Distinguishing Between the Validity and Utility of Psychiatric Diagnoses

Robert Kendell, M.D.

Assen Jablensky, M.D.

Objective: The meaning of the terms "validity" and "utility" as they apply to psychiatric diagnoses is examined.

Method: The authors discuss the concepts of validity, utility, and disease; review assumptions that have been made about mental disorders as disease entities; and examine the evidence that mental disorders are separated from one another and from normality by natural boundaries (zones of rarity).

Results: Despite historical and recent assumptions to the contrary, there is little evidence that most currently recognized mental disorders are separated by natural boundaries. Researchers are increasingly assuming that variation in symptoms is continuous and are therefore questioning the validity of contemporary classifications.

Conclusions: It is important to distinguish between validity and utility in considering psychiatric diagnoses. Diagnostic categories defined by their syndromes should be regarded as valid only if they have been shown to be discrete entities with natural boundaries that separate them from other disorders. Although most diagnostic concepts have not been shown to be valid in this sense, many possess high utility by virtue of the information about outcome, treatment response, and etiology that they convey. They are therefore invaluable working concepts for clinicians.

(Am J Psychiatry 2003; 160:4–12)

The introduction of explicit diagnostic criteria and rule-based classifications, such as DSM-III and its successors and ICD-10, has profoundly affected at least four domains of psychiatric practice. A standard frame of reference has been endorsed by most clinicians, enabling them to achieve better diagnostic agreement and improve communication, including statistical reporting on psychiatric morbidity, services, treatments, and outcomes. More precise diagnostic criteria and instruments have become the norm in research. Although most research diagnostic criteria are still provisional, they can be refined or rejected by using empirical evidence. Teaching is now based on an international reference system that provides a worldwide common language. And public access to the diagnostic criteria used by mental health professionals has helped improve communication with the users of services, caregivers, and society at large.

Although these gains must be acknowledged, it is important not to overlook the inadequacies and failings of contemporary classifications. At present, psychiatry is in a state of flux, and advances in neuroscience and genetics are soon likely to challenge many of its current theoretical underpinnings, particularly those related to the causation and definition of mental disorders. New treatments targeting specific functional systems in the brain will require a better differentiated classification of the clinical populations likely to benefit. The emerging realization that in every culture mental disorders account for a much larger share of the total burden of disease than previously assumed will also raise complex questions about the cost-benefit of interventions, the right to treatment, equity in access to treatment, and the feasibility of prevention.

Against this background, it is increasingly recognized that the validity of the diagnostic concepts enshrined in contemporary classifications of mental disorders cannot be taken for granted. The reliability of psychiatrists' diagnoses was dramatically improved, at least in research settings in which structured interviews were used, by the introduction of explicit definitions and decision rules in DSM-III and in the ICD-10 Diagnostic Criteria for Research (1). This at least partial solution to the reliability problem has shifted attention to the more fundamental issue of the validity of psychiatric diagnoses, and there is now a fairly widespread view that if future editions of these two classifications are to be a significant improvement on their predecessors, the validity of the diagnostic concepts they incorporate will have to be enhanced (2, 3). It has also been suggested (for example, at an American Psychiatric Association Research Planning Conference in Washington, D.C., in October 2000) that in the future it would be useful to score the various diagnoses listed in the DSM and ICD to indicate the extent to which each had been validated.

The Concept of Validity

Because the validity of diagnostic categories, and of their defining criteria, is becoming a topical issue, it is important that there should be no ambiguity about what is

implied by the term "validity." However, the meaning of validity in the context of diagnosis has never been adequately clarified. The word "valid" is derived from the Latin *validus*, meaning strong, and it is defined as "well-founded and applicable; sound and to the point; against which no objection can fairly be brought" (4). There is no single, agreed upon meaning of validity in science, although it is generally accepted that the concept addresses "the nature of reality" (5, pp. 456–476) and that its definition is an "epistemological and philosophical problem, not simply a question of measurement" (6, p. 121). In logic, validity is the characteristic of an inference that must be true if all its premises are true. Psychologists generally adopt the American Psychological Association's distinction between content, criterion-related, and construct validity (7), and their main concern has always been with the validity of psychological tests. Borrowing terms from psychometric theory, psychiatrists have mainly been concerned with concurrent and predictive validity, partly because of their relevance to the issue of the validity of diagnoses. Certainly, the ability to predict outcome, both in the absence of treatment and in response to specific therapies, has always been a crucial function both of physicians and of their diagnoses. Indeed, Goodwin and Guze (8) went so far as to assert that "diagnosis is prognosis" and referred approvingly to P.D. Scott's observation that "the follow up is the great exposer of truth.…It is to the psychiatrist what the post-mortem is to the physician."

Criteria for Establishing the Validity of Psychiatric Diagnoses

Robins and Guze (9) were probably the first to propose formal criteria for establishing the validity of psychiatric diagnoses, and their views have been very influential, particularly in North America. They listed five criteria: 1) clinical description (including symptom profiles, demographic characteristics, and typical precipitants), 2) laboratory studies (including psychological tests, radiology and post-mortem findings), 3) delimitation from other disorders (by means of exclusion criteria), 4) follow-up studies (including evidence of diagnostic stability), and 5) family studies. They used these five criteria to show that "good prognosis schizophrenia is not mild schizophrenia but a different illness," a demonstration that subsequently underpinned the distinction in DSM-III between schizophrenia and schizophreniform disorder.

