Symptom profiles of psychiatric disorders based on graded disease classes: an illustration using data from the WHO International Pilot Study of Schizophrenia

K. G. MANTON, A. KORTEN, M. A. WOODBURY, M. ANKER[1] AND A. JABLENSKY

From Duke University Department of Computer Science, Durham, NC, USA; NH&MRC Social Psychiatry Research Unit, the Australian National University, Canberra, ACT, Australia; WHO, Geneva, Switzerland; and University Department of Psychiatry, Perth, WA, Australia

SYNOPSIS The Grade of Membership (GoM) model is a classification procedure which allows a person to be a member of more than one diagnostic class. It simultaneously quantifies the degrees of membership in classes while generating the discrete symptom profiles or 'pure types' describing classes. The model was applied to the symptomatology, history, and follow-up of 1065 cases in the WHO International Pilot Study of Schizophrenia. The model produced an eight diagnostic class or 'pure type' solution, of which five were related to the diagnostic concepts of schizophrenia and paranoid disorder, two types were affective disorders, and one asymptomatic type. A subtype of paranoid schizophreniform disorder found primarily in developing countries was identified. There was a strong association between pure types and the original clinical and computer generated (CATEGO) diagnoses. A GoM based psychiatric classification might more clearly identify core disease processes than conventional classification models by filtering the confounding effects of individual heterogeneity from pure type definitions.

INTRODUCTION

The validity of a psychiatric classification is judged by: (*a*) its ability to describe clinically elicited symptoms and patterns of disease course and outcome; (*b*) its correlation with criteria such as age- and sex-prevalence, heritability, and treatment response with pathophysiological or morphological findings and, ultimately with aetiology; (*c*) whether it meaningfully distinguishes between diagnostic entities; and (*d*) whether it leaves few cases in residual diagnostic categories.

In the nineteenth century, clinical psychiatry identified the syndrome as the building block of nosology (Griesinger, 1861). The search for replicable, stable syndromes became the principal method in the taxonomic mapping of disorders to symptoms (Weber & Scharfetter,

1984). The World Health Organization's International Classification of Diseases, ICD-10 (WHO, 1992), and the American Psychiatric Association's DSM-IV (Task force on DSM-IV, 1991), which is in its final revisions, represent the latest stages of this process.

The identification of psychiatric syndromes has relied on two principal approaches. The first uses clinical insight and judgement to identify 'prototypes' which represent the necessary and sufficient attributes of the syndromes. A prototypical case is rarely clinically encountered in its pure sense but it is a 'standard' against which patients can be evaluated and possibly grouped. The prototype can generate clinical hypotheses, e.g. operational diagnostic criteria to be tested empirically. The second uses objective, statistical procedures to develop classifications. The earliest application of a purely statistical approach to psychiatric nosology dates to the 1920s (Moore, 1930). Attempts to derive statistically a comprehensive classification of mental disorders

[1] Address for correspondence: Ms M. Anker, Epidemiological and Statistical Methodology, WHO, 1211 Geneva 27, Switzerland.

were first made in the 1950s (e.g. Wittenborn, 1951; Lorr *et al.* 1963). Two statistical approaches to psychiatric classification have been tried: (*a*) construction of classifications on the basis of patient symptomatology using cluster analytic techniques such as taxonometric methods (Meehl, 1962), factor analysis and latent class analysis (McCutcheon, 1987; Woodbury & Manton, 1989); and (*b*) classification, or allocation, of patients to pre-established classes using linear or logistic discrimination (Everitt, 1980; Hand, 1981).

The application of statistical techniques to psychiatric classification problems has fallen, to date, short of expectations. Possible reasons include the dependence of existing methods on scaling and distributional assumptions which can distort classifications and the description of the profiles defining the classes. Efforts to interpret clinically statistically generated classes have proved difficult. A more general problem, associated with both the prototype, and statistical approaches to case categorization with psychiatric data, assumes that classes in an 'ideal' taxonomy must be mutually exclusive and jointly exhaustive of the universe of phenomena (Hempel, 1959), i.e. a case can either be, or *not* be, a member, of a given diagnostic class, but cannot exhibit 'partial' membership.

We report on the use of a classification procedure, the Grade of Membership (GoM) model (Woodbury & Clive, 1974; Woodbury *et al.* 1978; Woodbury & Manton, 1982; Clive *et al.* 1983), which deals with the scaling, distribution and discrete state assumptions referred to above. GoM differs from other classification models in that cases need not be assigned to a single diagnostic class. Instead, the symptoms of a case can be described as a weighted linear combination of profiles of symptoms, each associated with a specific diagnostic class, i.e. the individual can simultaneously and partially express symptoms in two or more classes. Thus, the model can describe complex cases who express only parts of the symptom profiles associated with prototypical diagnostic classes. The procedure has proved effective in generating classifications where categories are identified by large and complex sets of attributes, such as in the diagnosis of physical disease (Woodbury & Manton, 1982), describing the functional status of the very old (Manton & Soldo, 1985), the

diagnosis of somatization disorder (Swartz *et al.* 1986; Blazer *et al.* 1988) and of depression (Davidson *et al.* 1988, 1989). We report here the application of GoM to the classification of schizophrenia, using data from the WHO International Pilot Study of Schizophrenia (WHO, 1973, 1979; Leff *et al.* 1992).

