suggested the following issues shaped the statistics of insanity:

1. Growing practices of institutionalizing the insane.
2. Increasing average length of life.
3. New diagnostic methods in psychiatry leading to the detection of mental factors in physical cases.
4. Establishment of dispensaries.
5. Provision for voluntary and emergency commitment.
6. Better modes of transportation that made it possible to bring individuals in poor physical condition to mental hospitals.

The growing involvement of the Census Bureau with the statistics of mental illness was suggestive of the concern outside of psychiatry with policy implications and future trends. Grob noted that within psychiatry there was initially little interest in the activities of the Census Bureau. This began to change after the founding, in 1909, of the National Committee for Mental Hygiene.

In 1917 the American Medico-Psychological Association's Committee on Statistics urged all mental hospitals to adopt a uniform reporting system. With the assistance of the National Committee for Mental Hygiene, this Association produced the first uniform nomenclature of mental disease in 1918.

The adoption of a formal nomenclature reflecting statistical concerns was evidence of the growing maturity of psychiatric epidemiology. By 1920 the Bureau of the Census compiled its own nomenclature of diseases that included psychiatric disorders.

After World War II there was a literal explosion of community and demographic studies of the mentally ill. The National Mental Health Act was passed in 1946. A result of this Act was that responsibility for gathering data on the mentally ill was transferred to the Public Health Service and the soon-to-be-created National Institute of Mental Health.

Shepherd (1985) reported that the 1949 Milbank Memorial Fund Conference on the epidemiology of mental disorders was a seminal occasion, at which there was general agreement on the importance of epidemiology for causal research and for administrative policy. Its relevance to clinical psychiatry, however, was disputed by many of the practicing psychiatrists at the Conference who questioned how far epidemiological inquiry should be based on the conventional schemata of disease which, in their opinion, were inapplicable to mental disorders. Gordon (see Shepherd (1985), page 275), however, emphasized the study of the clinical case as the point of departure, pointing out that what we choose to call mental disease, is an artificial grouping of many morbid entities, and that in constructing an epidemiology of mental disorder investigators should focus attention on specific conditions which have features in common with other mass diseases well understood in their group relationship. In addition, Francis (see Shepherd (1985), page 275) reminded the Conference that epidemiology is basically dependent upon the accuracy of diagnosis and that until a valid basis for classification can be generally employed, data from different areas cannot be properly compared.

The third and current (1980) edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III) reflects a commitment to "data" as its basis. In brief, the American Psychiatric Association has recognized that progress in mental health research depends on adequate methods to collect, organize, and analyze data.

In order to more fully appreciate the current state of affairs, it is useful to note that the first edition of

the DSM (1952) reflected Adolf Meyer's theoretical, psychobiological view that mental disorders were reactions of the personality to psychological, social, and biological factors. Whereas, the current edition has a more descriptive approach which is an outgrowth of the growing recognition of the importance of diagnosis for both clinical pràctice and research. That is, clinicians and researchers must have a common language with which to communicate. For planning a treatment program must begin with an accurate diagnostic assessment; and, efficacy of various treatment modalities can be compared only if patient groups are described using diagnostic terms which are clearly defined.

Classification is implicit in the terms and methods of psychiatric epidemiology. Types of classification criteria and methods of observation and measurement change as our knowledge of the disease increases. Progress in developing objective procedures for observation and measurement correlate with the development of a scientific understanding of disease.

The trend in medical science has been from subjective observation of symptoms to objective and repeatable laboratory tests. For instance, positron-emission tomography is being used experimentally to visualize levels of activity of specific substances in various portions of the brain. These patterns of activity are correlated with diagnoses according to traditional procedures based on psychiatric symptomology. Bagne (1980) suggested that a standardized classification does this by helping to transcend different theoretical orientations and foci of attention that individuals bring to mental health practice and research. For instance:

Within research:
It contributes to the vigorous formulation of hypotheses to test. It also makes it easier to communicate methods and results so that others can repeat our observations. This is the hallmark of science.

Within practice:
Diagnosis (classification in a medical context) is a universal first step when patients are admitted to treatment. Disciplined diagnosis is based on relevant information, appropriate discriminations, and rigorous application of explicit criteria. Once established, the diagnosis is used to select the treatment and to communicate the information to other individuals who care for the patient.

Between research and practice:
A classification can bring the results of research to bear on clinical practice. It can bring insights gained from clinical practice to scientific test.

A mental health classification extends conceptualizations of mental disorders. As such the development of a classification is an expression of scientific creativity as well as a product of statistical procedures. A useful mental health classification serves as a tool to gather data which may eventually change our mental health concepts and result in a new classification. This process explains why a classification is usually revised after several years of use.

The approach taken in DSM-III is atheoretical with respect to etiology or pathophysiological process except for those disorders for which this is well established and therefore included in the definition of the disorder. The major justification for the generally atheoretical approach is that the inclusion of etiological theories would be an obstacle to the use of this classification by clinicians of varying theoretical orientations, since it would not be possible to present all reasonable etiological theories for each disroder. Because DSM-III is generally atheoretical with regard to etiology, it attempts to describe comprehensively what the manifestations of the mental disorders are. This approach can be said to be "descriptive" in that the definitions of the disorders generally consist of descriptions of the clinical features of the disorder.