This schema was elaborated by Kendler (10), who distinguished between antecedent validators (familial aggregation, premorbid personality, and precipitating factors), concurrent validators (including psychological tests), and predictive validators (diagnostic consistency over time, rates of relapse and recovery, and response to treatment). He then used these expanded criteria to demonstrate that paranoia (simple delusional disorder) is probably a distinct syndrome rather than a mild form of schizophrenia or a subtype of affective illness.

More recently, Andreasen (11) proposed "a second structural program for validating psychiatric diagnosis" and listed several additional validators—molecular genetics and molecular biology, neurochemistry, neuroanatomy, neurophysiology, and cognitive neuroscience—that are all potentially capable of linking symptoms and diagnoses to their neural substrates.

The Implicit "Disease Entity" Assumption

Thoughtful clinicians have long been aware that diagnostic categories are simply concepts, justified only by whether they provide a useful framework for organizing and explaining the complexity of clinical experience in order to derive inferences about outcome and to guide decisions about treatment. Unfortunately, once a diagnostic concept such as schizophrenia or Gulf War syndrome has come into general use, it tends to become reified. That is, people too easily assume that it is an entity of some kind that can be invoked to explain the patient's symptoms and whose validity need not be questioned. Even though the authors of contemporary nomenclatures may be careful to point out that "there is no assumption that each category of mental disorder is a completely discrete entity with absolute boundaries dividing it from other mental disorders or from no mental disorder" (DSM-IV, p. xxii), the mere fact that a diagnostic concept is listed in an official nomenclature and provided with a precise, complex definition tends to encourage this insidious reification.

The weakness of the validity criteria of both Robins and Guze and Kendler was that those criteria implicitly assumed that psychiatric disorders are discrete entities and that the role of validity criteria is to determine whether a putative disorder, such as "good-prognosis schizophrenia" or paranoia, is a valid entity in its own right or a mild form or variant of some other entity. The possibility that disorders might merge into one another with no natural boundary in between—what Sneath (12) called a "point of rarity," but what is better regarded as a zone of rarity—was simply not considered. Robins and Guze commented, for example, that "the finding of an increased prevalence of the same disorder among the close relatives of the original patients strongly indicates that one is dealing with a valid entity" (9). In reality, such a finding is equally compatible with the existence of continuous variation. It seems that the possibility of an increased prevalence of more than one disorder in the patients' first-degree relatives had not occurred to Robins and Guze. In a similarly revealing comment, they wrote that "the failure to achieve 100 percent success in predicting outcome and the overlap in the results of the family studies indicate that the criteria used for the separation need further refinement." They did not consider that the results might have occurred because no

natural boundary exists between good- and poor-prognosis cases, in which case the limitations they identified would have resisted all attempts at improvement based on refinement of the diagnostic criteria.

Robins and Guze's classic paper was written at a time when it was widely assumed that schizophrenia and manic-depressive (bipolar) disorder were transmitted by single—or at the most two or three—genes and before publication of the first studies examining whether there were "zones of rarity" between related syndromes. The situation now is quite different. Several attempts have been made to demonstrate natural boundaries between related syndromes or between a common syndrome such as major depression and normality, either by locating a zone of rarity between them (13, 14) or by demonstrating a nonlinear relationship between the symptom profiles and a validating variable such as outcome or heritability (15, 16). Most such attempts have ended in failure. Several general population surveys have also demonstrated that quite minor differences in the definition of individual syndromes such as major depression may result in large differences in recorded prevalence (17, 18), again suggesting that the boundary identified by the definition does not correspond with a natural zone of rarity.

At the same time, research is increasingly supporting the view that many different genes contribute to the etiology of most of psychiatry's major syndromes and that some of these genes are risk factors for what have until now been regarded as unrelated syndromes. Several other DSM/ICD disorders have been found to cluster among the relatives of individuals with schizophrenia, major depression, or bipolar affective disorder, and findings of such clusters have given rise to the concepts of "schizophrenia spectrum" and "affective spectrum" disorders. Increasing evidence also suggests that several genetic susceptibility loci may be common to two or more clinically distinct disorders. For example, three of the putative susceptibility loci associated with bipolar disorder (on chromosomes 13, 18, and 22) seem also to contribute to the risk of schizophrenia (19). In addition, the microdeletion in region q11 on chromosome 22, which underlies the velocardiofacial syndrome, appears to be associated with a higher incidence of mental retardation, schizophrenia, and bipolar affective disorder as well (20). Furthermore, the genetic basis of generalized anxiety disorder appears to be very similar to, if not indistinguishable from, that of major depression (21), and the genetic basis of schizophrenia seems to encompass a spectrum of other disorders, including schizotypal/paranoid personality disorder and even psychotic affective illness (22). It will not be surprising if in time such findings of overlapping genetic predisposition to seemingly unrelated disorders become the rule rather than the exception. It is equally likely that the same environmental factors contribute to the genesis of several different syndromes. Sexual or physical abuse and neglect in childhood, for example, seem to increase the risk of both anxiety and depression in adult life (23), and sexual abuse may increase the risk of bulimia nervosa and of alcohol and other substance dependence as well (24).

