Brit. J. Psychiat. (1979), **134**, 94–107

Psychotic and Neurotic Depression: 2. Clinical Characteristics

By MÁIRE NÍ BHROLCHÁIN, G. W. BROWN and T. O. HARRIS

SUMMARY A series of discriminant function analyses based only on clinical symptomatology suggests that psychotic and neurotic depressives do not differ globally but at a fairly specific level and that the principal, if not only, clinical difference between the groups is one of severity. This interpretation is reinforced by the fact that there are few symptoms more common in the neurotic group. There is suggestive evidence that symptoms generally thought to characterize neurotic depression may conduce to referral to psychiatric services. The neurotic depressive syndrome as classically conceived may therefore be an artificial one, created by selective factors bringing patients with particular symptoms into hospital populations.

In this paper, the second of a series of three, we examine whether the groups diagnosed as psychotic and neurotic depressives can be distinguished from each other by means of their clinical characteristics. To this end, we employ a number of discriminant function analyses. We ignore factor analysis since it cannot be used as a test of the two-group hypothesis; this is because, as Maxwell (1971) has indicated, the statistical requirements of the factor analytic model are contravened by the pooling of data, if there are two distinct homogeneous groups in the sample analysed. We have not used clustering methods, as we believe they will be profitable in this field only when good reasons are evident for choosing one particular target structure of clusters rather than another. There are many ways in which individuals might cluster with reference to one another; to search for groupings without taking this fact into account is a weak, not to say haphazard, strategy (Cormack, 1971).

The approach taken here to interpretation of the discriminant function analysis of psychiatric data has been discussed in detail in the preceding paper (Ní Bhrolcháin, 1978) from which two general points may be summarized briefly. First, it has been argued that, to avoid ambiguity, tests of the two-group hypothesis should be confined in the first instance to clinical symptomatology. Whether the groups differ aetiologically is a distinct issue and is dealt with in the third paper of this series (Brown, Ní Bhrolcháin and Harris, 1977). Second, in attempting to discriminate between two diagnostic groups, one may test for differences of a higher and lower order. A 'minimizing' analysis, tending to minimize the chances of finding bimodally distributed discriminant scores, includes all the symptoms and signs descriptive of depression and is conceived of as testing for differences at a general, illness level. A 'maximizing' analysis, tending to maximize the chances of finding bimodally distributed scores, is based on the features which differ between the groups and is construed as testing for the existence of lower-level group differences. In current psychiatric terminology, convenient and well-understood terms are not readily available, to refer to these levels of similarity and dissimilarity. Essen-Möller's (1973) distinction between 'gross' and 'specified' syndromes conveys some of the intended meaning, as does Foulds's notion of 'classes' and 'groups' (Foulds, 1965; Foulds and Bedford, 1975).

The notion of levels of psychiatric nosology has, of course, a direct parallel in biological classification in the idea of taxonomic rank. However, it does not necessarily imply a hierarchical arrangement of diagnostic groups.

Nosological studies of a large variety of diagnostic groups would be necessary for further clarification; the extent of similarities and dissimilarities between two diagnostic groups could then be evaluated by reference to their closeness to and distance from other diagnostic groups.

Design and Data Collection

The patient group on which the study is based has been described fully in Brown, Sklair, Harris and Birley (1973). The records of female patients admitted as in-patients to the Maudsley, Bethlem and St. Francis Hospitals were screened regularly and those aged 18 to 65, resident in the old borough of Camberwell in South London, were eligible for inclusion. To identify out-patients, psychiatrists at the clinics serving Camberwell were contacted regularly and asked for the names of patients recently admitted and presenting with depressive symptoms. In- and out-patients were admitted to the study if an important change had occurred in their condition during the 12 months prior to admission (in-patients) or first out-patient contact (out-patients). In subsequent discussion, 'admission' refers, for out-patients, to the first out-patient contact. In all, 114 patients were examined: 73 in-patients and 41 out-patients. There were two, and in some cases three, independent interviews: (a) the research psychiatrist, using a schedule based on the PSE (Wing *et al*, 1974), conducted the clinical interview without inquiring into social circumstances such as life events and other social difficulties and on this basis diagnosed the patient as psychotically or neurotically depressed. The diagnosis was made, therefore, purely on the symptoms and impairment presented at interview and was unbiased by any halo effects resulting from the assumption of a greater constitutional component in psychotic depression. Sixty-two patients were rated as suffering from psychotic and 49 from neurotic depression; 2 patients with definite manic symptoms were excluded. (b) One of two research sociologists interviewed each patient about her life events and other social circumstances during the year prior to interview. Although this interviewer also obtained details of the onset of the disorder, she was unaware of the psychiatrist's diagnosis. (c) The second sociological interviewer saw a close relative of the first 50 patients in the sample and collected the same social information, together with data on the timing of the onset of the patient's condition (Brown *et al*, 1973).

The clinical ratings comprise 74 distinct scales: a complete list of these is given in the Appendix. (Available on application to author). They fall into three groups: (i) *Symptoms:* 55 symptoms were rated present or absent; they cover a wide range, descriptive of depressive conditions, and include most of the items generally regarded as characteristic of psychotic and neurotic depression. (ii) *Severity ratings:* 15 severity scales, having from 4 to 7 points and measuring the severity of groups of symptoms, such as anxiety, retardation, sleep disturbance, etc. Two are ratings of the number of symptoms present. (iii) *Onset and development items:* 4 scales describing, e.g. fluctuating course of the disorder.

