Brit. J. Psychiat. (1965), 111, 659–674

The Diagnosis of Depressive Syndromes and the Prediction of E.C.T. Response

By M. W. P. CARNEY, M. ROTH and R. F. GARSIDE

Introduction

The establishment of a classification of affective disorders commanding wide agreement among clinical practitioners and investigators is one of the most pressing needs of contemporary psychiatry. This group of conditions has, in recent decades, displaced schizophrenia from the centre of the clinical stage. However, despite its prominence and importance in clinical practice, the territory remains inadequately charted. There is evidence to indicate that the uncertainty about the most clear and convenient lines of demarcation within this clinical territory makes a large contribution to the unreliability of psychiatric diagnosis. Thus, in a recent enquiry (Sandifer, Pettus and Quade, 1965) into the reliability of diagnoses made in 91 first admissions to a mental hospital by ten experienced psychiatrists, it was shown that the resolution of disagreement in the areas of "psychoneurosis—affective disorder" and "psychoneurosis—personality disorder" would have raised the overall reliability of diagnosis in this enquiry from 57 per cent. to 83 per cent.

The prediction of the results of treatment is not only of practical importance but raises issues that are to some extent related to the problem of classification. For the effects of new forms of psychiatric treatment introduced during the past few decades have, on the whole, confirmed the validity of main lines of demarcation between entities originally established in the clinic. However, a valid classification is unlikely to be arrived at from the results of treatment alone (as is well illustrated in the field of general medicine by the examples of cortisone and penicillin). Nor can it be assumed that the results of treatment may be predicted on the strength of a psychiatric diagnosis alone.

This paper reports studies concerned with problems of diagnosis and classification and the prediction of treatment. As far as classification and diagnosis are concerned, a number of enquiries have attempted in recent years to apply modern statistical techniques to try to decide whether depressive illnesses are drawn from different points along a single continuum, or whether a number of qualitatively distinct forms of depression exist. Kiloh and Garside (1963) summarized a large literature bearing on this problem and came to the conclusion that it provided strong evidence to support the traditional dichotomy of depressive cases into neurotic and endogenous varieties. It is unnecessary to review this evidence again, and only a few recent studies especially relevant to the present enquiry will be mentioned.

Hamilton and White (1959) carried out a factor analysis on data obtained from 64 severely depressed patients, making use of Hamilton's rating scale (1960). Four factors were isolated, the first covering such clinical features as depressive mood, guilt, retardation and loss of insight. The unequivocally endogenous and reactive groups showed significantly different mean scores on this first factor. However, although there was a suggestion of bimodality in the distribution of scores, it did not depart significantly from normality. This inconclusive result may have been due, as the authors suggested, to the small number of endogenous patients involved, among other factors. In their attempts to identify features of high predictive value for the outcome of depressive illness following treatment with E.C.T., the same authors (1960) carried out a multiple regression analysis on data obtained from 49 male depressives. Five quantified variables derived from the analysis had a multiple correlation of 0·62 with outcome. The material

was relatively small and confined to the male sex, and doubt has recently been cast (Rose, 1962) on the predictive value of one of the variables (falling blood-pressure with Mecholyl). Hobson (1953) extracted from 121 clinical items recorded in each of 127 in-patients those features which proved to be significantly correlated with a good outcome. He did not attempt any diagnostic sub-division of his material, but Roberts (1959), who followed up his patients 3 months after the completion of treatment, confirmed the prognostic value of the features selected by Hobson, and concluded that the correlations between clinical item score (Hobson), the symptom scores on Hamilton's rating scale and the outcome of treatment strongly supported the view that there were two qualitatively distinct groups of depressed patients. These corresponded broadly to neurotic and endogenous groups of depressive illness.

In the discriminant function analysis carried out on the data of 97 patients treated with imipramine, Kiloh *et al.* (1962) found that the cluster of features positively correlated with a good response to imipramine corresponded closely with those selected by Hobson as predictors of success with E.C.T. The second cluster, which correlated negatively with a good response to imipramine, approximated closely to the clinical picture of neurotic depression, whereas the first corresponded to that of endogenous depression. A factorial analysis carried out by Kiloh and Garside (1963) on data obtained from 92 cases of depression provided further strong confirmation for the nosological independence of endogenous and neurotic depression. Of the two factors extracted, the first was a general factor and the second a bipolar factor. The latter accounted for a greater part of the total variance than the general factor, and had therefore been more important in producing the original correlations between the 35 clinical features analysed. Moreover, there was a correlation of 0·986 between the factor loadings on this second bipolar factor and the correlations with diagnosis. This factor therefore clearly differentiated between neurotic and endogenous depressions.