Increasing Disenchantment With the "Disease Entity" Assumption

Although ubiquitous in both medical and lay discourse, the term "disease" has no unambiguous, generally accepted definition. However, as Scadding (25) pointed out, most of those using this term "allow themselves the comfortable delusion that everyone knows what it means." Albert et al. (26) catalogued six general views or concepts about "what types of conditions may be said to constitute a disease," ranging from nominalism and cultural-relativistic theories (i.e., something becomes a disease when a profession or society labels it as such) and social idealism (failure to attain a social ideal of perfect health) to culturally normative statistical concepts (deviation from statistically defined normality) and the "disease realism" view (objectively demonstrable departure from adaptive biological functioning). In adopting the last model as the one best suited to the present state of medicine, they emphasized that the clinical signs and symptoms do not constitute the disease and that it is not until causal mechanisms are clearly identified that "we can say we have 'really' discovered the disease" (26).

Although each of these general concepts of disease has been used by psychiatry at some time, it is the "disease realism" model (in both its biological and psychodynamic versions) that has dominated the debate since the end of the 19th century. Kraepelin—a staunch "disease realist"—long believed that dementia praecox and manic-depressive insanity, defined by painstaking clinical observation of their symptoms and outcome, represented distinct species of brain disease whose causal mechanisms would ultimately be discovered by neuropathology, experimental psychology, and genetics. Eventually, however, he abandoned his assumption that these two disorders were discrete entities and proposed instead a model that was essentially dimensional (27). About the same time, Jaspers (28) wrote that "the idea of the disease-entity is in truth an idea in Kant's sense of the word: the concept of an objective which one cannot reach…but all the same it indicates the path for fruitful research and supplies a *valid* point of orientation for particular empirical investigations" (p. 569). He then added that, although "the idea of disease-entities has become a fruitful orientation for the investigations of special psychiatry…no actual disease-entities exist" (p. 570).

The relevance of this view to the present taxonomic debate in psychiatry is twofold. First, discrete disease entities and dimensions of continuous variation are not mutually exclusive means of conceptualizing psychiatric disorders; both are compatible with a threshold model of disease and may account for different or even overlapping segments of

psychiatric morbidity. Second, the surface phenomena of psychiatric illness (i.e., the clustering of symptoms, signs, course, and outcome) provide no secure basis for deciding whether a diagnostic class or rubric is valid, in the sense of delineating a specific, necessary, and sufficient biological mechanism.

Several well-informed commentators have produced evidence suggesting that there may be no natural boundary between recognized mental disorder and normality or health (16, 29, 30). Widiger and Clark (31) suggested that variation in psychiatric symptoms may be better represented by "an ordered matrix of symptom-cluster dimensions" than by a set of discrete categories. Cloninger (32) stated firmly that "there is no empirical evidence" for "natural boundaries between major syndromes," that "no one has ever found a set of symptoms, signs, or tests that separate mental disorders fully into non-overlapping categories," and that "the categorical approach…is fundamentally flawed." Frustrated by the failure of two decades of laborious research to identify any of the genes underlying the major psychiatric syndromes, Ginsburg et al. (33) complained that "current nosology, now embedded in DSM-IV…does not define phenotypes for genetic study." Comorbidity poses a further problem that is becoming increasingly clamant as its full extent is revealed by community studies. As Sullivan and Kendler (34) commented, the scale of the apparent comorbidity between major depression, various anxiety disorders, and addictive syndromes is "not consistent with the orthodox conceptualization of these psychiatric disorders as discrete nosological entities." The accumulation of such evidence and opinions led Allen Frances, the chairperson of the task force that produced DSM-IV, and Helen Egger (35) to comment gloomily, but perhaps presciently, that "we are at the epicycle stage of psychiatry where astronomy was before Copernicus and biology before Darwin. Our inelegant and complex current descriptive system will undoubtedly be replaced by…simpler, more elegant models."

This disenchantment is understandable in the light of the failure of the revolutionary new nosology provided by DSM-III and its successors to lead to major insights into the etiology of any of the main syndromes. But disillusionment may not yet be justified. Although there is a growing assumption, at least within the research community, that most currently recognized psychiatric disorders are not disease entities, this belief has never been demonstrated, mainly because studies of the appropriate kind have rarely been mounted (36).