The mental state of the patient was rated on all these scales, both for the time of admission and for each of the critical change points including initial onset, occurring in the year prior to admission. The pre-admission data were obtained by retrospective questioning. Their reliability may be gauged from the 86 per cent agreement, for the first 50 patients examined, about the dating of onset between the account of the patient herself and that of the relative interviewed. One of the present authors (G. W. Brown) subsequently repeated the diagnostic classification without knowledge of the research psychiatrist's allocations. Overall agreement between the two sets of diagnoses was 78 per cent. Assuming the two raters have the same rate of error, the proportion of agreements between them is given, where p is the proportion of correct diagnoses made by each rater, by $p^2 + (1-p)^2$. This quantity is equal to 0.78 so that p, the accuracy figure for each rater, is 0.87 or 87 per cent.

Some preliminary points

In line with the arguments of the preceding paper, we have restricted our test of the two-group view to clinical items and have omitted aetiological variables as well as personality and

https://doi.org/10.1192/bjp.134.1.94 Published online by Cambridge University Press

other background features such as age. But even within this area, it is important to distinguish symptoms as such from their interference with the patient's daily life. It has been argued (e.g. by Wing, 1973) that severity of symptoms and degree of impairment are not strictly correlated and should be analysed independently. However, when it comes to differential diagnosis, we believe that impairment is as much part of the clinical picture as symptomatology and for this

TABLE I

Admission data: items and their weights in the discriminant function analyses and incidences in the psychotic and neurotic groups

	Weights	% positive Psychotics	% positive Neurotics	P < *
Analysis based on:				
(*a*) *8 items significantly associated with diagnosis at the 5 per cent level (A.1)*				
Retardation	2.102	73%	31%	.001
Waking early	1.518	56%	22%	.05
Severity of obsessions/preoccupations**	0.880	82%	63%	.05
Severity of impact on employment/household tasks**	0.628	84%	65%	.05
Loss of appetite	0.679	81%	63%	.05
Slowness of thinking/speech	0.189	44%	20%	.05
Overall severity****	−0.340	16%	41%	.01
Total number of symptoms**	−1.087	58%	39%	.05
(*b*) *17 items associated with diagnosis at the 20% level (A.2)*				
Retardation	2.166	73%	31%	.001
Waking early	1.623	56%	22%	.05
Severity of obsessions/preoccupations**	1.047	82%	63%	.05
Severity of impact on employment/household tasks**	0.644	84%	65%	.05
Middle insomnia	0.557	52%	37%	.2
Diurnal variation, worse am	0.532	45%	31%	.2
Initial insomnia	0.487	69%	55%	.1
Loss of weight	0.412	48%	31%	.1
Loss of appetite	0.147	81%	63%	.05
Severity of effect on day-to-day routine**	−0.069	71%	55%	.1
Slowness of thinking/speech	−0.147	44%	20%	.05
Severity of appetite disturbance***	−0.253	35%	47%	.1
Overall severity****	−0.538	16%	41%	.01
Decisions	−0.545	60%	47%	.2
Verbal attacks	−0.681	29%	43%	.2
Specific worry	−0.762	53%	65%	.2
Total number of symptoms**	−1.106	58%	39%	.05
		N = 62	N = 49	

(*c*) *39 items (A.3)*—The reliability of the estimates is so low that the weights are not given here (see Table II).

* This column gives the significance level of the association between the individual symptoms at onset and the psychotic/neurotic dichotomy, by uncorrected χ^2.

** These items, originally 4–5 point scales, have been dichotomized at the best discriminating point. All other items are binary ratings.

*** Dichotomized at marked and moderate loss vs little or no loss or some gain in weight: weight gain was very rare.

**** Severity is rated from 1 (severe) to 5 (mild) and dichotomized at the best discriminating point. On the dichotomization 0 represents, therefore, high and 1 low overall severity. The other severity scales are rated in the intuitively obvious way—a high score representing high severity.

reason, we have included items such as impairment of the patient's employment in the analyses to be described.

Using the discriminant function to test the distinctness of two groups raises the question of what statistical criterion should be relied on in judging whether one or two groups are really present. Three possible criteria for the decision suggest themselves:

(1) R^2, the proportion of the total score variance accounted for by the two-group division, is the statistic most commonly used for evaluating the success of a regression equation in predicting a dependent variable (here, group membership).

(2) The theoretical misclassification rate might also be considered; this allows an estimate of what percentage of errors would be made by the discriminant function in classifying new individuals into one or other diagnostic category.

(3) The modality of the score distribution has been a popular criterion in this field, being used by for example, Carney, Roth and Garside (1965) and Kendell (1968, 1970). Nevertheless the bimodality criterion is, as Hope (1969) has pointed out, an arbitrary one since it requires that the group means be separated to an arbitrary extent. Hope (1969) points out further that diagnostic unreliability may have the effect of concealing bimodality in apparently unimodal distributions. Hope's argument and calculations require modification, the details of which are available on request from the authors. Modality is the criterion of all three about which least is known statistically.

We shall consider all three criteria in judging the outcome of the analyses reported below. Analyses of this kind ideally require an independent sample for validation purposes. Since a second sample is unfortunately not available in this study, our results must be considered tentative and subject to confirmation.

Discriminant function analysis

In all, six discriminant analyses were carried out, based on data sets ranging from a 4-item set to a 41-item set. The items were chosen for inclusion as follows:

O.1 (Onset) and A.1 (Admission): 'Maximizing' analyses: These are based on the clinical features having incidences of 15 to 85 per cent in each of the two diagnostic groups and associated with diagnosis of the 5 per cent level of significance at onset (4 items) and admission (8 items). Significance levels are based on χ^2 to which, in order to include as many potential discriminators as possible, Yates's correction has not been applied. The items and their weights in the admission analysis are given in Table IA. These analyses correspond to the 'maximizing' strategy, outlined earlier and discussed in the previous paper.