Aims

The present enquiry had a number of objectives. In the first place it aimed at extending previous observations on the prediction of the results of E.C.T. to an adequately large and representative material drawn from both sexes. The study was begun in parallel with the later enquiries by Kiloh and his associates into the results of treatment with imipramine; it was hoped that the results obtained from another form of treatment would provide independent evidence in relation to the classification of depressive disorders. As the study was planned to cover patients admitted to a mental hospital, it was also considered that it would provide the necessary complement to a number of previous investigations, including those of Kiloh, which had been conducted on out-patients, among whom milder forms of depression may well have been over-represented. A further objective was the derivation from statistical analysis of diagnostic and predictive indices that would make it possible to translate the clinical data into quantitative terms and in this way ease communication and promote uniformity in clinical practice.

Method

This study was restricted to in-patients for three reasons: a more detailed assessment of individual features was made possible, errors due to failure to co-operate in the prescribed course of treatment were eliminated, and the controlled environment of hospital tended to exclude adventitious factors which might influence the results of treatment. The depressed patients interviewed were consecutive admissions for E.C.T., under the care of ten different consultant psychiatrists in three psychiatric units in Newcastle upon Tyne. The decision to admit to hospital was made by the appropriate consultant independently of the investigators. Before the first treatment each patient was assessed for each item (*vide infra*) by one of the authors in an interview of 45 minutes. In order to decide the presence or absence of some items it was found necessary to supplement the information from the interview with material from the case notes and the observations of the

nursing staff. This was frequently the case with items relating to: previous history of nervous breakdown, family history, premorbid personality, recent behaviour in respect of reactivity and variability of depression and hysterical phenomena, weight loss, and delusional ideas —the last often eludes detection in the course of a single interview. Patients in whom the depression was judged to be the principal component of the illness were admitted to the investigation: those in whom depression was judged to be secondary to other psychoses, organic cerebral disease, epilepsy, drug addiction and subnormality, were excluded.

The possibility of a dichotomous distribution of depressive states was initially investigated in a retrospective study of the case histories of 101 patients admitted to the psychiatric unit, Newcastle General Hospital, during the years 1956, 1957 and 1958. To these case histories was applied a diagnostic index compiled from items considered on clinical grounds to be valuable in discriminating between endogenous and neurotic depressions. On the basis of this study, 45 features were chosen for use in the present work. Owing to the limitations of the Ferranti Pegasus Computer, 35 items only could be used for statistical analysis; those excluded were features which showed the lowest frequency of occurrence or the smallest discriminating power. For the former reason five items, viz. previous attack of mania, hypnagogic hallucinations, depressive stupor, poverty of ideas and amenorrhoea, and for the latter reason, two items relating to libido and depersonalization, were discarded. Certain pairs of features occurred together so frequently that they were combined to form a single item; thus persistent suicidal ruminations and determined suicidal attempts became "psychotic suicidal", suicidal threats and half-hearted suicidal attempts became "neurotic suicidal" and agitation and retardation became "depressive psychomotor activity". Depressive "stupor" requires further consideration; as a principal presenting feature it could not be readily distinguished from stupor due to other causes, e.g. drugs or catatonia; as an event in the course of a clear-cut depressive illness it differed from retardation in degree only and since it was also a rare occurrence it was recorded with it under depressive psychomotor activity. The 35 items finally used are listed in Table I.

TABLE I

Items Selected for Analysis

1. Age over 40 when first seen.
2. Adequate personality.
3. No adequate precipitants for illness.
4. No adequate psychogenesis.
5. Unvarying depression.
6. No reactivity of depression.
7. Distinct quality of depression.
8. Weight loss in excess of 7 lb.
9. Constipation.
10. Pyknic physique.
11. Previous episodes of depression.
12. Family history of endogenous depression.
13. Early waking.
14. Depression worse in mornings.
15. Depressive psychomotor activity.
16. Anxiety.
17. Delusions of retribution.
18. Nihilistic delusions.
19. Somatic delusions.
20. Paranoid delusions or gross ideas of reference.
21. Persistent suicidal ruminations, and/or determined suicidal attempt (psychotic suicidal).
22. Depressive hallucinations.
23. History of over 1 year with no symptom-free intervals.
24. Family history of neurosis, psychopathy or drug addiction.
25. Depression worse in evenings.
26. Tendency to blame others for illness.
27. Self-pity.
28. Hopeful attitude towards illness.
29. Hypochondriasis.
30. Suicidal threats or half-hearted suicidal gestures (neurotic suicidal).
31. Irritability.
32. Phobias.
33. Hysterical features or attitude.
34. Initial insomnia.
35. Guilt: (*a*) delusions of guilt;
    (*b*) guilt feelings; or
    (*c*) guilt free.