Distinguishing Between Discrete Entities and Continuous Variation

The crucial issue is whether psychiatric syndromes are separated from one another, and from normality, by zones of rarity or whether they are merely arbitrary loci in a multidimensional space in which variation in both symptoms and etiology is more or less continuous. The mere existence of interforms between one syndrome and another, or between a syndrome and normality, is not evidence that those syndromes are not valid categories. There are many interforms between the two biological sexes (Klinefelter's and Turner's syndromes, the adrenogenital syndrome, etc.). There are even interforms—mosaics and partial trisomies—between trisomy 21 and normal chromosomal architecture. But in neither case do these interforms threaten the validity of the categories concerned, because they are uncommon compared with the defined conditions. Statistical techniques such as discriminant function analysis for testing whether related syndromes are indeed separated by a zone of rarity have existed for more than 50 years (37), and these means have been used at least once to demonstrate that schizophrenia is distinguishable from other syndromes (38). Various forms of cluster analysis have also been used to demonstrate that, for example, distinct clusters of patients corresponding to the clinical concepts of mania, depression, and acute schizophrenia stand out clearly and consistently against the undifferentiated background of heterogenous symptoms (39). Other more elaborate statistical techniques have been developed more recently, and their potential has been demonstrated in analyses of clinical data sets of various kinds. A means of identifying natural clinical groupings by a combination of discriminant function analysis and admixture analysis was described by Sigvardsson et al. (40) and was used to demonstrate two distinct patterns of somatization in Swedish men. Meehl (41) developed the MAXCOV-HITMAX procedure, a taxonomic approach designed to detect and separate latent "loose syndromes" such as "schizotaxia" by using the observed covariances between multiple indicator variables. Woodbury and colleagues (42) developed a "grade of membership" model for assigning individuals to diagnostic categories; the model explicitly recognizes that natural classes have fuzzy boundaries and therefore allows individuals to be partly assigned to more than one class. Kendell (36) described a number of clinical research strategies, all based on a population deliberately chosen to represent a broader grouping than a single diagnostic category, which could be used to test or validate existing classifications.

The central problem, therefore, is not that it has now been demonstrated that there are no natural boundaries (zones of rarity) between existing diagnostic categories, or even that there are no appropriate statistical techniques, data sets, or clinical research strategies for determining whether natural boundaries exist within the main territories of mental disorder. The problem is that the requisite research has, for the most part, not yet been done. The resulting uncertainty makes it all the more important to clarify what is implied when a diagnostic category is described as having high validity, or simply as being valid.

The Concept of Validity Applied to Diagnoses

If the variation in psychiatric symptoms is indeed continuous and does not coalesce into fairly well-defined clusters, and if most of our familiar diagnostic categories are nothing but arbitrary loci in a multidimensional space, it is difficult to see how these categories can legitimately be regarded as valid, however useful they may be to practicing clinicians. On the other hand, a few diagnostic categories in psychiatry are almost universally accepted as valid. Most of these categories designate causes of mental retardation or dementia, such as Down's syndrome, fragile X syndrome, phenylketonuria, Huntington's disease, and Jacob-Creutzfeldt disease. We suggest, therefore, that a diagnostic category should be described as valid only if one of two conditions has been met. If the *defining characteristic* of the category is a syndrome, this syndrome must be demonstrated to be an entity, separated from neighboring syndromes and normality by a zone of rarity. Alternatively, if the category's defining characteristics are more fundamental—that is, if the category is defined by a physiological, anatomical, histological, chromosomal, or molecular abnormality—clear, qualitative differences must exist between these defining characteristics and those of other conditions with a similar syndrome (25). This distinction would imply that Down's syndrome and the other disorders listed earlier would all be valid, not because their etiology is known, but because, at the conceptual level at which they are defined, they are clearly different from other superficially similar conditions. The defining characteristic of Down's syndrome is the presence of an additional chromosome 21, and, although some individuals possess only part of that additional chromosome or possess it only in a proportion of their cells (mosaicism), these interforms are comparatively infrequent. The defining characteristic of Huntington's disease is an abnormal gene (the gene for huntingtin), which can clearly be identified as present or absent, at the tip of the short arm of chromosome 4. The defining characteristic of Creutzfeldt-Jacob disease is a characteristic histology (spongiform encephalopathy), which neuropathologists can reliably distinguish from other cerebral pathologies.

There are several reasons why the crucial issue in determining validity is not understanding of etiology but rather the existence of clear boundaries or qualitative differences at the level of the defining characteristic. First, understanding of etiology is not an all-or-none issue. It often emerges in stages as a complex network of interacting events is elucidated. Second, a clear boundary may be apparent, or be demonstrated, long before the underlying etiology is known. The histology of Creutzfeldt-Jacob disease and the other spongiform encephalopathies was recognized to be different from that of other brain diseases before abnormal prions were even conceived. The atheromatous plaques in the walls of coronary arteries, which are the defining characteristic of ischemic heart disease, were recognized as a distinct pathology, albeit one that often long preceded the onset of symptoms, several decades before their etiology was at all well understood. It is now apparent that the etiology of plaque formation is extremely complex, both genetically and environmentally.