Note that 'maximizing' is not strictly true of these analyses. Choosing e.g. the eight items most strongly associated with the psychotic/neurotic dichotomy at the zero order level does not ensure that these are the best 8 discriminators taken together. Moran (1966) discusses this general regression problem in the psychiatric context and explains that there is no practicable way of finding 'best' joint m discriminators of n possible variables, when n and m are large. 'Maximizing' is used here in the sense outlined in the preceding paper, viz. that the object in using a given set of items is to maximize, within the constraints mentioned in this footnote, the chances of finding bimodality in the score distribution.

O.2 (Onset) and A.2 (Admission): 'Intermediate' analyses: These include the clinical items having 15 to 85 per cent incidence and associated with the diagnostic dichotomy at the 20 per cent level of significance: 15 items at onset and 17 at admission. The items and their weights in the admission analysis are given in Table IB. We refer to these analyses as 'intermediate'. (For economy, items, incidences and weights for the onset analyses are not given; they closely resemble the admission results).

O.3 (Onset) and A.3 (Admission): 'Minimizing' analyses: These are based on symptoms and severity ratings which have incidences of between 15 per cent and 85 per cent in each diagnostic group, regardless of their association with diagnosis: 41 at onset and 39 at admission. They correspond to the 'minimizing' strategy outlined above.

The elimination of items with low or high frequencies in *either* group is necessary by virtue of the assumption of equal variance—covariance matrices for the groups involved in the derivation of the linear discriminant function. Maxwell (1971) suggests the use with binary data of items having incidences in the range 20 to 80 per cent. The number and nature of the items which would have been excluded thereby from the admission analyses were considered sufficient grounds for which to relax the range down to 15 per cent and up to 85 per cent. It might be objected that this restriction in itself reduces the likelihood of discriminating successfully between the groups, since some clinical items might be very common in one group and very rare in the other. This was not true of any of the items excluded on these grounds, all of which had either high or low incidences in *both* groups —for example 77 per cent of psychotics and 92 per cent of neurotics were positive on 'Crying' at admission ($p < .05$) and 8 and 24 per cent, respectively, on 'Diurnal variation, worse p.m.' ($p < .05$). There are, in fact, statistical procedures for the derivation of a discriminant function when the variance-covariance matrices are not similar in the two groups (cf. Anderson, 1975 and references), but these came to our attention too late to investigate their applicability to data of the kind analysed here.

Results

The statistics associated with the six analyses are given in Table II.

(1) *R^2:* If we were to judge solely by R^2, interpretable as the ratio of between-groups to total variance, we should probably reject the two-group hypothesis out of hand. For the six analyses, R^2 ranges between 0.15 and 0.45 (see

TABLE II

Statistics associated with the discriminant function analyses based on symptoms and severity ratings at onset and admission. N's are 62 psychotic depressives and 49 neurotic depressives

Data sets include items with incidences of 15%–85% associated by χ^2 with the psychotic/neurotic dichotomy at:

	5% level		20% level		All items with incidences of 15%–85%	
	O.1 Onset	A.1 Admission	O.2 Onset	A.2 Admission	O.3 Onset	A.3 Admission
Number of items	4	8	15	17	41	39
R^2	0.15	0.29	0.23	0.35	0.41	0.45
Corrected R^2*	0.13	0.24	0.12	0.24	0.07	0.16
F test	$p < .002$	$p < .0001$	$p < .05$	$p < .001$	$p > .3$	$p = .08$
r_{10}**	0.72	0.75	0.42	0.63	0.00	0.26
χ^2 goodness-of fit test***	$\chi^2 = 144.3$ df = 5 p = 0.0	$\chi^2 = 22.1$ df = 6 $p < .002$	$\chi^2 = 24.4$ df = 8 $p < .002$	$\chi^2 = 5.2$ df = 8 $p > .7$	$\chi^2 = 4.7$ df = 8 $p > .75$	(i) $\chi^2 = 15.46$ df = 8, $p = .051$ (ii) $\chi^2 = 7.6$ df = 6, $p > .25$
Theoretical mis-classification rates						
Psychotics	33%	24%	29%	21%	17%	16%
Neurotics	34%	28%	30%	26%	25%	22%
Total	33%	26%	29%	23%	20%	19%

* This correction is designed to adjust for capitalization on chance effects. See e.g., Guilford (1965) p. 401.

** r_{10} is the minimum correlation another discriminant function must have with a given function in order not to differ from it at the 10% level of significance. It is, thus, a measure of the precision of estimation of the discriminant function; see Hope (1968) Chapter 7, and Fisher (1940).

*** The χ^2 tests are based on cells obtained by dividing the score distribution into intervals half a (total group) standard deviation wide. Cells with expected frequencies of less than 5 are combined. See text for discussion of the two values of χ^2 given for analysis A.3.

https://doi.org/10.1192/bjp.134.1.94 Published online by Cambridge University Press

Table II). By any standards, these figures are extremely low; at most, the diagnostic divide accounts for less than half of the score variance (45 per cent in analysis A.3). Furthermore, the 'raw'-R^2 is always an overestimate since R^2 will certainly decrease when the functions are applied to an independent sample: R^2 will be overestimated in any particular sample and, with a given sample size, the greater the number of items used the greater will its standard error be. The adjustment used here is that given in Guildford (1965) p. 401 and must be considered only an approximation. Some idea of how sharp the reduction of R^2 would be is given by the 'Corrected R^2' figures in Table II. These adjusted figures range from 0.07 to 0.24, indicating that we could expect less than a quarter of the score variance to be attributable to diagnosis if the best of these functions were used on a new set of patients. It is true that the F test is significant in the maximizing and in the intermediate analyses (but not in the minimizing analyses): all this means, however, is that their R^2's differ significantly from zero. In summary, the very low R^2 figures give little or no support to the two-group view and clearly favour the one-group hypothesis.