DEFINITIONS

Most of the terms used are self-explanatory. However, the following require further explanation.

Adequate Personality: This describes subjects free from any history of neurotic breakdown and without disabling neurotic symptoms or serious social maladjustment.

Precipitants: Psychological and/or physical events

https://doi.org/10.1192/bjp.111.477.659 Published online by Cambridge University Press

impressively related to the onset of symptoms and which appeared likely on other grounds to have played some part in the development of the illness.

No Adequate Psychogenesis: No psychological stress or difficulty continuing to operate after the onset of symptoms and adequate to explain perpetuation of the illness.

Unvarying Depression: Absence of marked fluctuations of affect from day to day or week to week.

No Reactivity: Absence of definite mood change in response to changes in the external environment.

Distinct Quality: Some patients may describe their depression as similar to "normal" sadness or gloom, differing in degree only; others describe their mood as having a quality quite distinct from the depression with which they normally react to adversity. It is to this latter type of depression that this feature refers.

Pyknic Physique: The assessment of somatotype was made by clinical impression according to criteria laid down by Kretschmer (1926).

Depressive Psychomotor Activity: This term is used inclusively to describe any objective evidence of psychomotor slowing, stupor or agitation.

Nihilistic Delusions: Delusions of doom, imminent destruction, somatic dissolution or poverty of the patient and/or his family.

Somatic Delusions: Delusions of bodily change or disease, usually of a bizarre nature.

Hypochondriasis: Excessive or morbid preoccupation with bodily sensations which have little or no organic basis.

A score of one was assigned to each clinical feature if present, and a score of nought if absent, except in the case of "guilt" when "delusions" scored two, "feelings" one and "guilt free" nought.

Patients who received three or more E.C.T. were followed up for a total period of 6 months. Assessments of improvement were made before discharge from hospital by clinical examination and also consultation with medical staff and charge nurses. Follow-up assessments were made at 3 months and 6 months by one of the authors. Four degrees of improvement were recognized; A, symptom-free with full social recovery; B, full social recovery with residual symptoms; C, marked improvement with incomplete social recovery; D, slight improvement, no better or worse. In order to avoid spontaneous mood change being attributed to the effect of E.C.T., patients who had failed to make a satisfactory immediate response to E.C.T. (C + D) were graded D at the later assessments, and the gradings at 3 and 6 months could be no higher than the lowest grade previously achieved. At each assessment, the details of any medication received during the follow-up period were recorded.

Criterion diagnoses were established within a few days of admission to hospital by a review of all the available clinical information, including the case history, social history, consultant's opinion and nurses' observations. There were three possible diagnoses, definite endogenous depression, definite neurotic depression and doubtful type depression ("endogenous" and "neurotic" are used in the sense defined by Kiloh and Garside (1963)). The number of doubtful diagnoses was small because observation in hospital permitted greater certainty in diagnosis.

In order to ascertain whether the data were consistent with the hypothesis that endogenous and neurotic depression are more or less severe manifestations of a single depressive condition or, on the other hand, indicated that they are different entities, a principal component factor analysis (Thomson, 1956) was carried out on a Pegasus Computer. This analysis was based upon the correlations among the 35 features, self-correlations being equated to unity. The statistical significance of the residual correlations was calculated as indicated by Burt (1952).

The difference between endogenous and neurotic depression having been established, multiple regression analyses (Snedecor, 1956) were carried out on the computer to ascertain the weighting coefficients which, when applied to the feature scores, give the best prediction of diagnostic grouping (i.e. endogenous or neurotic depression) and of the effect of E.C.T. at both 3 and 6 months. In the diagnostic analysis, 13 cases in which some uncertainty attached to the diagnosis were excluded; they were included again in all the statistical analyses that followed. It should perhaps be pointed out that when the predicted variable has a dichotomous distribution, as in the case of two diagnostic groups, multiple regression analysis, as Fisher (1941) has indicated, is formally equivalent to discriminant function analysis.

The principal component method of factor analysis used in the present study is more

accurate than that used (simple summation) in a similar context by Kiloh and Garside (1963), and permits more precise tests of significance of the residual correlations. Moreover, the milder cases in the Kiloh and Garside study might have lent themselves to diagnostic differentiations which did not hold in the more severe forms of depression considered here.

