The Biometric Approach to Neuropsychopharmacology

Joseph Zubin*

Veterans Administration Hospital, Pittsburgh, Pennsylvania 15206; and The University of Pittsburgh, School of Medicine, Pittsburgh, Pennsylvania 15260

CURRENT STATUS

The impact of biometric methods on neuropsychopharmacology may be likened to that of the microscope on biology for it opened up new vistas for objectively distinguishing between previously unknown effects of various drugs. In the words of Karl Jaspers (4): "The biometric methods give us more than figures and correlations. They foster clarity in all fields in which biometric variations can be established. Moreover, through the application of these methods we have concrete experiences which we would never have had without them . . ." Biometrics became an independent discipline under Karl Pearson toward the end of the 19th century. It was an outgrowth of the quantitative developments in psychology, biology, medicine, and public health; as the demand for quantification increased, the biometric method was developed to meet these needs. When psychopathologists began to feel the need for quantification in the wake of therapeutic interventions that produced rapid, observable changes in behavior, biometrically minded psychologists and psychiatrists attempted to provide such measures. It soon became apparent that the available clinical tools used in assessing patient behavior, the clinical interviews, case histories and tests, were insufficient to meet the task, and rating scales were introduced to capture the essence of the observed changes. This gave rise to such instruments as the Phipps Psychiatric Clinic Behavior Chart, the Malamud Sands Rating Scale (8), the Wittenborn Psychiatric Rating Scale (11), the Lorr Multidimensional Scale (7), and a variety of other scales for rating the changes in behavior brought on by the somotherapies from the 1930s to the 1950s. The golden age of biometrics, when the creative thrust was most apparent, reached its zenith in the 1950s. Since then, the rating scales have been consolidating their gains and bringing law and order into the captured beachheads of quantification.

As the treatments introduced by the development of neuropsychopharmacology became more specific to only certain types of patients, it became necessary to develop more precise classification of patients to reduce the tremendous variability among patients belonging to classic diagnostic groups. This would permit more homogeneous groups for testing the variety of drugs that became available. Fortunately, the need for better diagnostic methods had been foreseen even before the drug era. The Mental Status Examination, which had been the vade mecum of the psychiatric resident, provided the starting point for improving diagnosis. The first step was taken in 1944 by N.D.C. Lewis when he invited this author to help him in the revision of the Mental Status Examination (6). It became clear that the narrative report of this examination often failed to record all the information which the examiner required; a mimeographed check-list was provided for recording each of the items in the examination, including ratings of a variety of traits. Unfortunately, World War II intervened and the project had to be given up until the Biometrics Research Unit was created in 1956 (15). Thus this volume marks the 20th anniversary of the establishment of the Biometrics Research Unit, which was given the specific mandate of improving diagnosis and prognosis of mental disorders with the view of better evaluation of outcome of treatment. The first effort in this direction was the development of a Ward Behavior Inventory (1) in which nurses and attendants rated patient behavior before and after treatment to determine the amount of change observed.

* Formerly Chief of Psychiatric Research (Biometrics), Biometrics Research Unit, New York State Department of Mental Hygiene.
The next step was to apply this approach of rating specific behavior to the assessment of the patient's mental status as revealed by the Mental Status Examination. To this end a systematic structured interview based on the Mental Status Examination was developed (9,14). This led to further development of interview methods, which turned the unreliable and often invalid blunderbuss of the clinical interview into a sharpshooting rifle aimed at specifying the signs and symptoms of psychopathology present in the patient. The first application of these techniques was not in neuropsychopharmacology but in the cross-cultural studies of diagnosis, which required the assessment of mental patients in different cultural settings by means of the same instruments of known reliability. The U.S.-U.K. (2) study of diagnosis, which limited itself to two countries but covered the entire spectrum of the mental disorders, and the complementary study of the WHO Pilot Study of Schizophrenia (12), which, while limiting itself to only one disorder, covered the spectrum of nine different cultures, demonstrated the utility of the structured clinical interview and its superiority to the unstructured approach.

Besides providing better interviewing methods which gave a more precise description of the behavior of the patient, it also became necessary to provide better methods for integrating the results of the interview into some type of classification system. Interview results could be analyzed into underlying dimensions of psychopathology. Each patient could be given quantitative scores for each of the dimensions that constitute the anatomy of psychopathology, but this dimensional analysis did not provide classificatory categories of diagnosis. In order to provide such diagnoses, despite the absence of any known etiology, Hempel (3) and Stengel (10) suggested that an operational approach be provided which established the specific consensual criteria that a patient must meet in order to qualify for a given diagnosis. This operational approach to classification did not gain much headway until the newer therapeutic pharmacological methods forced the clinicians to become more precise in their diagnosis and until the knowledge of the findings of the U.S.-U.K. project regarding the misdiagnosis of patients in both countries became widespread. The development of the operational approach involving the Research Diagnostic Criteria (RDC) is detailed in the chapter by Spitzer et al. (*this volume*).

In reviewing the contributions of the biometric approach to the solution of the problems of neuropsychopharmacology during the past decade, it became apparent that varying degrees of progress have been made in the following areas: (a) reliability of interviewing method for descriptive psychopathology, (b) reliability of diagnosis, (c) validity of diagnosis, and (d) evaluation of outcome of treatment.

Considerable progress has been made in the development of reliable instruments and in the provision of criteria for more reliable diagnostic methods.