(2) *Misclassification rates:* The theoretical misclassification rates in the six analyses range from 19 per cent for the minimizing admission data analysis to 33 per cent for the maximizing onset data analysis. The usefulness of the misclassification rate as a criterion is doubtful, however, since there is no conventionally accepted level beyond which an error rate is considered 'too high'. The 19 per cent and 20 per cent figures obtained for the minimizing analyses (A.3 and O.3) may be fairly low for this kind of data, but some might argue that a 20 per cent overlap is too large and that the separation could only be considered satisfactory if the overlap were, say, below 10 per cent. Besides, if the relatively low misclassification rates of the minimizing analyses are taken as supporting the two-group view, these same analyses' poor showing on the R^2 and modality criterions, as we shall see below, counteracts this support.

(3) *Modality:* The χ^2 goodness-to-fit test was applied to test for departures from normality in the score distributions of the six discriminant analyses. This is a rather strong test of unimodality since there is no reason to suppose that if there is only one mode, the null hypothesis distribution should be normal, at any rate with a fairly small sample. Since there appears to be no test for modality in existence, a test for normality is the only alternative to simple visual inspection. Results are given in Table II. The score distributions for all six analyses are given in Fig 1.

Both of the maximizing analyses (O.1 and A.1) have score distributions that depart significantly from normal and both tend fairly clearly to bimodality; this is seen in Fig 1a. The minimizing analyses (O.3 and A.3), on the other hand, are not significantly different from normal in their score patterns and appear fairly clearly unimodal (cf. Fig 1c). This conclusion might be questioned in the case of the minimizing admission data distribution (A.3) which in the goodness-of-fit test with 8 df has a χ^2 of 15.459 and an associated probability value of .051. The main contributors to this χ^2 are the three extreme cells at each end of the distribution where the observed numbers substantially exceed the expectations from normality. Since this sort of departure fron unimodality is not critical, the test which excludes these, having a χ^2 of 7.6, 6 df and $p > 0.25$ is probably a better guide.

The intermediate analyses—which may be considered, for those who would think the 5 per cent significance level too restrictive, a possible alternative to the maximizing analyses —are discrepant; the score distribution of the intermediate onset analysis (O.2) is significantly non-normal and that of the comparable admission analysis (A.2) does not depart from normality (cf. Fig 1b). It will be seen, however, that the discrepancy in this respect between the two intermediate analyses is probably of little consequence.

In summary, the results of the minimizing analyses (O.3 and A.3), judged by both R^2 and the bimodality criterion, would suggest that the psychotic and neurotic groups cannot be distinguished at a general, 'illness' level. (The

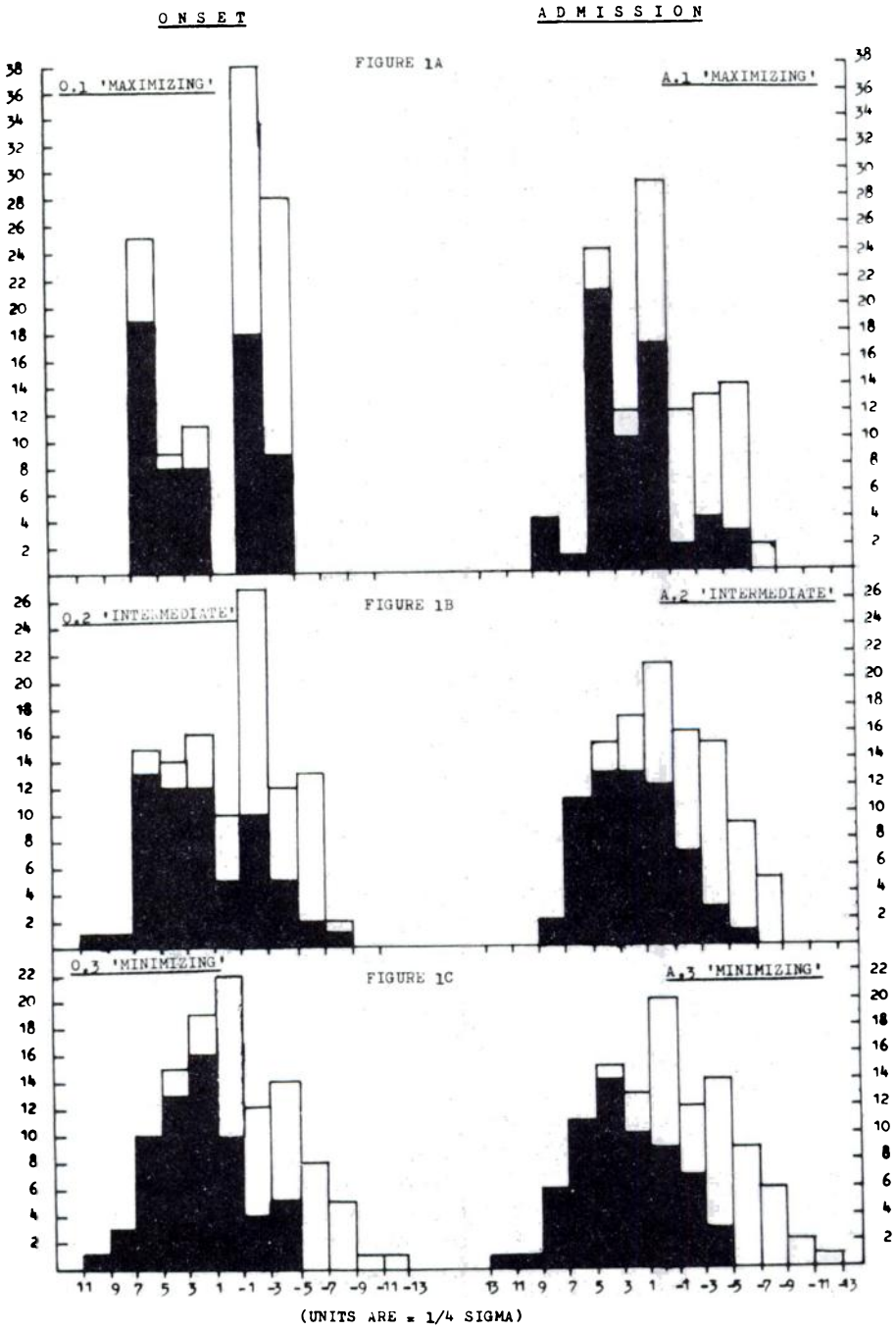

FIG 1.—Distributions of discriminant function scores for onset and admission. The shaded area represents psychotic patients and the unshaded area neurotic patients.