Third, most psychiatric disorders (and some neurological disorders, such as torticollis, dystonia deformans, and migraine) are still defined by their clinical syndromes because their etiology is still largely unknown. It would be perverse to define validity such that no syndrome of unknown etiology could be accepted as a valid category, even if discriminant function analysis demonstrated that such a syndrome is separated by a zone of rarity from neighboring syndromes. Such evidence of a natural boundary would strongly suggest—but not prove—that the etiology of that syndrome was different from that of its neighbors, and such evidence would act as a powerful stimulus for research to elucidate the syndrome's etiology. Although it was never done, research demonstrating a zone of rarity between the facial and bodily features of children with Down's syndrome and those of other mentally retarded children would probably have been possible long before the discovery in 1959 that the former had an additional chromosome. Certainly, clinicians rarely had any difficulty deciding whether individual children were suffering from Down's syndrome (43, 44).

Finally, if many of our existing syndromal concepts do not reflect genuine discontinuities in the variation in symptoms, they are unlikely to survive successful exploration of their biological substrate. It would surely be folly to give the accolade of valid—meaning "well-founded… sound…against which no objection can fairly be brought" (4)—to categories that, as Frances and Egger have warned (35), may well be discarded within a decade or two. The syndromes of Down and Huntington have survived the identification of their biological substrates because they were based on genuine discontinuities in symptoms and signs, but others—such as dropsy, chlorosis, and Banti's syndrome—did not, and many of our present syndrome-based categories must face an uncertain future. Indeed, if no detectable discontinuities in symptoms are found in large tracts of the territory of psychiatric disorder, it is likely that, sooner or later, our existing typology will be abandoned and replaced by a dimensional classification. If that happens—and it may be about to happen for personality disorder—all existing categories will disappear and will do so with the implication that they have been discarded because they were not valid. In their place will be a set of dimensions, and important questions will then need to be asked about the number and identity of these dimensions and perhaps about their validity as well.

We concede that the criterion of a zone of rarity that we are advocating is not used—or at least not widely and deliberately used—in the classification of other medical disorders. This is because in other branches of contemporary

medicine nearly all diseases are now defined at a more fundamental level than their syndrome and are distinguished from one another by fairly well-established differences in pathology or etiology. As a result, issues of diagnostic validity rarely arise, and are certainly rarely discussed, even though many diseases share very similar syndromes. Many infectious fevers, for example, are characterized by malaise, pyrexia, sweating, headache, gastrointestinal disturbances, and a skin rash, but their causal organisms, which are different and readily distinguishable, are crucial defining characteristics. Pulmonary tuberculosis and bronchial carcinoma may present in almost identical ways, but no one doubts the validity of the distinction between them because of the key role of the tubercle bacillus in the former and because their histopathologies are different. Similarly, the clinical syndromes associated with chronic bronchitis, emphysema, and asthma have many common features (productive cough, wheeze, dyspnea, etc.), but the defining characteristics of these three conditions are qualitatively distinct—excessive mucus secretion in the bronchial mucosa in chronic bronchitis, enlarged air spaces distal to the terminal bronchioles in emphysema, and variable, widespread narrowing of peripheral airways in asthma (45). Indeed, these qualitative differences in their defining characteristics make it possible, and common, for two or three of these respiratory conditions to be diagnosed simultaneously (in sharp contrast to psychiatric disorders, where issues of comorbidity can only be decided by the adoption of arbitrary conventions). In the few psychiatric disorders, such as Alzheimer's disease, that are already defined by their pathology rather than their syndrome, there is already a clear tendency for the pathologically defined disorder to be subdivided into a series of genetically defined variants (46), as in other branches of medicine.

Psychiatry is in the position—that most of medicine was in 200 years ago—of still having to define most of its disorders by their syndromes. Because of the consequent need to distinguish one disorder from another by differences between syndromes, the validity of diagnostic concepts remains an important issue in psychiatry. In this situation, to search for boundaries between syndromes and to use zones of rarity as criteria of validity is, we contend, the best strategy available to us, as it was for bacteria when Sneath (12) was discussing their classification more than 40 years ago. It ensures that the lines of division are drawn, as Sneath (12) wrote, "where the resultant groups will hold the greatest content of information," and it creates an approximation to an Adansonian classification.

Although the problems involved in distinguishing one condition from another and in resolving issues of comorbidity are quite different, and more intractable, for psychiatric disorders than for most other medical disorders, the related problem of distinguishing between disorder and normality—the problem of where to draw the boundary—is often similar for both. As with major depression (16) and generalized anxiety, so with hypertension, diabetes, and irritable bowel syndrome. The variation between extensive, handicapping symptoms or pathology and an almost total absence of symptoms or pathology appears to be continuous in each case with no demonstrable zone of rarity. As a result, the boundary between normality and disorder has to be decided arbitrarily on pragmatic grounds.

The Utility of Diagnoses

The consequence of defining diagnostic validity in the way we are proposing is, of course, that most contemporary psychiatric disorders, even those such as schizophrenia that have a pedigree stretching back to the 19th century, cannot yet be described as valid disease categories. This does not mean, though, that they are not valuable concepts. In our view, it is crucial to maintain a clear distinction between validity and utility, and at present these two terms are often used as if they were synonyms. Indeed, Spitzer (47), the principal architect of DSM-III and DSM-III-R, refers to "clinical utility (validity)" and states that "a diagnostic concept is assumed to have validity to the extent that the defining features of the disorder provide useful information not contained in the definition of the disorder. This…may be about etiology, risk factors, usual course of the illness, whether it is more common among family members, and…whether it helps in decisions about management and treatment."