Results I

Differentiation between endogenous and neurotic depression

One hundred and forty-eight subjects were finally admitted to the study after rejection of 17 patients in whose case a few days' observation or the effects of the early stages of treatment revealed that schizophrenia, organic psychosis or obsessive compulsive neurosis were the underlying disorders. Thirteen of the 148 cases had to be excluded because treatment was discontinued by the clinician in charge before they had received the three E.C.T. necessary to qualify for admission to the study.

Of the 135 remaining patients, six could not be traced for follow-up purposes. Follow-up observations were completed at 3 months in 129, and at 6 months in 108 cases.

Table II shows the outcome with E.C.T. after 3 months and 6 months, according to diagnosis. It will be seen that 3 months after treatment patients with endogenous depression had a satisfactory outcome (Categories A and B) far more frequently than those with neurotic depression. At 6 months the discrepancy is of a similar order (77 and 15 per cent. respectively). It will be recalled that these figures do not take account of any marked improvement (A and B) that first appeared some time after the completion of E.C.T., as the change in these cases was regarded as probably spontaneous.

For comparison with the multiple regression analyses, the product moment correlations between diagnosis (excluding doubtful cases) and outcome after 3 and 6 months were calculated. These were 0·62 and 0·61 respectively.

The product moment correlation coefficients among the 35 features are given in Table III. The loadings of the first three factors are given in Table IV. These three factors are significant; the remaining residuals are not significant. For comparison, the correlations with diagnosis and with outcome 6 months after E.C.T. are also given. Table V gives the results of the tests of significance on the residual correlations after the extraction of 0, 1, 2 and 3 factors.

When the first factor loadings are compared with the correlations with diagnosis (Table IV), it is clear that the first factor and diagnosis are closely related. It is also clear that the first factor is related to E.C.T. outcome. The correlations between the first factor loadings and diagnostic correlations is 0·99, and that between loadings and E.C.T. outcome correlations is 0·96. The latter correlation is of particular importance because it indicates that the first factor loadings are closely related to the relatively objective criterion of E.C.T. outcome. The magnitude of these correlations is not to be interpreted as meaning that diagnosis is almost perfect or that the result of treatment can be predicted with virtual certainty. They refer to

Table II

Diagnosis and Outcome

Diagnosis	Outcome after 3 months					Outcome after 6 months					
	A	B	C	D	Total	A	B	C	D	?	Total
Neurotic	5	7	14	37	63	3	5	10	35	10	63
Endogenous	30	14	2	7	53	21	13	2	8	9	53
Doubtful	2	2	1	8	13	0	2	2	7	2	13
Total	37	23	17	52	129	24	20	14	50	21	129

https://doi.org/10.1192/bjp.111.477.659 Published online by Cambridge University Press

TABLE IV

Factor Loadings, Diagnosis and E.C.T. Outcome Correlations
(decimal points omitted)

	Factor Loadings N = 129			Correlations with	
				Diagnosis	E.C.T. Outcome (6 months)
	I	II	III	N = 116	N = 108
1. Age (40+)	209	−458	−167	202	218
2. No precipitants	348	065	−179	393	236
3. Adequate personality	385	−381	051	419	234
4. No adequate psychogenesis	505	−014	−218	534	315
5. Varying depression	572	212	−203	534	261
6. No reactivity	568	301	−187	565	303
7. Distinct quality of depression	420	098	146	425	155
8. Weight loss (7 lb. +)	362	−071	117	433	395
9. Constipation	138	−083	059	270	043
10. Pyknic	186	273	−223	263	249
11. Previous episodes	486	005	−264	459	413
12. Family history endog. dep.	164	294	−182	155	113
13. Early wakening	227	124	−207	271	270
14. Worse a.m.	060	267	−694	143	065
15. Depressive psychomotor activity	736	053	−023	675	440
16. Anxiety	−686	212	−117	−619	−458
17. Delusion of retribution	591	112	441	470	250
18. Nihilistic delusions	673	305	063	657	384
19. Somatic delusions	313	173	184	317	284
20. Paranoid delusions	446	276	410	381	206
21. Suicidal rumination/attempt	225	358	−103	257	139
22. Depressive hallucinations	400	275	394	354	172
23. History over 1 year	−608	000	208	−493	−433
24. Family history—neurosis	−186	326	047	−179	−103
25. Worse p.m.	−276	021	580	−265	−252
26. Blames others	−450	430	219	−439	−395
27. Self pity	−554	081	−278	−486	−274
28. Hopeful outlook	−672	−399	−032	−652	−358
29. Hypochondriacal	−616	152	−041	−483	−497
30. Suicidal through frustration	−394	257	091	−341	−332
31. Irritable	−473	384	−043	−371	−255
32. Phobias	−442	458	−048	−338	−258
33. Hysterical	−354	233	−036	−312	−325
34. Initial insomnia	−308	161	219	−356	−012
35. Delusion of guilt	673	000	132	611	392

associations between correlations and loadings, which are both essentially averages. Thus, the individual variation of patients is largely eliminated from the two correlations obtained; it is this variation which makes both diagnosis and prediction of treatment in the individual case less accurate. The first factor loadings account for 20·8 per cent. of the total variance.