https://doi.org/10.1192/bjp.134.1.94 Published online by Cambridge University Press

reasons for dealing with such an apparently trite question are discussed in the preceding paper). Their misclassification rates are low enough to suggest the opposite conclusion, but this is very much a matter of opinion and, in view particularly of the low adjusted R^2's, we should be inclined to discount them. Taking modality as the criterion, the intermediate analyses give an equivocal result—distribution 0.2 tending to bimodality and A.2 to normality —but their low R^2's and fairly substantial error rates would seem, again, to give little support to the two-group view. The maximizing analyses, on the other hand, give clear evidence of bimodality and, despite their low R^2 values and fairly high error rates, might be interpreted on these grounds, in view of the popularity of the bimodality criterion, as giving limited support to the view that *at some fairly specific level* the groups are distinguishable. Nevertheless, the nature of the items with which these two bimodal distributions have been obtained points to the possibility that even a tentative conclusion in favour of two sub-groups would be ill-founded; this is because indices of severity predominate among the items included in the maximizing analyses which produce the bimodal score distributions.

Three of the four items included in the maximizing onset analysis—i.e. those associated at onset with the psychotic/neurotic dichotomy at the 5 per cent level—are indices of severity. These are overall severity, severity of obsessions/preoccupations and total number of symptoms; the fourth (retardation) might also be interpreted as essentially a measure of severity of depression. Retardation could be seen as primarily a matter of severity if severity of depression is construed as an index of the degree of 'hopelessness' and of the 'giving up-given up' complex described by Engel (1967). Its association with the rating of 'overall severity' is 0.42, as measured by the phi coefficient: 90 per cent of patients positive on retardation are rated high on overall severity, compared with 53 per cent of the non-retarded patients. Similarly, four of the eight clinical items significantly associated with the dichotomy at admission, and included therefore in the maximizing admission analyses (A.1), are severity measures (see Table Ia). Retardation is also among these eight and has in all the analyses reported the highest weighting. It is clear that severity indices are over-represented among the statistically significant clinical differences between the groups; four of the 15 severity items compared with 4 of the 55 symptom items are significantly associated at admission with diagnosis. We may ask, therefore, whether the clinical difference between psychotic and neurotic depressions is either wholly or principally a matter of severity. (There are many possible concepts of 'severity', e.g. severity of hopelessness, severity as retarded quality, as depth of depression or degree of impairment. While it is probable that all of these are closely associated with one another, we must for the moment leave open which one or combination of these might be critical. Our rating of 'overall severity' has, in fact, been based more on the component of impairment and on the intensity of any psychiatric symptoms, not necessarily depressive in character, such as panic attacks. It is thus not simply an index of the severity of depressed mood). We refer to this proposition as the 'severity hypothesis'.

One rather unsatisfactory way of dealing with this question involves the discriminant function; a second—simpler and more intuitive—is based on the raw data on symptomatology. We consider each of these in turn.

(1) *The argument from the discriminant function:* The discriminant function cannot tackle the question whether the only difference between the psychotic and neurotic groups is in the severity of disturbance. Discriminant function analysis can, however, establish how successfully the groups can be separated by means of severity ratings. It might be suggested that confirmation of the severity hypothesis would require that the severity measures alone should perform better than symptom items alone in discriminating between the groups. While this is evident in a simple way from the fact that many more severity ratings, proportionately, than symptoms are associated with the diagnostic dichotomy, this suggestion would impose a false dichotomy on the clinical ratings. The severity scales are partly presence/absence ratings, since a severity rating of nil is given if

none of the group of symptoms covered by a given severity scale is present. The symptom ratings, similarly, are partly measures of severity; they are rated present if above a certain threshold laid down by the rating criteria.

Two further discriminant analyses were carried out based on (a) all but two of the severity indices—a total of 13 items (A.4) ('Severity of elated mood' is excluded since only one patient received a non-zero rating and 'number of new symptoms' since it is a measure of acuteness of change in symptoms rather than strictly of severity), and (b) the 13 items in (a) above together with 'retardation' (A.5). Admission data only will be considered as the results for onset are similar.

Do these analyses succeed in separating the groups? Analysis A.4, based on the thirteen severity items alone, does not perform particularly well; its R^2 is low at 0.18 (0.08 corrected), it has a fairly high misclassification rate of 32 per cent and its score distribution does not depart significantly from normality ($\chi^2 = 7.51$, df = 6, $p > 0.25$). Adding retardation as a discriminator improves the separation; analysis A.5 has an R^2 of 0.28 (0.18 corrected), an overall misclassification rate of 27 per cent and its score distribution is significantly non-normal ($\chi^2 = 13.37$, df = 6, $p < 0.05$), appearing visually to tend to bimodality. (For the comparable analyses of onset data, the reverse is the case: the 13 severity items score distribution is significantly non-normal and the 14-item score distribution (including retardation) does not depart from normality). In these three respects, analysis A.5 is comparable to the maximizing admission analysis (A.1). In other words, the psychotic and neurotic groups can be separated as effectively by using the severity measures together with retardation as they can by choosing, *post factum*, the 'best' eight discriminators taken singly.