METHOD

The Grade of Membership (GoM) model generates classifications by assigning grades of membership in analytically defined disease categories to cases. Disease categories are known as 'pure types', which are intended to be similar to prototypical or 'textbook' cases. Because cases with symptoms corresponding exactly to a pure type are rare, often the profile of symptoms observed in a case is related to more than one pure type. This is handled by assigning scores, or Grades of Membership (g_{ik}), describing the degree to which an individual belongs to a pure type. GoM scores are non-negative and sum to one for each individual. Thus, a case may belong entirely to one pure type or partially to several. Pure types are described by symptom profiles whose coefficients are estimated for each symptom to indicate the probability (λ_{ij}) of a symptom being manifest by a case belonging entirely ($g_{ik} = 1 \cdot 0$) to one pure type.

GoM represents the manifest symptoms of cases as a weighted linear combination of pure type coefficients and GoM scores. Formally, suppose K pure types are needed to describe the symptomatology expressed in the population. We can describe the model by letting i be the index for cases, j the index for symptoms (which are either present or absent), and k the index for pure types. The probability that case i manifests symptom j is P_{ij}, where

$$P_{ij} = \sum_{k} g_{ik} \lambda_{kj}$$

and g_{ik} is the GoM score for the case i on the kth pure type, and λ_{kj} is the set of probabilities for pure type k and symptom j. Pure type coefficients (λ_{kj}) and GoM scores (g_{ik}s) are estimated simultaneously by maximum likelihood using as data the observed presence or absence of symptoms. The model can be extended to polychotomous signs or symptoms by adding an

index, l_j, to represent the number of exclusive, possible responses to the jth variable.

The GoM model has similarities to factor analysis. The pure type coefficients (λ_{kj}) of GoM are similar in function to 'loadings' in factor analysis. Grade of Membership scores are similar in function to factor scores. There are, however, clear differences between the two methods. The focus of GoM is on characterizing individual cases on a large number of symptoms, i.e. two individuals are similar if their symptomatology is similar. In factor analysis the focus is on identifying clusters of variables from pairwise correlations or covariances. The correlations, which are functions of second-order moments, do not retain information on individuals. In factor analysis, variables are assumed continuous and normally distributed. In GoM, variables are categorical and no specific distribution is assumed. Thus, GoM can describe non-linear relations and categorical data with significant higher-order moments. GoM scores are non-negative and sum to one for a case. Factor scores range from minus to plus infinity. The symptoms for a case can be decomposed, using g_{ik}, into membership in each pure type. In factor analysis, there is no analogous decomposition of the properties of individuals.

Furthermore, because no specific distribution is assumed, GoM is more robust to sample differences than factor analysis. In GoM, the heterogeneity of cases is represented by g_{ik} while leaving λ_{kj} largely unaffected. The limitation of factor analysis to second-order moments does not permit individual heterogeneity to be filtered from its loadings.

GoM parameters are estimated by maximum likelihood so there is an explicit probability model and decisions can be made about model structure using classical statistical inference. For example, a decision has to be made about K, i.e. the number of profiles necessary to reproduce individual symptom patterns. An initial value of K is selected from theoretical considerations, and the likelihood calculated. The likelihood values for solutions with $K - 1$ and $K + 1$ pure types are then calculated and compared to the model with K pure types to see if the changes in the likelihood are significant. This is repeated until the addition or deletion of a pure type produces a likelihood ratio χ^2 for the change which is not significant. We then accept the

model with that value of K as reproducing the data except for sampling error. GoM scores can be used to determine the prevalence of pure type profiles in the population. The total population burden of a symptom profile is represented by summing the g_{ik}s for a pure type and dividing by the sample size. This is a more general definition of prevalence, weighted by disease or symptom intensity, than using discrete classes where a case belongs to one homogeneous group.

A simple example illustrates these concepts. Suppose a study examined the cultural diversity of a community largely of French or German descent. Information available includes the nationalities of the four grandparents of each member of the community, categorized as 'French', 'German' or 'Other'. A classification could be developed which would conform to a GoM model with three pure types: 'French', 'German' or 'Other'. The degree to which an individual belongs to each of the pure types is the proportion of grandparents with the corresponding nationality. Thus, someone who had one French grandparent and three German grandparents would have a g_{ik} of 0·25 in the 'French' pure type and a g_{ik} of 0·75 in the 'German' pure type and zero on 'Other'; someone whose grandparents were all British would have g_{ik} of zero in the 'French' and 'German' pure types and 1·0 on the 'Other' type. The prevalence of 'French', 'German' and 'Other' backgrounds are calculated by adding the g_{ik}s across all individuals in each pure type and dividing by the total. This is a more general definition of prevalence than if population members were classified on current nationality. One way to examine the structure of GoM classes is by calculating the correlation over individuals between g_{ik}s for each profile. These correlations will tend to be negative since their total is fixed for each individual (i.e. the g_{ik}s sum to 1·0 for each person). Thus, a zero correlation implies a moderate relation of the g_{ik}s of any two given pure types. In the example, suppose that the community had a long history of inter-marriage between French and German communities with other nationalities being recent arrivals. This could mean that the correlation between 'French' and 'German' pure types would be higher than between 'Other' and either the 'French' or 'German' pure types.

The relation of pure types to variables *not*

used in their definition can be estimated by calculating λ_{kj}s for 'external' variables without allowing them to affect the likelihood function, i.e. the external variables are not used to form the pure type profiles. The values of λ_{kj}s for these variables represent their association with the symptom profiles.