Only six of the second factor loadings are negative. This, together with the nature of the features with these negative loadings, indicates that the second factor is one of general depression, as defined by the sum of 35 features. The correlation coefficient between the second factor loading and the diagnosis correlations is only 0·12, a non-significant value. The second factor accounts for 6·3 per cent. of the total variance. It may be wondered how it is that the second and not the first factor is the general one. The reason for this is that the principal com-

TABLE V
Tests of Significance

No. of Factors Extracted	χ^2	Df	Significance
0	1,189	629	Yes
1	724·7	594	Yes
2	652·7	560	Yes
3	584·0	527	No

ponent method of analysis extracts the factors in order of their contribution to the total variance. In the more commonly used but probably less precise methods of analysis the general factor is extracted first.

The third factor is related to paranoid and retributive delusions and depressive hallucinations, and thus refers to a cluster of psychotic features found in a proportion of depressed patients. It accounts for 5·9 per cent. of the total variance.

If depression were a single disease entity, so-called neurotic and endogenous depressions merely being less or more severe manifestations of depressive illness, one would expect the general factor to account for more of the total variance than any of the bimodal factors. Moreover, the correlations of each feature with diagnosis would be closely similar to the general factor loading of each feature. But in the present analysis the first bipolar factor accounts for three times as much of the total variance as does the general factor. Furthermore, the correlations of each feature with diagnosis are closely similar to the bipolar factor loadings, but not to those for the general factor. Hence the data are inconsistent with the hypothesis that endogenous and neurotic depressions merely describe degrees of severity of depression; they support the view that the disorders differ in kind.

The possibility that the results may have been largely determined by the original selection of the features has to be considered. However, if endogenous and neurotic depressions merely differ in degree, *any* large collection of depressive features should on analysis yield a general factor which would account for more of the total variance than any bipolar factor. If it is possible to choose a wide range of features that produce contrary results, as in the present study, there is clear indication that the two conditions can be differentiated from each other on grounds other than severity of the illness.

RESULTS II

PREDICTION OF DIAGNOSIS AND OUTCOME

It is, therefore, appropriate to analyse the data further to ascertain how this differentiation may best be carried out on the basis of the 35 features. The results of this analysis, i.e. the multiple regression analysis, are given in Table VI. This table also includes the regression analysis of the prediction of E.C.T. outcome at 3 and 6 months. The correlations with diagnosis and E.C.T. outcome at 6 months are not given in Table VI, as these appear in Table IV. The weighting coefficients are similar for E.C.T. at 3 and 6 months.

The ten items which best predicted outcome with E.C.T., and the ten which gave the best prediction of diagnosis were considered separately; their weighting coefficients and the resulting multiple correlations are given in Table VII. These two groups of ten features overlap to some extent and include 18 features altogether. The weights and multiple correlation of the group of 18 features are also given in Table VII. The multiple correlations derived from ten features in Table VII are not much less than the corresponding correlations in Table VI which are derived from all 35 features.

A diagnostic index, constructed by multiplying each of the 18 weighting coefficients (Table VII) by ten and converting to the nearest whole number, was used to calculate a diagnostic score for each patient. The distribution of these scores is given in Table VIII, together with the corresponding theoretical frequencies. The distribution is shown in Figure 1; it will be seen that it is bimodal; it departs from normality very significantly ($\chi^2 > 500$; df = 17). The fact that the distribution is clearly bimodal supports the hypothesis that endogenous and neurotic depression are distinct.

It will be seen from Table VII that the

https://doi.org/10.1192/bjp.111.477.659 Published online by Cambridge University Press

diagnostic weights of the 18 features are not very similar to the E.C.T. outcome weights. In fact, the correlations between the diagnostic weights and E.C.T. outcome weights at 3 and 6 months are 0·61 and 0·62 respectively. This suggests that a better prediction of E.C.T. outcome can be made from the features directly than can be made from the diagnosis. This conclusion is supported by the fact that the correlations between diagnosis and E.C.T. outcome at 3 and 6 months are 0·62 and 0·61 (Table II), whereas the corresponding multiple correlations based on 18 features are 0·72 and 0·74 (Table VII). An E.C.T. prediction index