It may be objected at this point that the separation achieved in analysis A.5 is nevertheless not a particularly good one—in that its multiple R^2 is low and the misclassification rates are fairly large. This is indeed so, but need not be fatal for what we have called the severity hypothesis. In fact, there are two forms that the severity hypothesis might take: (1) that on a true scale of severity (which, of course, the discriminant function is not) there is a bimodal distribution, comprising a mixture of the (higher severity) psychotic and (lower severity) neurotic groups' distributions; or (2) that on a true scale of severity, the total group of depressed patients is unimodally distributed and that psychotics tend to be drawn on average from the upper end of the distribution and neurotics from the lower end. If the first were the true state of affairs, we might expect to obtain a better separation with the discriminant function than if the second held. In either case, we should not expect the separation to be strikingly good, since the research psychiatrist diagnosed the patients on the assumption that other clinical characteristics besides severity of disturbance differentiate the groups. If severity is in fact the over-riding clinical difference between psychotic and neurotic depression, then diagnostic allocations based on other criteria would to some extent reduce the power of severity measures as discriminators. Furthermore, one cannot assume that existing measures of severity are sufficiently sensitive to discriminate well. More and better-validated indices of severity would be necessary before the hypothesis could be tested at all adequately by means of the discriminant function. In all, the evidence for the severity view from the discriminant function analyses presented here is, in our view, as good as it can be with the present data—encouraging and, at minimum, not negative.

Before leaving this approach, the role of retardation as a discriminator must be mentioned. Its large weightings, together with the large mean difference between the groups, indicate that its 'coefficient of individual determination' is by far the largest of all the items used throughout these analyses. This does not necessarily mean, as Hope (1968, Chs. 7 and 10) explains, that retardation is 'more important' in separating the groups than the severity items. However, it is worth noting that the single retardation item alone is quite an effective discriminator; if all patients who are retarded at admission were classified as psychotic and all the non-retarded as neurotic, 27 per cent

of psychotics and 31 per cent of neurotics would be misclassified, giving an overall misclassification rate of 29 per cent. This compares with an error rate of 35 per cent if overall severity were used alone as a discriminator. Whether the effectiveness of the retardation item reflects the greater reliability of presence/absence symptom ratings and how far its success is due to the fairly sizeable association between retardation and the overall severity rating cannot be established here. Retardation might also, as we have mentioned, be seen as itself an index of severity.

(2) *The argument from symptomatology:* If the difference between psychotic and neurotic depression is simply one of severity, then the incidence of all symptoms should be higher among psychotics than among neurotics. Of the 55 symptoms rated in this study, there are only two on which the group differences reach the 5 per cent level of significance and which are also more common in the neurotic group: 'Crying' and 'Diurnal variation, worse in the evening'. Table III gives the percentages of each group having, at onset and admission, these and other symptoms commonly held to be neurotic features. These figures give little evidence of specifically 'neurotic' signs or symptoms of depression and underline the remark of Pilowsky, Levine and Boulton (1969) that neurotic depression is often diagnosed by default. Kendell's (1976, p. 18) characterization of Type B depression—as denoting '. . . milder illness, prone to fluctuate from day to day and lacking the characteristic features of (Type A)'—is in a similar vein. Foulds (1973), in proposing a more general version of the severity hypothesis presented here—viz. that the relationship between psychotic and neurotic depression is an inclusive one—has cited as supporting evidence the results of Foulds and Caine (1959), Foulds (1965) and Garside *et al* (1971), all of which are consistent in revealing few, if any, clinical features more common in neurotic than in psychotic depression. The lack of symptoms and signs characteristic of neurotic depression is also reflected in Kendell's data: of the 39 clinical items (including those covering 'history of present illness') given in Tables I and II of Kendell (1968), only two—'duration before admission over 12 months' and 'serious suicide attempt'—are significantly more frequent among the neurotic depressives than among the psychotics.

It could of course be argued that the present results are due to the absence from the present study of such traditionally 'neurotic' items as reactivity of depression, self-pity, hysterical symptoms and hypochondriasis (e.g. Kiloh and Garside, 1963). However, there are plausible grounds for supposing that even the marginally higher incidences among diagnosed neurotics of those 'neurotic' symptoms on which we do have ratings—those given in Table III below—may be artefacts of the selective factors bringing some patients into treatment and allowing

TABLE III

The incidence of 'neurotic' symptoms in the psychotic and neurotic groups

	At Onset			At Admission		
	Psychotics (N = 62) %	Neurotics (N = 49) %	p*	Psychotics (N = 62) %	Neurotics (N = 49) %	p*
Crying	65	78	ns	77	92	< .05
Diurnal variation, worse pm	8	18	ns	8	24	< .05
Specific worry	60	51	ns	53	65	ns
Panic attacks	16	20	ns	21	29	ns
Irritability	65	71	ns	74	80	ns
Verbal attacks	22	33	ns	29	43	ns
Violence/destructive behaviour	5	12	ns	8	18	ns

* Uncorrected χ^2 has been used. The significance of the association between diagnosis and 'Crying' and 'Diurnal variation, worse p.m.' at admission is not affected by applying Yates's correction.

others to evade the treatment net. The argument here is that psychotic (or severly disturbed) depressives who reach psychiatric treatment are a more representative sample of their population of origin than are neurotics who reach hospital services. There are a number of reasons for this supposition: (i) Psychotics are more severely depressed and they are, because of this, less likely to be subject to the vagaries of the social selective factors which conduce to treatment-seeking and referral. The vast majority of schizophrenics, for example, are likely to receive hospital treatment at some time during their illnesses; their conditions being perhaps less 'normal' and more severe than depressive ones they are more readily seen as needing psychiatric care. See Brown, Ní Bhrolcháin and Harris (1975) for a brief discussion of some potential social selective factors. (ii) Psychotic depression is generally assumed to occur in patients with stable pre-morbid personalities and 'neurotic' depression among those with a previous history of neurotic traits. The appearance of psychotic symptoms among 'normal' personalities might, therefore, be seen as so unexpected and inexplicable that the problem is immediately defined as 'psychiatric' and thus a matter for referral to psychiatric services. In the neurotically depressed, on the other hand, depressive symptoms may appear to be no more than an exacerbation of enduring personality problems. Those of the neurotics who do reach in- and out-patient status may, it would be predicted, have had psychiatric referral suggested to them by a person with some technical experience of psychiatric illness, who would have recognized the symptoms as such. In particular, it might be expected that certain kinds of symptoms may conduce to referral. Far-fetched though they may appear, these conjectures receive some support from our data.