The International Pilot Study of Schizophrenia

The IPSS is a transcultural investigation of 1202 cases in nine countries – China, Colombia, the former Czechoslovakia, Denmark, India, Nigeria, the former Union of Soviet Socialist Republics, the UK, and the US. Cases were selected among consecutive admissions on the basis of criteria excluding gross organic cerebral pathology, chronic medical conditions, sensory defects and mental retardation, as well as individuals below the age of 15 and above the age of 44. Initial clinical diagnoses were: schizophrenia (811 cases), affective psychosis (164 cases), and other psychoses and non-psychotic illnesses (277 cases). Each case had a comprehensive initial clinical and social evaluation, which was repeated after 2 and 5 years. Symptomatology was assessed by psychiatrists using the Present State Examination (PSE; WHO, 1973; Wing *et al.* 1974) translated into the local language and back-translated into English to ensure equivalence. Similar procedures were used to elicit medical and social history data. Disorders were diagnosed by three methods. First, a clinical diagnosis was made and coded to the ICD-8 by the project psychiatrist in each centre. Secondly, the PSE ratings on symptoms were input to a computer diagnostic program (CATEGO) to produce a deterministic reference classification based on a logical decision tree designed to implement the diagnostic concepts of schizophrenia and other disorders. Thirdly, clustering was used to group patients with shared characteristics. A final diagnosis was made for each case by the coordinating team in Geneva, using all diagnostic and follow-up data including three consecutive CATEGO evaluations.

This analysis explores the extent to which clinical symptoms and signs at initial examination can be used to define syndromal entities represented by pure type profiles; how the symptom patterns manifested by cases are

related to pure types; and how pure types relate to external variables such as field research centre (FRC), age, sex, final project diagnosis, social isolation prior to inclusion in the study, and pattern of disease course over 5 years.

The variables used to define pure types were: (*a*) psychopathological symptoms, elicited with the PSE, present in the month before the diagnostic interview. These were organized as units of analysis (UAs), coded as either present or absent (for a detailed description of the derivation and content of the UAs, see WHO, 1973); and (*b*) symptoms present at any time during the index episode of illness, elicited in a psychiatric history interview with a key informant (or, in a few instances, the patient). These symptoms were recorded as present or absent on a 57-item checklist in the Psychiatric History Schedule.

RESULTS

GoM was applied to 1065 cases from the IPSS with sufficient data. We first determined the number of profiles required to explain heterogeneity on 170 UA and symptom variables. Solutions for 6, 7, 8, and 9 pure types were obtained; on the basis of the likelihood ratio χ^2 values, an 8 pure type solution was chosen.

Description of the syndromal pure types

All symptoms used to define pure types are dichotomous. Thus, pure types may be described by the probability (λ_{kj}) of the presence of specific symptoms. Symptoms with $\lambda_{kj} \geqslant 0.5$ are considered characteristic of a pure type, and are given in Tables 1, 2 and 3. The membership for some persons is shared in that $\lambda_{kj} \geqslant 0.5$ for more than one pure type. For Types I–V, only unique features are listed. Types VI and VII consist of shared features. Type VIII was an asymptomatic pure type characterized by the absence of symptoms with $\lambda_{kj} \geqslant 0.5$. Few symptoms derived from the psychiatric history feature in the pure types, and unless otherwise indicated, symptoms refer to UAs from the initial PSE.

A clinical evaluation of the statistically derived profiles suggests that they can be grouped into four sets, as follows.

Table 1. *Symptom profiles of schizophrenic syndromes*

(Unique features only are listed, defined as symptoms for which lambda values *kj* are greater than 0·5, where λ_{kj} is the probability that a symptom be expressed, given the pure type. λ_{kj} values are shown in brackets.)

Pure Type I (Positive psychotic syndrome with predominance of auditory hallucinations)	Pure Type II (Positive psychotic syndrome with predominance of depersonalization/derealization)	Pure Type III (Syndrome with negative disturbances)
Verbal hallucinations (1·00)	Changed appearance (1·00)	Apathy (1·00)
Voices speak to patient (1·00)	Distorted time perception (1·00)	Restricted speech (1·00)
Frequent auditory hallucinations (1·00)	Looking at self (1·00)	Negativism (0·61)
Voices speak full sentences (1·00)	Break of self identity (1·00)	
Voices discussing patient (1·00)	Nihilistic delusions (0·77)	
Non-verbal auditory hallucinations (0·99)	Gustatory hallucinations (0·83)	
Voices comment on patient's thoughts (1·00)	Sexual hallucinations (0·45) †	
Auditory pseudohallucinations (0·48) †	Delusions of impending doom (0·44) †	
Hallucinations from body (0·42) †	*Depersonalization (1·00)	
	*Sexual delusions (1·00)	

* Psychiatric history items.
† Although *kj* less than 0·5, these items have very low values on all other pure types.

Table 2. *Symptom profiles of affective syndromes*

(Unique features only are listed, defined as symptoms for which lambda values (λ_{kj}) are greater than 0·5, where λ_{kj} is the probability that a symptom be expressed, given the pure type. λ_{kj} values are shown in brackets.)