TABLE VI

Weights and Multiple Correlations

Feature	Diagnosis N = 116		E.C.T. (3 months) N = 129		E.C.T. (6 months) N = 108
	No. of Patients Showing Feature	Weighting Coefficient	Correlation	Weighting Coefficient	Weighting Coefficient
1. Age (40+)	85	−036	248	112	−069
2. No precipitants	38	080	206	−062	058
3. Adequate personality	68	116	235	078	222
4. No adequate psychogenesis	65	108	373	219	187
5. Unvarying depression	56	060	196	−320	−206
6. No reactivity	43	057	262	−154	−236
7. Distinct quality of depression	70	099	194	−050	−124
8. Weight loss (7 lb. +)	58	178	331	390	606
9. Constipation	53	112	050	−198	−090
10. Pyknic	39	115	195	605	609
11. Previous episodes	48	090	399	309	085
12. Family history—endogenous depression	19	018	094	−339	−081
13. Early wakening	78	070	227	434	432
14. Worse a.m.	48	010	101	379	197
15. Depressive psychomotor activity	58	069	367	−413	−253
16. Anxiety	71	−109	−446	−521	−481
17. Delusions of retribution	29	−017	268	−278	−511
18. Nihilistic delusions	36	124	351	098	−027
19. Somatic delusions	15	027	254	377	351
20. Paranoid delusions	33	028	231	481	297
21. Suicidal rumination/attempt	56	091	012	−163	−020
22. Depressive hallucinations	17	115	154	−103	108
23. History over 1 year	38	020	−381	−029	−274
24. Family history—neurosis	22	−061	−163	−083	−006
25. Worse p.m.	30	−059	−285	−408	−439
26. Blames others	25	−103	−367	−204	−179
27. Self pity	57	−048	−306	−444	−247
28. Hopeful patient's attitude	65	−018	−290	107	120
29. Hypochondriacal	59	−007	−494	−741	−693
30. Neurotically suicidal	21	007	−320	−424	−146
31. Irritable	69	−030	−321	−111	−048
32. Phobias	40	015	−294	114	138
33. Hysterical	25	−073	−219	−203	−477
34. Initial insomnia	87	−119	−097	274	514
35. Guilt	+	044	359	091	159
Multiple correlation		·94		·76	·78

+ delusion of guilt, n = 39; guilt feelings, n = 30; guilt free, n = 47

https://doi.org/10.1192/bjp.111.477.659 Published online by Cambridge University Press

TABLE VII

Weights and Multiple Correlations

Feature	Diagnosis Weights		E.C.T. (3 months) Weights		E.C.T. (6 months) Weights	
	10 Features	18 Features	10 Features	18 Features	10 Features	18 Features
3. Adequate personality	134	150		144		189
4. No adequate psychogenesis	155	128		267		151
7. Distinct quality	103	093		−028		−159
8. Weight loss	161	164	429	411	571	594
10. Pyknic		141	447	526	555	601
11. Previous episodes	112	081		212		227
12. Early wakening		076	430	459	402	417
15. Depressive psychomotor activity	180	098		−370		−255
16. Anxiety	−140	−159	−427	−430	−340	−359
18. Nihilistic delusions	249	209		−168		−113
19. Somatic delusions		058	307	326	354	384
20. Paranoid delusions		064	211	346	088	189
25. Worse p.m.		−096	−624	−501	−461	−337
26. Blames others	−124	−093		−394		−333
27. Self pity		−044	−193	−205	−092	−092
29. Hypochondriacal		004	−717	−746	−623	−623
33. Hysterical		−051	−248	−160	−500	−368
35. Guilt	094	081		−054		−012
Multiple correlation	·90	·91	·69	·72	·72	·74

TABLE VIII

Distribution of Diagnosis Scores

Score (× 10)	Obtained Frequency	Theoretical Normal Frequency
14–15	3	0·03
13–14	3	0·09
12–13	5	0·31
11–12	10	0·92
10–11	7	2·36
9–10	12	5·08
8– 9	8	9·00
7– 8	4	14·24
6– 7	3	18·74
5– 6	6	20·91
4– 5	3	19·31
3– 4	9	15·93
2– 3	8	10·86
1– 2	15	6·26
0– 1	12	3·02
(−1)– 0	10	1·30
(−2)–(−1)	6	0·45
(−3)–(−2)	5	0·18
Total	129	128·99