From the social interview, we have information on who first suggested to the patient that she seek psychiatric help. We have grouped these sources under four heads: (1) the patient herself (S); (2) the patient's husband or boyfriend; (3) 'official' referrals, e.g. by a G.P., social worker, probation officer and (4) other sources, e.g. friends, relatives, workmates. Although psychiatric treatment was first suggested to only 15 per cent of patients by 'official' sources, the suggestion came from an 'official' source more often in the neurotic (22 per cent) than in the psychotic group (10 per cent), as is seen in Table IVa. This slight difference is increased when neurotics and psychotics are compared according to their scores on a neurotic symptoms index: this index is the number of the seven 'neurotic' symptoms given in Table III on which the patient is rated positive. Neurotics having low 'neurotic' scores are no more likely than psychotics with low scores to have had treatment first suggested by an 'official' source, but of the high scorers, none of the psychotic patients, compared with one-third of the neurotic patients, were originally referred by an 'official' source (see Table IVb). The same trend is evident when neurotic and psychotic patients, positive on each of the seven 'neurotic' symptoms, are compared for the origin of referral: neurotic patients positive on these 'neurotic' symptoms have had referral originally suggested to them by an 'official' source more frequently than psychotic patients positive on the 'neurotic' items. While only two of the individual symptom tables reach significance ('irritability' and 'verbal attacks') such a result would be expected by chance among seven tabulations with a probability of less than 5 per cent and all seven tables go in the same direction. If the number of referrals which were originally suggested by an 'official' source can be construed as an index of biased selection, then the proportions of hospital neurotics displaying features such as violence, irritability, verbal attacks and so on may overestimate the frequency of these symptoms among neurotic depressives as a whole—that is, among those in and out of hospital. The apparent existence of symptoms 'characteristic' of neurotic depression may, in other words, be an artefact of selection, since these symptoms in themselves may bring about referral to psychiatric services of individuals not displaying psychotic symptoms. Three of the seven symptoms considered here are un-social, acting-out features: 'verbal attacks', 'irritability' and 'violence/destructive behaviour'—a point which raises the possibility that it is the social distress these symptoms cause which conduces to the overreferral of the non-psychotic patients

TABLE IV

Source of referral (a) by diagnosis, (b) by 'neurotic' symptoms score. Figures in brackets are row percentages

(a)

Diagnosis	Official	Husband	S & Other	Total
Psychotic	6 (10)	24 (39)	32 (52)	62
Neurotic	11 (22)	10 (20)	28 (57)	49
Total	17 (15)	34 (31)	60 (54)	111

$\chi^2 = 6.06$, df = 2, $p < .05$
Phi (Official v rest) = .18; Maximum phi = 0.48

(b)

	Patients with 'neurotic' symptoms score of 0–3			Patients with 'neurotic' symptoms score of 4–7		
	Source of referral			Source of referral		
Diagnosis	Official	S, Hb & Other	Total	Official	S, Hb & Other	Total
Psychotic	6 (13)	40 (87)	46	0 (0)	16 (100)	16
Neurotic	3 (13)	20 (87)	23	8 (33)	18 (67)	26
Total	9 (13)	60 (87)	69	8 (21)	34 (79)	42
	$p > .3$*			$p < .02$		
	Phi = 0.0; Maximum phi = 0.55			Phi = 0.38; Maximum phi = 0.62		

* Significance levels are based on Fisher's exact test.

who display them. Hysterical symptoms and other features commonly supposed to characterize 'neurotic' depression but not represented here might well conform to the same expectation.

In short, the fact that some symptoms are marginally more common among neurotic depressives may be due to social selective influences on referral patterns. The evidence presented here is undoubtedly weak but does, we feel, give sufficient cause for (i) urging caution in assuming that diagnostic groups in hospital samples are equally representative of their populations of origin and (ii) raising the question whether the neurotic depressive syndrome, as classically conceived, may be an artificial one created by selective factors.

Conclusion

We believe that the evidence for the severity hypothesis is highly suggestive. Our results cannot be definitive and require the confirmation of an independent sample, which unfortunately has not been available to us. We hope that others will attempt to replicate our findings.

For clarity, it must be mentioned that the view that severity is the principal if not only clinical difference between psychotic and neurotic depression is not in conflict with the recurrent finding that depressive phenomena require more than one factor to represent them adequately. While the severity view certainly implies that depressed patients form a clinically homogeneous group this does not imply, as Maxwell (1971) has indicated, that the matrix of clinical correlations should be resolvable into a single factor. The severity view may be taken to imply, however, that factor solutions which incorporate a general severity factor might be preferable to the simple structure solutions commonly presented. A further implication is that severity is probably the major source of clinical differences between depressed patients and that, for prognostic and treatment purposes, it may be more useful in the long run to know how severely disturbed/retarded a patient is, providing other relevant details are also available, than to have a psychotic or neurotic diagnosis.