We propose that a diagnostic rubric may be said to possess utility if it provides nontrivial information about prognosis and likely treatment outcomes, and/or testable propositions about biological and social correlates. (Utility as defined here is therefore not the same as the idea of subjective utility in decision theory, but it is close to Spitzer's definition of validity.) To the best of our knowledge, the term "utility" was first used in this sense by Meehl (48) who wrote that "the fundamental argument for the utility of formal diagnosis…amounts to the same kind of thing one would say in defending formal diagnosis in organic medicine. One holds that there is a sufficient amount of etiological and prognostic homogeneity among patients belonging to a given diagnostic group so that the assignment of a patient to this group has probability implications which it is clinically unsound to ignore" (p. 92).

Many, though not all, of the diagnostic concepts represented by the categories of disorder listed in contemporary nomenclatures such as DSM-IV and ICD-10 are extremely useful to practicing clinicians, and most clinicians would be hard put to cope without them. Diagnostic categories provide invaluable information about the likelihood of future recovery, relapse, deterioration, and social handicap; they guide decisions about treatment; and they provide a wealth of information about similar patients encountered in clinical populations or community surveys throughout the world—their frequency and demographic characteristics, their family backgrounds and premorbid

personalities, their symptom profiles and their evolution over time; the results of clinical trials of several alternative therapies; and research on the etiology of the syndrome. This is all useful and sometimes invaluable information, whether or not the category in question is valid. Its usefulness depends mainly on two things: 1) the quantity and quality of the information in the literature (which depends on how long the category has been recognized and provided with adequate diagnostic criteria and how much competent research the category has generated) and 2) whether the implications of that information—particularly about etiology, prognosis, and treatment—are substantially different from the implications of analogous information about other related syndromes.

There is another crucial difference between validity and utility. Validity, as we define it, is an invariate characteristic of a diagnostic category. There may be considerable uncertainty about the category's validity because the relevant empirical information is lacking, but in principle a category cannot be partly valid. Either it is or it is not valid, and its validity does not depend on the context. Utility, on the other hand, is a graded characteristic that is partly context specific. Schizophrenia may be an invaluable concept to practicing psychiatrists but of little use to criminal lawyers or to scientists exploring the genetic basis of psychosis. Bipolar disorder may be a very useful concept in an acute admission unit, where it is important to distinguish between psychotic states that are and are not induced by stimulant drugs and that do and do not require long-term medication. However, it may be less useful in a rehabilitation program where the crucial issues may be which residents take their medication regularly, which are likely to have psychotic relapses, and which ones upset the others. Borderline personality disorder is a useful concept to many psychotherapists but not to most biologically oriented psychiatrists.

The rival or alternative operational definitions of a single diagnostic concept also have different implications for validity and utility. The existence of several rival definitions of a syndrome, embracing overlapping populations of patients, should raise the suspicion that it is not a valid category because these rival definitions suggest that variation in symptoms is continuous and that no identified zone of rarity indicating the boundaries of the syndrome can be drawn. Alternatively, if in practice all the rival definitions identify almost identical populations of patients, there is no important difference between them and the category may well be valid. Only one definition can be valid, unless all the alternatives identify virtually the same population of individuals. On the other hand, several alternative definitions can all be useful and can have different utilities in different contexts. The DSM-IV definition of schizophrenia, for example, is particularly useful for predicting outcome, largely because some degree of chronicity is in-built. But a much broader definition, embracing a heterogeneous "schizophrenia spectrum," is more useful for defining a syndrome with high heritability (22).

Implications for Research

Ever since the publication of DSM-III in 1980 and of the research version of ICD-10 in 1993, researchers have been under varying degrees of pressure, both from grant-giving organizations and from journal editors, to define the subjects of their research by using the explicit definitions provided in one or the other of these nomenclatures. There are good reasons for this. Before the 1970s, psychiatric research had been severely hampered by the low reliability of diagnostic assignments and by the fact that key terms such as "schizophrenia" were used in different ways in different countries and even in different centers within a single country (49, 50). Explicit definitions provide, if not a guarantee, at least an indication of adequate reliability, and they make it clear what meaning is being ascribed to the diagnostic terms employed. It is also a fundamental requirement of all scientific research that the subject matter of the investigation should be described with sufficient accuracy and in sufficient detail to enable others to repeat the study if they wish to. However, three elements of this policy need to be distinguished: 1) insistence on the use of explicit definitions to define the subjects of the research, 2) insistence that those subjects should be defined by their syndromes, and 3) insistence on the use of particular definitions of those syndromes, usually the definitions provided in the "official" nomenclature.