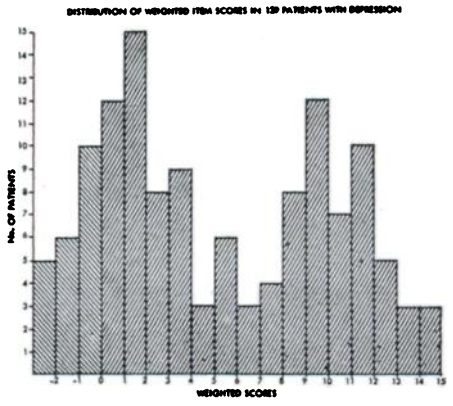

FIG. 1.

was constructed from the 18 appropriate weights (6 months) in the same fashion as the diagnosis index, and used to calculate the E.C.T. predictions score for each patient. The distribution of prediction scores at 6 months is shown in Figure 2. If a dividing line is taken

https://doi.org/10.1192/bjp.111.477.659 Published online by Cambridge University Press

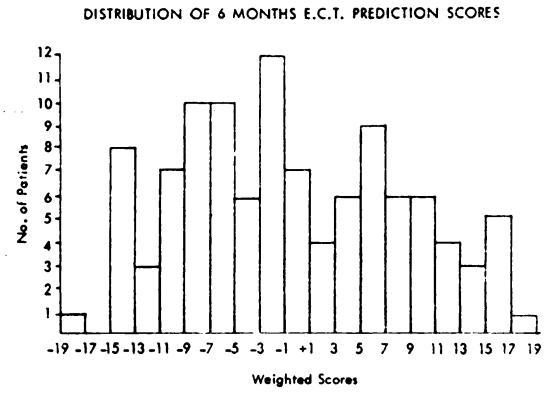

FIG. 2.

at +1, only 13 per cent. of subjects are misclassified as to outcome (taking outcome groups A and B together and C and D together). In applying a scale such as this, it should be remembered that the predictive accuracy of any set of weights relating to prognosis is necessarily diminished by the operation of random social and environmental factors during the follow-up period, a disadvantage to which a diagnostic index is not subject. But if outcome is predicted by diagnosis alone on the assumption that endogenous depression responds favourably to E.C.T. (Table II) the outcome (6 months) of 19 per cent. is wrongly predicted. This figure is based upon those patients who were definitely diagnosed, that is ignoring those with doubtful diagnosis. If the latter were included, then the percentage misclassified would presumably be more than 19.

Ideally, no treatments other than E.C.T. should have been given in the follow-up period. However, the ideal experimental design can rarely be adhered to in clinical practice. Moreover, the majority of the patients in this study were not under the personal care of the investigators, and anti-depressant drugs were prescribed in many cases by the clinician in charge as a standard measure to reduce the chances of relapse after E.C.T. or to treat such a relapse. It is possible that these drugs could have affected the outcome after treatment with E.C.T. and thus might have upset the conclusions of this part of the investigation. To test whether this appeared to be the case, the data given in Table IX were examined (diagnosis being arrived at from Table VII) to ascertain whether drugs given affected the relationship between outcome and diagnosis. The method used was that described by Plackett (1962). This method cannot be used with zero frequencies, and accordingly all the patients who had received M.A.O. inhibitors were grouped together, leaving three drug groups: imipramine series (49), M.A.O. inhibitors (34), and no drugs (20).

The result obtained was $\chi^2 = 1 \cdot 18$, df = 2, which corresponds to a probability of about 0·6. There is, therefore, no evidence that drugs

TABLE IX

Outcome of 103 Cases (five uncertain omitted) Treated with E.C.T. and Assessed at 6 Months Classified by Diagnostic Index and Drugs

	Tofranil Series of Drugs	M.A.O. Inhibitors	Tofranil and M.A.C. Inhibitors	No Anti-depressant Drugs	Totals
Endogenous (A and B)	18	5	0	9	32
Endogenous (C and D)	4	3	1	2	10
Neurotic (A and B)	3	0	2	3	8
Neurotic (C and D)	24	13	10	6	53
Totals	49	21	13	20	103

https://doi.org/10.1192/bjp.111.477.659 Published online by Cambridge University Press

given (or not given) had any effect upon the relationship between diagnosis and outcome after E.C.T. Thus, in the context of the present investigation, the effect of drugs can be ignored in drawing conclusions about the differential effect of E.C.T.

The weights given in Table VII are complex for clinical use. Accordingly Table X has been constructed: this table is derived from the weights (ten features) of Table VII, but the weights have been converted to whole numbers. To use this table the presence of any relevant feature is scored as indicated by the weights given in the table. If a feature is absent it is ignored. For example, if a patient has adequate personality, weight loss, early wakening, blames others and has guilt feelings and/or delusions of guilt, then his diagnostic index score is $+1+2-1+1 = +3$, and his E.C.T. index score is $+3+2 = +5$.