Concurrent validity (a form of reliability) seems to be satisfactory and computerized diagnoses add to the attainment of concurrent validity. With regard to content validity, we have also been successful insofar as the interviewing methods contain sufficient items to cover the entire spectrum of psychopathology or can be extended to do so.

The degree of success in attaining predictive validity is rather modest. Generally, patients diagnosed as schizophrenic tend to require longer hospital stays than those diagnosed as depressives; even depressives who are misdiagnosed as schizophrenics seem to have shorter hospital stays. The process of selecting the optimum treatment has not reached a high level of accuracy when evaluated against outcome.

Regarding construct validity, only the very beginnings have been achieved.

**CRITIQUE AND FUTURE DEVELOPMENTS**

Despite the generally optimistic assessment of the current status of the application of biometric methods to neuropsychopharmacology, some serious shortcomings exist.

Since the etiology of the mental disorders remains largely unknown, the success attained in concurrent and content validity is not enough because these two types of validity are based on consensual agreement rather than on objective criteria. Even predictive validity is not sufficient since prediction is generally rather weak. Demonstrations that certain diagnostic categories tend to profit from specific therapies, while encouraging the belief in the validity of diagnoses, are not in themselves convincing as long as we do not know how these therapies work. Do they tackle the focal disorder which the diagnostic category presumably represents or do they merely mitigate the peripheral effects of this
disorder? We must remember that even in the physical disorders whose focal source is known, there is considerable variability in how the illness expresses itself, depending on the premorbid personality, the milieu, and other environmental factors. In the mental disorders, where there is no demonstrable focal disorder, it is difficult to know whether the therapeutic agent affects the focal disorder or its penumbra, i.e., the behavioral response to the focal disorder. For this reason it is hazardous to accept therapeutic response as evidence for the validity of the diagnosis.

What can be done to establish the validity of a diagnosis when its focal cause or causes are unknown? The only method available is the conceptualization of “as if” causes and testing their tenability. This is the process of developing scientific models, and the testing of the hypotheses emanating from these models could provide the bases for the construct validity of diagnostic systems.

There are currently six scientific models that preempt the field (16) of etiology of the mental disorders. Thus the focal disorder may lurk in (a) the ecological niche the person occupies; (b) the quirks of his/her development; (c) his/her learning and conditioning history; (d) his/her genetic inheritance; (e) his/her learning and conditioning history; (d) his/her genetic inheritance; (e) his/her deviant internal environment; or (f) his/her neurophysiological anomalies. Each of these models can provide hypotheses for testing the construct validity of the diagnostic schema. Here we deal with only the last three models: genetic, internal environment, and neurophysiological models, the other models being discussed elsewhere (J. Zunin, in preparation). Spitzer et al. (this volume) have already pointed to a construct validity test based on the genetic model. They indicated that Gottesman and Shields found that neither the loose American diagnosis of schizophrenia nor the more rigid European diagnosis indicated the highest concordance rate for schizophrenia in identical twins, while a “middle of the road” diagnostic approach showed the highest degree of concordance.

Thus the “middle of the road” diagnostic schema shows the highest construct validity for the genetic model for etiology.

Another test of the construct validity of diagnoses based on the genetic model is provided by the demonstration that bipolar but not unipolar depressives show a prevalence of depression in their close relatives that is higher than expected from the general population. This lends more support for the differentiation of bipolar from unipolar depression.

With respect to the internal environmental model, such indicators as monoamine oxidase levels in the blood platelets of schizophrenic probands and in their unaffected monozygotic co-twins (13) afford a construct validity test of the diagnosis of schizophrenia; quite a number of other indicators emanating from the internal environment are referred to in this book.

The neurophysiological model searches for indications of psychopathology in the processing of information via the central nervous system of the patient. Such measures as evoked potentials, sensory integration, auditory thresholds, reaction time, pupillography, eye tracking, etc., have been found to differentiate between same diagnostic groups and constitute construct validity evidence for the diagnoses. Among the recent developments in this area is the introduction of signal detection methods (18) and forced choice method, which have been introduced into the testing of mental patients in order to differentiate between sensitivity and criterion differences. Another new development is the search for techniques in which the patient “excels” over his/her controls (18). This trend has been developed in order to overcome the ever present danger of finding differences between patients and normals simply because patients are generally not as well motivated or attentive. Still another problem that is now being tackled is the attempt to separate trait-related from state-related differences between patients and normals (17).

One of the major stumbling blocks in the application of biometric methodology to investigation in psychopharmacology is the need for large samples, which are often difficult to collect. It is interesting to note that sequential analysis is rarely reported as a technique in studies in neuropsychopharmacology. Yet such an analysis could, in most instances, reduce the number of cases required for establishing or disestablishing the presence of a significant difference (5).

In summary, although considerable progress in the application of the biometric method to the problems of neuropsychopharmacology has taken place, this progress has opened new territories with ever increasing new problems to solve. It is heartening, as Professor Hamilton (this volume) indicates, that the progress has not been unilateral. The impact has been reciprocal and neuropsychopharmacology has also benefited from the biometric thrust. This thrust
has so permeated the field that this section devoted to biometrics does not contain all of its contributions; many of them will be found scattered in other sections of this volume. If there is a common denominator to neuropsychopharmacology with all its interdisciplinary sciences, it will be found in the common language of objectivity and measurement, which biometrics provides.

ACKNOWLEDGMENTS

The preparation of this chapter was aided by a grant from the Veterans Administration on “Vulnerability to episodes of functional psychosis in adults” and the Department of Psychiatry, University of Pittsburgh.

REFERENCES