The first element is fundamental and common to all scientific research. The second and third are not, and whether they are appropriate requirements will depend on circumstances. Syndromes are basic concepts for most clinicians. They think in syndromal terms, and much of their clinical knowledge is stored in this format. It is therefore appropriate for most epidemiological research, for most studies of clinical course, and above all for most clinical trials to be based on precisely defined syndromes, whether or not those syndromes have been shown to be valid. Research into etiology is quite different. If the syndrome has not been shown to be valid (that is, to have demonstrable boundaries), there may be excellent reasons for using quite different criteria, for example, the presence of a single key symptom; a minimum score on a rating scale; a cognitive, pharmacological, or neurophysiological abnormality; or some combination of these. The frustration of geneticists (33), psychologists (2, 51), and psychiatrists themselves (35) with contemporary syndromal classifications and their definitions is therefore understandable and frequently justified.

An editor's or a funding organization's insistence on the use of the "official" definition of a syndrome that has not been shown to be valid is rarely justified and usually suggests that political considerations are intruding on what should be a purely scientific decision. The widespread use

of a single definition has many advantages, but researchers must be free to use other definitions if they wish, if only because that is how the shortcomings of the standard definition are most likely to be overcome.

In conclusion, it is important to distinguish between the validity and the utility of all diagnostic concepts and of their formal definitions. Otherwise, the term "valid" will continue to mislead, implying some kind of scientific respectability but actually meaning little more than "useful." At present there is little evidence that most contemporary psychiatric diagnoses are valid, because they are still defined by syndromes that have not been demonstrated to have natural boundaries. This does not mean, though, that most psychiatric diagnoses are not useful concepts. In fact, many of them are invaluable. But, because utility often varies with the context, statements about utility must always be related to context, including who is using the diagnosis, in what circumstances, and for what purposes.

Received Dec. 27, 2001; revisions received April 24 and June 20, 2002; accepted June 26, 2002. From the Department of Psychiatry and Behavioural Science, University of Western Australia, Perth, Australia. Address reprint requests to Dr. Kendell, 3 West Castle Rd., Edinburgh EH10 5AT, U.K.; randakendell@yahoo.co.uk (e-mail).

Supported by a grant from the Raine Medical Research Foundation of the University of Western Australia (Dr. Kendell).

References

1. World Health Organization: The ICD-10 Classification of Mental and Behavioural Disorders Diagnostic Criteria for Research. Geneva, WHO, 1993
2. Carson RC: Dilemmas in the pathway of the DSM-IV. J Abnorm Psychol 1991; 100:302–307
3. Kendell RE: Five criteria for an improved taxonomy of mental disorders, in Defining Psychopathology in the 21st Century: DSM-V and Beyond. Edited by Helzer JE, Hudziak J. Washington, DC, American Psychiatric Publishing, 2002, pp 3–18
4. The Shorter Oxford English Dictionary, 3rd ed. Oxford, UK, Clarendon Press, 1978
5. Kerlinger FN: Foundations of Behavioural Research, 2nd ed. London, Holt, Rinehart & Winston, 1973
6. Galtung J: Theory and Methods of Social Research. Oslo, Universitetsforlaget, 1969
7. American Psychological Association: Standards for Educational and Psychological Tests and Manuals. Washington, DC, APA, 1966
8. Goodwin DW, Guze SB: Preface, in Psychiatric Diagnosis. New York, Oxford University Press, 1974, pp ix–xii
9. Robins E, Guze SB: Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. Am J Psychiatry 1970; 126:983–987
10. Kendler KS: The nosologic validity of paranoia (simple delusional disorder): a review. Arch Gen Psychiatry 1980; 37:699–706
11. Andreasen NC: The validation of psychiatric diagnosis: new models and approaches (editorial). Am J Psychiatry 1995; 152: 161–162
12. Sneath PHA: Some thoughts on bacterial classification. J Gen Microbiol 1957; 17:184–200
13. Kendell RE, Gourlay J: The clinical distinction between psychotic and neurotic depression. Br J Psychiatry 1970; 117:257–260
14. Kendell RE, Gourlay J: The clinical distinction between the affective psychoses and schizophrenia. Br J Psychiatry 1970; 117: 261–266
15. Kendell RE, Brockington IF: The identification of disease entities and the relationship between schizophrenic and affective psychoses. Br J Psychiatry 1980; 137:324–331
16. Kendler KS, Gardner CO Jr: Boundaries of major depression: an evaluation of DSM-IV criteria. Am J Psychiatry 1998; 155:172–177
17. Kendler KS: The impact of diagnostic hierarchies on prevalence estimates for psychiatric disorders. Compr Psychiatry 1988; 29:218–227
18. Regier DA, Kaelber CT, Rae DS, Farmer ME, Knauper B, Kessler RC, Norquist GS: Limitations of diagnostic and assessment instruments for mental disorders. Arch Gen Psychiatry 1998; 55: 109–115
19. Berrettini WH: Susceptibility loci for bipolar disorder: overlap with inherited vulnerability to schizophrenia. Biol Psychiatry 2000; 47:245–251
20. Swillen A, Vogels A, Devriendt K, Fryns JP: Chromosome 22q11 syndrome: update and review of the clinical features, cognitive-behavioral spectrum, and psychiatric complications. Am J Med Genet 2000; 97:128–135
21. Kendler KS: Major depression and generalised anxiety disorder: same genes, (partly) different environments—revisited. Br J Psychiatry Suppl 1996; 30:68–75
22. Kendler KS, Neale MC, Walsh D: Evaluating the spectrum concept of schizophrenia in the Roscommon Family Study. Am J Psychiatry 1995; 152:749–754
23. Brown GW, Harris TO, Eales MJ: Social factors and co-morbidity of depressive and anxiety disorders. Br J Psychiatry Suppl 1996; 30:50–57
24. Bulik CM, Prescott CA, Kendler KS: Features of childhood sexual abuse and the development of psychiatric and substance use disorders. Br J Psychiatry 2001; 179:444–449
25. Scadding JG: Diagnosis: the clinician and the computer. Lancet 1967; 2:877–882
26. Albert DA, Munson R, Resnik MD: Reasoning in Medicine. Baltimore, Johns Hopkins University Press, 1988
27. Kraepelin E: Die Erscheinungsformen des Irreseins (1920): English translation, in Themes and Variations in European Psychiatry. Edited by Hirsch SR, Shepherd M. Bristol, UK, John Wright & Sons, pp 7–30
28. Jaspers K: Allgemeine Psychopathologie: English Translation of the 7th (1946) Edition. Translated and edited by Hoenig J, Hamilton MW. Manchester, UK, Manchester University Press, 1963, pp 569–570
29. Judd LL, Paulus MP, Wells KB, Rapaport MH: Socioeconomic burden of subsyndromal depressive symptoms and major depression in a sample of the general population. Am J Psychiatry 1996; 153:1411–1417
30. Lewinsohn PM, Klein DN, Seeley JR: Bipolar disorders in a community sample of older adolescents: prevalence, phenomenology, comorbidity and course. J Am Acad Child Adolesc Psychiatry 1995; 34:454–463
31. Widiger TA, Clark LA: Towards DSM-V and the classification of psychopathology. Psychol Bull 2000; 126:946–963
32. Cloninger CR: A new conceptual paradigm from genetics and psychobiology for the science of mental health. Aust N Z J Psychiatry 1999; 33:174–186
33. Ginsburg BE, Werick TM, Escobar JI: Molecular genetics of psychopathologies: a search for simple answers to complex problems. Behav Genet 1996; 26:325–333
34. Sullivan PF, Kendler KS: Typology of common psychiatric syndromes: an empirical study. Br J Psychiatry 1998; 173:312–319
35. Frances AJ, Egger HL: Whither psychiatric diagnosis. Aust N Z J Psychiatry 1999; 33:161–165