These conclusions agree in broad outline with the classic view of Mapother (1926) and Lewis (1938). We would be cautious, however, about Lewis's apparent assumption (1938, p. 877) that the severe condition is usually

acute and the mild condition usually chronic. This may tend to be true of hospital patients, but it may also be the case that, for example, acute mild conditions, although common in the community at large, do not reach hospital services. Similarly, psychotic or severe depression may occur more frequently in older patients with more stable personalities and neurotic or mild conditions in younger patients with a history of neurotic traits. On the other hand, selective factors may be such that mild conditions in stable personalities, though common, do not reach psychiatric services. But even if it is true that the psychotically or severely depressed, in contrast to the neurotically or mildly depressed, are older, more stable and subject more to recurrent episodes, this is not incompatible with the proposition that severity of disturbance may be the clinical, symptomatological difference between psychotic and neurotic depression.

In conclusion, we would suggest that the relationships between chronicity, periodicity and symptoms, along with non-illness variables such as age and personality, need thorough investigation in unselected samples of depressed patients in the general community before valid and stable schemes of classification and diagnosis can be arrived at.

Acknowledgements

The research was supported by the Medical Research Council and the Foundations Fund for Research in Psychiatry. We are indebted to Professor J. R. M. Copeland who made the original clinical diagnoses and to Freda Sklair and Jenny Frankland who also interviewed the patients. We are particularly grateful to Dr K. Hope for his valuable comments on an earlier draft of this paper.

References

ANDERSON, J. A. (1975) Quadratic logistic discrimination. *Biometrika*, **62**, 149–54.

BROWN, G. W., SKLAIR, F., HARRIS, T. O. & BIRLEY, J. L. T. (1973) Life events and psychiatric disorders. Part I: some methodological issues. *Psychological Medicine*, **3**, 74–87.

—— Ní BHROLCHÁIN, M. & HARRIS, T. O. (1975) Social class and psychiatric disturbance among women in an urban population. *Sociology*, **9**, 225–54.

—— —— —— (1977) Psychotic and neurotic depression: 3. Aetiological and background factors, (forthcoming).

CARNEY, M. W., ROTH, M. & GARSIDE, R. F. (1965) The diagnosis of depressive syndromes and the prediction of ECT response. *British Journal of Psychiatry*, **111**, 659–74.

CORMACK, R. M. (1971) A review of classification. *Journal of the Royal Statistical Society* (A), **134**, 321–67.

ENGEL, G. L. (1967) A psychological setting of somatic disease: the 'giving up-given up' complex. *Proceedings of the Royal Society of Medicine*, **60**, 553–5.

ESSEN-MÖLLER, E. (1973) Standard lists for three-fold classification of mental disorders. *Acta Psychiatrica Scandanavica*, **49**, 198–212.

FISHER, R. A. (1940) The precision of discriminant functions. *Annals of Eugenics*, **10**, 422–9.

FOULDS, G. A. (1965) *Personality and Personal Illness*. London: Tavistock.

—— (1973) The relationship between the depressive illnesses. *British Journal of Psychiatry*, **122**, 531–3.

—— & BEDFORD, A. (1975) Hierarchy of classes of personal illness. *Psychological Medicine*, **5**, 181–92.

—— & CAINE, T. M. (1959) Some symptoms and signs of depression in women. *Journal of Mental Science*, **105**, 182–9.

GARSIDE, R. F., KAY, D. W. K., WILSON, I. C., DEATON, I. D. & ROTH, M. (1971) Depressive syndromes and the classification of patients. *Psychological Medicine*, **1**, 333–8.

GUILFORD, J. P. (1965) *Fundamental Statistics in Psychology and Education*. 4th edition. New York: McGraw Hill.

HOPE, K. (1968) *Methods of Multivariate Analysis*. London: University of London Press.

—— (1969) Review of 'The Classification of Depressive Illnesses' by R. E. Kendell. *British Journal of Psychiatry*, **115**, 731–4.

KENDELL, R. E. (1968) *The Classification of Depressive Illnesses*. Maudsley Monograph No. 18. London: Oxford University Press.

—— (1970) The problem of classification. *Recent Developments in Affective Disorders*. Special Publication No. 2, 15–26.

—— (1976) The classification of depression: a review of contemporary confusion. *British Journal of Psychiatry*, **129**, 15–28.

KILOH, L. G. & GARSIDE, R. F. (1963) The independence of neurotic depression and endogenous depression. *British Journal of Psychiatry*, **109**, 451–63.

LEWIS, A. J. (1938) States of depression: their clinical and aetiological differentiation. *British Medical Journal*, *ii*, 875–8.

MAPOTHER, E. (1926) Manic-depressive psychoses. *British Medical Journal*, *ii*, 872–9.

MAXWELL, A. E. (1971) Multivariate statistical methods and classification problems. *British Journal of Psychiatry*, **119**, 121–7.

MORAN, P. A. P. (1966) The establishment of a psychiatric syndrome. *British Journal of Psychiatry*, **112**, 1165–71.

Ní Bhrolcháin, M. (1978) Psychotic and neurotic depression: 1. Some points of method. *British Journal of Psychiatry*, **134**, 94–107.

Pilowsky, I., Levine, S. & Boulton, D. M. (1969) The classification of depression by numerical taxonomy. *British Journal of Psychiatry*, **115**, 937–45.

Wing, J. K. (1973) Social and familial factors in the causation and treatment of schizophrenia. In: *Biochemistry and Mental Illness* (edited by L. L. Iversen and S. Rose). London: The Biochemical Society.

—— Cooper, J. E. & Sartorius, N. (1974) *The Measurement and Classification of Schizophrenic Symptoms*. London: Cambridge University Press.

Máire Ní Bhrolcháin, B.A., M.Sc., *Research Fellow*,

Centre for Population Studies, London School of Hygiene and Tropical Medicine, Keppel Street, London W.C.1.,

George W. Brown, Ph.D., *Professor of Sociology*,

Tirril Harris, M.A., *Research Officer*,

Bedford College, Regent's Park, London N.W.1

(*Received 1 July; revised 21 November 1977*)