Pure Type IV Severe depression	Pure Type V Manic or schizomanic syndrome
Depressed mood (1·00)	Observed elated mood (1·00)
Hopelessness (1·00)	Lability of affect (1·00)
Worse in morning (1·00)	Demonstrative (1·00)
Special depression (1·00)	Flight of ideas (1·00)
Diminished appetite (1·00)	Change of interests (0·88)
Decreased energy (1·00)	Pressure of speech (1·00)
Retardation (1·00)	Perseveration (0·93)
Decreased libido (0·75)	Suggestibility (0·89)
Indecision (0·85)	Ecstatic mood (0·70)
Worse in evening (0·58)	Klang associations (0·50)
*Suicidal tendencies (1·00)	*Sudden changes in direction of interests (1·00)
*Delusions of self-depreciation or guilt (1·00)	*Affective disorders – mood swings (1·00)
*Decreased interest (1·00)	

* Psychiatric history items.

A. Schizophrenic syndromes containing unique features

Three pure types correspond to diagnostic concepts in the ICD category of schizophrenia.

I. A positive psychotic syndrome characterized by a predominance of auditory hallucinations. Among symptoms unique to this syndrome (all of them representing different types of hallucinations) are at least two of Schneider's 'first rank' symptoms of schizophrenia (Schneider, 1959): voices discussing the subject; voices commenting on his/her thoughts. Three other first-rank symptoms (thought alienation, thoughts spoken aloud, and delusions of control) along with delusions of persecution, tactile, visual and olfactory hallucinations and flatness of affect are shared with other profiles.

II. A positive psychotic syndrome characterized by depersonalization/derealization. Several symptoms unique to this profile (changed appearance, experience of looking at one's self from outside, experience of a break in the identity of the self, distorted time perception)

Table 3. *Symptom profiles of non-specific paranoid psychosis*

(Profiles contain shared symptoms only. Symptoms are listed for which kj are greater than 0·5, where λ_{kj} is the probability that a symptom be expressed, given the pure type.)

Pure Type VI Paranoid with aggression	Pure Type VII Paranoid
*Verbal aggression (1·00)	Suspiciousness (1·00)
*Physical aggression (1·00)	Delusions of persecution (1·00)
*Overactivity and agitation (1·00)	Unwilling to cooperate (1·00)
*Bizarre behaviour (1·00)	Delusions of reference (1·00)
Irritability (0·76)	Delusional mood (1·00)
*Disorders of pitch and modulation (1·00)	Inadequate description (1·00)
Delusional system (0·67)	Poor rapport (1·00)
Early waking (0·91)	Ideas of reference (0·77)
Worse in evening (0·59)	
Suspiciousness (1·00)	
Visual hallucinations (0·59)	
Sleep problems (0·68)	
Delusions of persecution (1·00)	
Ideas of reference (0·78)	

* Psychiatric history items.

are expressions of psychotic depersonalization or derealization. There are over 40 shared features, covering many positive psychotic symptoms including three first-rank phenomena.

III. A negative or 'clinical poverty syndrome' (Wing & Brown, 1970) uniquely characterized by 'negative' symptoms, such as apathy and restricted speech. Shared characteristics include multiple attributes of chronic and residual schizophrenia, e.g. autism, flatness of affect, self-neglect, speech dissociation, and catatonic symptoms such as stereotypies, ambivalence, and decreased interests.

B. Affective syndromes

Two pure types relate to affective disorders.

IV. Depressive syndrome: this profile includes a number of manifestations of major depressive disorder in both unique and shared features. Among its unique features are depressed mood, hopelessness, decreased energy and libido, retardation, as well as the PSE item 'special' depression (presence of a special quality to the affective disturbance, e.g. 'like a heavy black cloud over my head'), and diurnal fluctuations of severity. Shared features are delusions of guilt, suicidal thoughts, obsessive thoughts, and retardation.

V. Manic or schizomanic syndrome: while unique features of this profile provide a typical picture of manic disorder, there is a large number of attributes shared with other profiles including

delusions, hallucinations, autism, ambivalence, and several other features encountered in schizophrenic disorders. This suggests that this pure type encompasses symptoms currently classified as schizo-affective and, in particular, schizo-manic.

C. Paranoid psychosis

The two profiles VI and VII are similar. Both are characterized solely by symptoms shared with other pure types. Both profiles describe some of the common presenting symptoms of schizophrenic illnesses, such as delusions of reference and persecution, suspiciousness, and uncooperativeness (see Table 3). In contrast to Type VII, Type VI is described by several history-derived items related to bizarre behaviour, excitement, and aggression.

D. Asymptomatic type

There is only one pure type in this group (Type VIII). In contrast to all other pure types, this is characterized by absence of symptoms. As an asymptomatic type, when mixed with another type, the symptom intensity of the other syndrome is decreased.

Frequency of different pure types and correlations between pure types

Pure type profiles (λ_{kj}) and g_{ik}s for individuals are derived simultaneously. The g_{ik}s indicate the degree to which the individual's symptoms reflect

Table 4. *Percentage of disease burden associated with each pure type*

Type		%
I	Positive psychotic, auditory hallucinations	11·1
II	Positive psychotic, depersonalization/derealization	6·0
III	Negative schizophrenic syndrome	10·7
IV	Severe depression	14·2
V	Manic/schizomanic syndrome	5·6
VI	Paranoid psychosis with aggression	9·4
VII	Paranoid psychosis	14·8
VIII	Asymptomatic syndrome	28·2