36. Kendell RE: Clinical validity. Psychol Med 1989; 19:45–55
37. Rao CR: The utilisation of multiple measurements in problems of biological classification. J Royal Statistical Society (Series B) 1948; 10:159–193
38. Cloninger CR, Martin RL, Guze SB, Clayton PJ: Diagnosis and prognosis in schizophrenia. Arch Gen Psychiatry 1985; 42:15–25
39. Everitt BS, Gourlay AJ, Kendell RE: An attempt at validation of traditional psychiatric syndromes by cluster analysis. Br J Psychiatry 1971; 119:399–412
40. Sigvardsson S, Bohman M, von Knorring A-L, Cloninger CR: Symptom patterns and causes of somatization in men: differentiation of two discrete disorders. Genet Epidemiol 1986; 3: 153–169
41. Meehl PE: Towards an integrated theory of schizotaxia, schizotypy, and schizophrenia. J Personal Disord 1990; 4:1–99
42. Manton KG, Korten A, Woodbury MA, Jablensky A: Symptom profiles of psychiatric disorders based on graded disease classes: an illustration using data from the WHO International Pilot Study of Schizophrenia. Psychol Med 1994; 24:133–144
43. Tredgold RF, Soddy K: A Textbook of Mental Deficiency, 9th ed. London, Bailliere, Tindall & Cox, 1956
44. Hilliard LT, Kirman BH: Mental Deficiency. London, J & A Churchill, 1957
45. Scadding JG: Meaning of diagnostic terms in broncho-pulmonary disease. Br Med J 1963; 2:1425–1430
46. Tanzi RE, Bertram L: New frontiers in Alzheimer's disease genetics. Neuron 2001; 32:181–184
47. Spitzer RL: Values and assumptions in the development of DSM-III and DSM-III-R: an insider's perspective and a belated response to Sadler, Hulgus, and Agich's "On values in recent American psychiatric classification." J Nerv Ment Dis 2001; 189: 351–359
48. Meehl PE: Psychodiagnosis, in Selected Papers. Minneapolis, University of Minnesota Press, 1959
49. Cooper JE, Kendell RE, Gurland BJ, Sharpe L, Copeland JRM, Simon R: Psychiatric Diagnosis in New York and London: Maudsley Monograph 20. London, Oxford University Press, 1972
50. World Health Organization: The International Pilot Study of Schizophrenia, vol 1. Geneva, WHO, 1973
51. Follette WC, Houts AC: Models of scientific progress and the role of theory in taxonomy development: a case study of the DSM. J Consult Clin Psychol 1996; 64:1120–1132