Everitt has noted an increased use of statistical methods in the psychiatric literature. He cites data by DeGroot and Mezzich (1985) which show that in 1980 the most commonly used statistical techniques were simple t and χ^2 tests, viz. 28% for the *American Journal of Psychiatry*, 51% for the *British Journal of Psychiatry*, and 60% for the *Archives of General Psychiatry*. It is interesting to note that these percentages compare to 39% reported for the *New England Journal of Medicine* for 1978 and 1979 (see Emerson and Colditz, 1983). Emerson and Colditz' review of the *New England Journal of Medicine* also indicated that the reader who was, at most, conversant with descriptive statistics had access to 58% of its articles. This figure compares with 58% for the *American Journal of Psychiatry*, 22% for the *British Journal of Psychiatry,* and 15% for the *Archives of General Psychiatry*, (see DeGroot and Mezzich, 1985).

These data suggest the consulting statistician might consider offering his/her expertise to health colleagues in order to design and conduct appropriate quantitative studies which employ state of the art statistical methods. In brief, more powerful statistical procedures are available for potential use than are currently reported in much of the psychiatric literature.

For instance, the following remarks highlight Everitt's excellent demonstration of the use of Cox's (1972) regression model to analyze censored survival data. His reanalysis of a clinical trial of bromocriptine demonstrates how one might attempt to provide a clear answer to a query which on the surface may appear straightforward, but which in fact is constrained by differential dropouts and failures to achieve some "improvement" criterion. His methods provided estimates regarding the query of interest which were more realistic than the simple averages of

uncensored observations offered by the original investigators. In addition, his reanalysis of data regarding the factors associated with the length of stay of mentally ill patients at Broadmoor is an example of the types of analytical gains which can be had by using the actual length of stay as a dependent variable in a Cox's proportional hazards model, given that observations which pertained to patients still in detention at the time of the study were considered censored. His findings were substantially more revealing than those given by the original investigators since they merely categorized the length of stay variable (short, intermediate, and long) and compared it with other explanatory variables via a series of simple χ^2 tests.

Everitt offers an interesting alternative approach regarding teaching statistics to psychiatrists. I concur with his appraisal that in order to conduct the type of course he has suggested considerably more time and effort, by both the instructor and the student, would be required as compared to a conventional service course. Also, the type of course Everitt describes would be appropriate for a homogeneous set of students— say, medical residents.

I have taught service statistics courses in Schools of Public Health, Medicine, and Nursing for a number of years. As such I've experienced the seemingly myriad of competing priorities which impinge on these students. In order to attempt to deal with some of these factors, our faculty (Department of Biostatistics, School of Public Health, The University of North Carolina at Chapel Hill) currently offers three different service courses. Each of these courses covers elements of descriptive and inferential statistics; but, they differ in student backgrounds assumed, the depth to which they go into theoretical issues, and the speed

with which they move through the material. However, only our third-level course requires students to complete assignments using various computer packages; and, these students are a heterogeneous set of bachelor's, master's, doctoral, and postdoctoral individuals from many health disciplines. The group discussion aspects which Everitt suggests would probably not fare well for such classes.

I would suggest the following to any applied statistician who contemplates collaborating with allied health professionals:

1. There are vast differences in the types and analytical levels of training to which the myriad of allied health professionals are exposed.
2. Health professionals operate in subgroups— areas of specialization. It is necessary to know and work through existing hierarchies.
3. Understanding and co-operation is fundamental to collaboration.

ADDITIONAL REFERENCES

AMERICAN PSYCHIATRIC ASSOCIATION. (1980). *Diagnostic and Statistical Manual of Mental Disorders*, 3rd ed. American Psychiatric Assoc., Washington.

BAGNE, C. (1980). Personal communication. Dr. Bagne was a postdoctoral fellow under the direction of Dr. Turnbull, 1979–1980.

EMERSON, J. D. and COLDITZ, G. A. (1983). Statistics in practice: use of statsitical analysis in the *New England Journal of Medicine*. New England J. Med. **309** 709–713.

GROB, G. N. (1985). The origins of American psychiatric epidemiology. *Amer. J. Public Health* **75** 229–236.

SHEPHERD, M. (1985). Psychiatric epidemiology and epidemiological psychiatry. *Amer. J. Public Health* **75** 275–276.

U. S. BUREAU OF THE CENSUS. (1914). *Insane and Feeble-Minded in Institutions: 1910*. Government Printing Office, Washington, D.C.

Comment

Joseph S. Verducci

I congratulate Dr. Everitt for his sure-footed climb up the mountain that is psychiatric statistics. His narration of consulting encounters strikes a few shivers of recognition from my own work at the Mental Research Institute (MRI) in Palo Alto and the Western Psychiatric Institute and Clinic (WPIC) in Pittsburgh. The chilling effect is from the enormity of the work that psychiatric researchers have undertaken.

Joseph S. Verducci is Assistant Professor, Department of Statistics, Ohio State University, 1958 Neil Avenue, Columbus, Ohio 43210.

The practice of "mind-healing" has grown from Franz Mesmer's gazes into the 18th century psyche to PEP scans of glucose glowing in 20th century brains. Since its inception, psychiatry has been developing much too rapidly to accommodate the slow scrutiny of physical scientists. For example, it wasn't until Thurstone's work (1927) that Fechner's (1859) experiments on psychophysics could be analyzed, and not until Mosteller's work (1951) that they could be formally analyzed. Psychiatry would have probably evolved much differently if Freud had waited for a statistician to analyze the data he had amassed on "free" word associations.