Table 5. *Correlation among the GoM scores* (g_{1k}) *for the eight pure types*

		I	II	III	IV	V	VI	VII	VIII
I	Auditory	1·000							
II	Depersonalization/derealization	0·19	1·000						
III	Negative	−0·065	−0·111	1·000					
IV	Depression	−0·078	−0·038	−0·202	1·000				
V	Manic	−0·086	0·111	−0·065	−0·303	1·000			
VI	Paranoid/aggression	−0·002	−0·144	0·004	−0·265	−0·064	1·000		
VII	Paranoid	−0·32	0·092	−0·14	−0·235	0·007	0·025	1·000	
VIII	Asymptomatic	−0·425	−0·379	−0·276	−0·005	−0·199	−0·294	−0·325	1·000

Table 6. λ_{kj} *relating pure types to Field Research Centres*

	I	II	III	IV	V	VI	VII	VIII
	Auditory	Depersonalization/ Derealization	Negative	Depression	Manic	Paranoid/ Aggression	Paranoid	Asymptomatic
Aarhus	0·062	0·296	0·000	0·036	0·002	0·000	0·189	0·195
Agra	0·000	0·000	0·302	0·063	0·000	0·604	0·000	0·000
Cali	0·184	0·185	0·000	0·000	0·000	0·080	0·000	0·244
Ibadan	0·412	0·000	0·331	0·102	0·000	0·316	0·087	0·000
London	0·247	0·360	0·000	0·100	0·000	0·000	0·024	0·104
Moscow	0·016	0·000	0·000	0·342	0·111	0·000	0·104	0·295
Washington	0·000	0·159	0·000	0·000	0·268	0·000	0·459	0·162
Prague	0·079	0·000	0·367	0·357	0·619	0·000	0·137	0·000

the profile of each pure type. These can be used to examine the frequency distribution of symptoms over individuals. Since IPSS cases are non-random clinical samples, the frequency data are not estimates of population prevalences, but the attribution of pure type associated symptom burden to cases. The relative percentages of the eight pure types are shown in Table 4. This distribution, which indicates the degree to which patients express the different pure types of psychopathology, can be complemented by descriptions of the extent to which the profiles co-occur in individual cases. This is represented by the correlation of g_{ik}s as shown in Table 5.

Pure Types I and II are strongly correlated, i.e. they tend to occur in the same person, although phenomenologically they represent different and separate syndromal entities. On the other hand, both Type I and II are least likely to occur (have the largest negative correlations) with Type VIII ('asymptomatic'). This suggests that these two profiles are the most intensively expressed syndromes. The depressive syndrome (Type IV) is negatively correlated with the manic/schizomanic syndrome (Type V) and also with Types VI, VII, and III (paranoid psychosis and schizophrenia with 'negative' disturbances). Most individuals had non-zero g_{ik}s in several pure types. Only 12 individuals had a g_{ik} greater than 0·9. The average non-zero g_{ik} ranged from 0·10 for Type II (positive psychotic schizophrenic) to 0·25 for Type V

(manic), 0·28 for Type I (positive psychotic schizophrenic) and 0·38 for Type VIII (asymptomatic).

Relation of other variables to pure types

The validity of the pure types can be tested by examining their relation to 'external' variables not used to define profiles. For the IPSS, external variables were geographical and cultural location (FRC), age, sex, level of social functioning, and pattern of course over 5 years. The λ_{kj}s for external variables are interpreted similarly to the other λ_{kj}, i.e. they represent, for a given pure type, the probability of exhibiting the characteristic.

Geographical and cultural location

The distribution of pure types over FRC is in Table 6. Each column is associated with a pure type, and gives the λ_{kj}s for FRCs. The columns total to unity. Thus, in the IPSS patient sample, the probability that a case totally associated with Pure Type I comes from Ibadan is 41·2%. There is a zero probability that such a case comes from Washington. The first pure type thus describes a profile in Cali, Ibadan and London, rather than in Agra or Washington. This is not to say, of course, that cases from Washington or Agra cannot display partial membership in this pure type. Looking at the other pure types, we see that the depressive syndrome is characteristic of patients found primarily in Moscow and Prague, while the manic syndrome is characteristic in Washington and Prague patients. Pure Type VI, paranoid with aggression, is found only in developing countries.

Age, sex and social isolation

While the distribution of the pure types by age (above and below age 30) did not present any clinically important associations (except for Type II, Type III, and Type V being over-represented in below age 30), the distribution by gender suggests significant differences (see Table 7).

Types I and III are almost equally distributed between males and females. However, Type II is expressed predominantly in young males. Type IV and Type V occur exclusively in females, especially the young. As pure types are ideal-

izations of clinical reality, this pattern can be seen as a confirmation of previously established associations of affective syndromes to gender.

Social isolation characterized the social functioning of IPSS cases. Since social isolation (both pre-morbid and prior to initial examination) was among the best predictors of subsequent course and outcome (WHO, 1979), its relation to pure types is of interest. Table 7 indicates that the agitated, aggressive paranoid syndrome (Type VI), characteristic of the three FRCs in developing countries, is *not* associated with social isolation. A low association with social isolation is observed for Types VI (paranoid/aggression), V (manic/schizomanic), VIII (neurotic and personality disorders), and I (florid schizophrenia with auditory hallucinations). In contrast, Type III ('negative' syndrome) and Type II (positive psychotic syndrome) are associated with social isolation as either a personality trait, or illness-induced social withdrawal, at the initial examination.

Disease course and outcome over five years

Two outcome variables are the most important external criteria. These are: (*a*) the total number of discrete illness episodes (both psychotic and non-psychotic, and of any clinical type, see WHO, 1979) over five years of follow-up; and (*b*) the proportion of follow-up time occupied by such episodes. Results are presented in Table 8. The manic/schizomanic syndrome (Type V) is associated with the least time spent in illness episodes; the probability of a Type V case spending more than 40% of the time in psychotic episodes is 0·1%, i.e. the majority of manic/schizomanic episodes are brief. In contrast, although the 'negative' schizophrenic syndrome (Type III) is associated with a 64% probability of a single episode, there is a 32% probability that the episode lasted most of the follow-up. Pure Type II is associated with multiple episodes which last, in total, at least half of the follow-up.

Diagnostic classification used at initial evaluation

A final test of the validity of pure types is their relation to the diagnostic classification of cases carried out at the FRCs (clinician's diagnosis) and the coordinating centre (CATEGO computer diagnosis). The two approaches are different in that assignments to ICD-8 categories

Table 7. λ_{kj} *relating pure types to age, sex, and social isolation*

	I	II	III	IV	V	VI	VII	VIII
		Depersonalization/				Paranoid/		
	Auditory	Derealization	Negative	Depression	Manic	Aggression	Paranoid	Asymptomatic
Age (in years)								
< 30	0·421	0·978	0·895	0·412	0·728	0·592	0·363	0·469
> 30	0·589	0·022	0·105	0·588	0·272	0·408	0·637	0·531
Sex								
Male	0·505	0·973	0·524	0·000	0·000	0·873	0·400	0·579
Female	0·495	0·027	0·476	1·000	1·000	0·127	0·600	0·421
Social isolation								
None	0·557	0·000	0·178	0·306	0·622	1·000	0·295	0·496
Little	0·114	0·221	0·000	0·248	0·152	0·000	0·253	0·234
Some	0·111	0·156	0·218	0·197	0·226	0·000	0·225	0·082
Moderate	0·132	0·462	0·299	0·204	0·000	0·000	0·177	0·157
Severe	0·086	0·161	0·305	0·045	0·000	0·000	0·050	0·031

Table 8. λ_{kj} *relating pure types to total number of and percentage of time spent in psychotic episodes*

	I	II	III	IV	V	VI	VII	VIII
		Depersonalization/				Paranoid/		
	Auditory	Derealization	Negative	Depression	Manic	Aggression	Paranoid	Asymptomatic
Total number of psychotic episodes in 5 years								
1	0·432	0·211	0·640	0·495	0·029	0·571	0·431	0·684
2	0·309	0·558	0·229	0·276	0·232	0·258	0·196	0·105
3	0·098	0·071	0·116	0·158	0·195	0·171	0·140	0·103
4	0·000	0·160	0·015	0·039	0·318	0·000	0·060	0·041
5	0·161	0·000	0·000	0·032	0·226	0·000	0·173	0·067
Time (%) spent in psychotic episodes in 5 years								
< 10	0·401	0·232	0·403	0·661	0·217	0·781	0·454	0·820
10–40	0·143	0·111	0·106	0·278	0·782	0·162	0·205	0·074
40–90	0·213	0·179	0·173	0·061	0·001	0·057	0·127	0·016
90–100	0·243	0·478	0·318	0·000	0·000	0·000	0·215	0·090

were made by the psychiatrist using all available information. CATEGO classifications are based on PSE symptoms only. Both approaches have acceptable predictive validity (Wing, 1980).

The relation of the pure types to these two schemes is shown in Table 9. There is a high degree of concordance between the GoM typology and the ICD-8 clinical diagnostic classification. The association of ICD-8 category 295 (schizophrenia) and Type II or Type III is 100%; 93·4% for Type I, 92·3% for Type VI, and 81% for Type VII. The remainder of Type I or Type VII cases are almost exclusively associated with ICD-8 paranoia (297.0). Type IV (depression) is associated with clinical diagnoses of either manic-depressive psychosis (296.0 and 296.2) or depressive neurosis (300.4). Type V (manic/schizomanic) is associated not only with clinically diagnosed manic-depressive psychosis, but also has a 23·3% probability of

falling into the ICD-8 category of schizophrenia (295.0).

The concordance of GoM profiles and CATEGO classes is striking. All of Type I and II symptomatology falls within CATEGO class S, which contains 'central' schizophrenic conditions (WHO, 1979). This class is closely associated with Type VI (66·8%) but only marginally with Type III (2·6%). Type III is, however, closely associated (85·6%) with the CATEGO class O which is characterized by psychotic conditions with catatonic symptoms and hallucinatory behaviour. Type VII is closely associated with CATEGO class P (paranoid psychoses). Similarly, there is a close association of CATEGO classes D (depressive psychoses) and M (manic and mixed affective psychoses) with Pure Types IV and V respectively.

The association of asymptomatic GoM Type VIII with both ICD-8 and CATEGO is less clear.

Table 9. λ_{kj} *relating pure types to nosologic schemes*

	I	II	III	IV	V	VI	VII	VIII
		Depersonalization/				Paranoid/		
	Auditory	Derealization	Negative	Depression	Manic	Aggression	Paranoid	Asymptomatic
ICD-8 diagnoses								
Schizophrenia	0·934	1·0	1·0	0·0	0·233	0·923	0·810	0·395
Manic depr./depression	0·0	0·0	0·0	0·689	0·0	0·0	0·005	0·0
Mania	0·0	0·0	0·0	0·0	0·732	0·077	0·0	0·026
Paranoid, reactive	0·066	0·0	0·0	0·0	0·0	0·0	0·177	0·039
Neurotic depression	0·0	0·0	0·0	0·241	0·0	0·0	0·008	0·165
Other	0·0	0·0	0·0	0·070	0·035	0·0	0·0	0·375
CATEGO class								
O (Other psychotic)	0·0	0·0	0·856	0·0	0·033	0·0	0·025	0·0
S (Schizophrenic)	1·0	1·0	0·026	0·0	0·0	0·668	0·0	0·0
D (Depression)	0·0	0·0	0·128	1·0	0·0	0·0	0·020	0·0
M (Mania)	0·0	0·0	0·0	0·0	0·967	0·185	0·052	0·188
P (Paranoid)	0·0	0·0	0·0	0·0	0·0	0·147	0·874	0·096
B (Personality disorder)	0·0	0·0	0·0	0·0	0·0	0·0	0·005	0·117
N (Neurotic)	0·0	0·0	0·0	0·0	0·0	0·0	0·024	0·599

This is not surprising, given its residual nature. Its association with an ICD-8 diagnosis of schizophrenia (39·5 %) possibly arises because of cases diagnosed using history variables not used in the GoM analysis, or because of treated cases who were asymptomatic at assessment. Associations are high for the appropriate rubrics of neurotic and personality disorders (ICD-8, 300 and 301; CATEGO B and N).

DISCUSSION

We established the applicability of GoM, as a statistical tool, to a multi-cultural psychiatric database, indicated its potential uses, and demonstrated its validity in mental health research. GoM differs from other classification methods in that: (a) a person may be a member of more than one class with membership in different classes quantified; (b) there is an explicit probability model on which objective statistical decisions can be made about the optimal solution; (c) the χ^2 used to assess the fit of GoM classes to the data are not dependent upon assuming the data are distributed in a particular way; (d) GoM can be used to determine the prevalence of specific symptom profiles in a population more flexibly than a 'crisp' classification since the latter assumes that each case in a class is exactly described by the class characteristics; (e) λ_{kj}s can be estimated for variables not used to define profiles.

GoM was applied to 1065 cases assessed in eight FRCs participating in the WHO Inter-

national Pilot Study of Schizophrenia (IPSS), followed for 5 years. These cases, and the psychiatrists evaluating them, belong to different cultural settings, a fact which may cause confounding, in spite of standardized assessments and efforts in training and reliability monitoring.

In decomposing the mental state (Present State Examination – PSE) and psychiatric history data and in constructing pure types of clinical syndromes (symptom profiles), the GoM model identified eight pure types. The eight pure types or syndromes are listed together with coefficients indicating the degree of association of each symptom or sign with the profile. The g_{tk} represent the degree of membership in each group.

Of the eight pure types, five (three specific, i.e. having unique attributes, and two non-specific) have clear relations to the diagnostic concepts of schizophrenia and paranoid psychoses. Two are affective in nature, one depressive and one manic or schizomanic. The remaining type is asymptomatic.

The eight pure types showed differences in their distribution over FRCs, age, association with degree of social isolation prior to the initial examination, number of illness episodes over 5 years and proportion of follow-up time spent in such episodes. The findings are in the direction predicted from prior analyses of IPSS data (WHO, 1973, 1979) using different methods.

The consistency of the diagnostic classification of the IPSS initial examination material which has been achieved, on one hand, with traditional

clinical judgement (ICD-8) and its computer simulation (CATEGO), and on the other, with a statistical method (GoM), is only partly due to all three methods employing PSE data, since the ways those data are used are very different. The emergence of a concordant classification is evidence for the existence of replicable symptom patterns in psychiatric patients from different cultures.

Assuming, given these findings, the clinical validity of the GoM profiles, we can consider how the analysis refines prior conclusions from the IPSS. First, the findings suggest that the schizophrenic syndrome can be subdivided into three types: one characterized predominantly by auditory hallucinations; one by psychotic depersonalization/derealization and multiple positive psychotic features; and one by negative disturbances. They are identifiable by symptoms not appearing in other pure types. They occur with different frequencies in FRCs and have different prognoses.

Secondly, a paranoid schizophreniform disorder was identified in FRCs in developing countries. This disorder, although exhibiting features in common with paranoid and schizophrenic illnesses occurring in other FRCs, is also characterized by behavioural manifestations elicited from the psychiatric history data, i.e. hyperactivity and agitation, verbal and physical aggression, and bizarre behaviour. Its duration is brief and prognosis good. In the light of its mixed clinical presentation and benign course, the relation of the syndrome to schizophrenia requires further investigation. The identification of this constellation of symptoms and history items suggests hypotheses to be tested by exploring the association of this syndrome with psychiatric and personal history, course and outcome, and sociocultural data.

Thirdly, GoM types corresponding to manic disorders suggests no clear-cut division can be drawn between 'pure' manic symptomatology and a wide range of symptoms usually considered part of the manifestation of schizophrenia. Thus, a schizomanic syndrome, closely related to the manic syndrome or merging into it, is suggested. In contrast, the GoM derived syndrome of depression did not contain schizophrenic features.

Fourthly, and most importantly, the GoM model provides a new way to resolve the apparent contradiction of a categorical classification of psychiatric syndromes, and the notion of a continuum of psychopathological symptoms where no clear boundaries can be drawn and 'entities' are artificially designated. The continuum notion may owe its existence to the difficulty in disentangling two factors: the existence of descriptive 'types' and the variable degree of their expression in individual cases. A promising use of the model is to identify and represent individual differences in disease manifestation. Thus, it can be used to evaluate hypotheses about psychiatric disorders and classifications that may allow us to identify core disease processes better than conventional discrete classification models.

This paper is based on the data and experience obtained during the WHO coordinated International Pilot Study of Schizophrenia (IPSS), a project sponsored by the World Health Organization, the National Institute of Mental Health (USA), and the participating field research centres. A full list of investigators is presented in *Schizophrenia, An International Follow-Up Study, WHO, Geneva* (John Wiley & Sons: New York, 1979, reprinted 1981).

REFERENCES

American Psychiatric Association, Task Force on DSM-IV. (1990). *DSM-IV*. Work in Progress. American Psychiatric Association: Washington, DC.

Blazer, D., Schwartz, M., Woodbury, M. A., Manton, K. G., Hughes, D. & George, L. K. (1988). Depressive symptoms and depressive diagnoses in a community population: use of a new procedure for analysis of psychiatric classification. *Archives of General Psychiatry* **45**, 1078–1084.

Clive, J., Woodbury, M. A. & Siegler, I. C. (1983). Fuzzy and crisp set-theoretical-based classification of health and disease. *Journal of Medical Systems* **7**, 317–322.

Davidson, J. R. T., Woodbury, M. A., Pelton, S. & Krishnan, K. R. R. (1988). A study of depressive typologies using Grade of Membership analysis. *Psychological Medicine* **18**, 179–189.

Davidson, J. R. T., Woodbury, M. A., Zisook, S. & Giller, E. L. (1989). Classification of depression by Grade of Membership: a confirmation study. *Psychological Medicine* **19**, 987–998.

Everitt, B. S. (1990). *Cluster Analysis* (2nd edn). Gower Press: London.

Griesinger, W. (1861). *Die Pathologie und Therapie der Psychischen Krankheiten*. Krabbe: Stuttgart.

Hand, D. J. (1981). *Discrimination and Classification*. Wiley: New York.

Hempel, C. G. (1959). Introduction to problems of taxonomy. In *Field Studies in the Mental Disorders* (ed. J. Zubin), pp. 3–22. Grune and Stratton: New York.

Leff, J., Sartorius, N., Jablensky, A., Korten, A. & Ernberg, G. (1992). The International Pilot Study of Schizophrenia: five-year follow-up findings. *Psychological Medicine* **22**, 131–145.

Lorr, M., Klett, C. J. & McNair, D. M. (1963). *Syndromes of Psychosis*. Pergamon Press: New York.

McCutcheon, A. (1987). *Latent Class Analysis*. Sage: Beverly Hills, CA.

Manton, K. G. & Solds, B. J. (1985). Dynamics of health changes in the oldest old; new perspectives and evidence. *Milbank Memorial Fund Quarterly.* **63**, 206–285.

Meehl, P. E. (1962). Schizotaxia, schizotypy, schizophrenia. *American Psychologist* **17**, 827–838.

Moore, T. V. (1930). *The Essential Psychoses and Their Functional Syndromes.* Williams and Wilkins: Baltimore.

Schneider, K. (1959). *Clinical Psychopathology.* Grune and Stratton: New York.

Swartz, M., Blazer, D., Woodbury, M. A., George, L. & Landerman, R. (1986). Somatization disorder in the U.S. southern community: use of a new procedure for analysis of medical classification. *Psychological Medicine* **16**, 595–609.

Weber, A. C. & Scharfetter, C. (1984). The syndrome concept: history and statistical operationalizations. *Psychological Medicine* **14**, 315–325.

Wing, J. K. (1980). Methodological issues in psychiatric case identification. *Psychological Medicine* **10**, 5–10.

Wing, J. K. & Brown, G. W. (1970). *Institutionalism and Schizophrenia.* Cambridge University Press: London.

Wing, J. K., Cooper, J. E. & Sartorius, N. (1974). *The Measurement and Classification of Psychiatric Symptoms.* Cambridge University Press: Cambridge.

Wittenborn, J. R. (1951). Symptom pattern in a group of mental hospital patients. *Journal of Consulting Psychology* **16** (4). Reprinted in *Research in Psychopathology* (ed. H. C. Quay), van Nostrand: Princeton, N.J. (1963).

Woodbury, M. A. & Clive, J. (1974). Clinical pure types as a fuzzy partition. *Journal of Cybernetics* **4**, 111–121.

Woodbury, M. A., Clive, J. & Garson, A. (1978). Mathematical typology: a Grade of Membership technique for obtaining disease definition. *Computers and Biomedical Research* **11**, 277–298.

Woodbury, M. A. & Manton, K. G. (1982). A new procedure for the analysis of medical classifications. *Methods of Information in Medicine* **21**, 210–220.

Woodbury, M. A. & Manton, K. G. (1989). Grade of Membership analysis of depression-related psychiatric disorders. *Sociological Methods and Research* **18**, 126–163.

World Health Organization (1973). *Report of the International Pilot Study of Schizophrenia*, Vol. I. *Results of the Initial Evaluation Phase.* WHO: Geneva.

World Health Organization (1979). *Schizophrenia: An International Follow-up Study.* Wiley and Sons: Chichester.

World Health Organization (1992). *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines.* World Health Organization: Geneva.