Before any meaning can be attached either to diagnostic or E.C.T. index scores, it is necessary to know the distribution the scores obtained. These distributions are given in Table XI. It is clear from this table that anyone with a diagnostic score of 6 or more is likely to be endogenous, and with a score of 5 or less neurotic. The E.C.T. scores are not so clear-cut, but a score of 1 or more suggests a good result with E.C.T., whereas a score of 0 or less suggests a poor result.

TABLE X
Weights for Deriving Diagnostic and E.C.T. Prediction Indices

Feature	Weights	
	Diagnostic Indices	E.C.T. Indices
3. Adequate personality	+1	
4. No adequate psychogenesis	+2	
7. Distinct quality	+1	
8. Weight loss	+2	+3
10. Pyknic		+3
11. Previous episode	+1	
12. Early wakening		+2
15. Depressive psychomotor activity	+2	
16. Anxiety	−1	−2
18. Nihilistic delusions	+2	
19. Somatic delusions		+2
20. Paranoid delusions		+1
25. Worse p.m.		−3
26. Blame others	−1	
27. Self pity		−1
29. Hypochondriacal		−3
33. Hysterical		−3
35. Guilt	+1	

The correlations of the scores derived from Table X with diagnosis and E.C.T. outcome are 0·89 and 0·67 respectively—very little lower than if the more accurate weights given in Table VII are used. Thus the use of the simplified weights of Table X does not reduce the accuracy of either diagnosis or the prediction of E.C.T. outcome at 6 months to any great extent.

In doubtful cases the more accurate weights based on 18 features (Table VII) may be used. When this is done, a diagnostic score of above 6 is indicative of endogenous and a score of below 6 of neurotic depression. An E.C.T. prediction score (6 months) of above 1 suggests a good response and a score of below 1 a poor response.

Discussion

The results of the factorial analysis are closely similar to those reported by Kiloh and Garside (1963). In their studies, as in our own, the bipolar factor isolated accounted for a greater part of the total variance than the general factor. The correlation of 0·99 they obtained between the bipolar factor loadings (unrotated) and diagnostic correlations is identical with that recorded in this investigation. The cluster of features correlating significantly with a diagnosis of endogenous depression is also closely similar. This measure of agreement was found despite the fact that the two analyses were carried out by different methods. As the present study was with patients for whom admission was considered imperative and who were therefore probably more severely depressed than the out-patients of Kiloh and Garside's investigation, the observations of these authors are shown by the present enquiry to have more general significance.

Certain differences from the results recorded by Kiloh and Garside are perhaps worthy of comment. A significant negative correlation between anxiety and a diagnosis of endogenous

https://doi.org/10.1192/bjp.111.477.659 Published online by Cambridge University Press

TABLE XI

Distribution of Diagnostic and E.C.T. Scores

	Diagnosis			E.C.T. (6 months)	
Score	Endogenous	Neurotic	Doubtful	Good Response (A + B)	Poor Response (C + D)
12	5				
11	7				
10	9			2	
9	11			2	
8	8		1	4	1
7	8		1	1	
6	4	3	2	4	1
5	1	3	1	5	
4		2	3	3	1
3		12	3	5	3
2		12	1	6	1
1		13		3	1
0		12		2	3
−1		6	1	2	6
−2				1	2
−3				1	10
−4				2	6
−5					5
−6					14
−7				1	6
−8					1
−9					2
−10					1
Total ..	53	63	13	44	64

depression was found in the present enquiry, in contrast to the small negative correlation recorded in their studies. This is also reminiscent of the negative correlation between depressed mood and guilt, on the one hand, and agitation and psychic anxiety on the other, in Hamilton's investigation (Hamilton, 1960). These observations may at first sight appear anomalous in that a state of anguished, hand-wringing restlessness is widely thought to be a common feature in, and even pathognomonic of one form of endogenous depression. However, a much less specific form of anxiety with motor unrest is a common component of the neurotic form of depression, which is one of the most common syndromes encountered in psychiatric practice. The two phenomena are not easily differentiated and failure to discriminate between them may have caused the discrepancy to which reference has been made. More detailed clinical analysis and definition may serve to eliminate disagreements of this kind in future studies. Physique was not commented on by Kiloh and Garside, but a relationship between pyknic physique and manic-depressive psychosis has been established by anthropometric studies (Rees, 1944) and the association of a pyknic physique with favourable outcome following treatment with E.C.T. has been demonstrated by other investigators (Hamilton and White, 1960; Roberts, 1959). Neither "no adequate psychogenesis" nor "nihilistic delusions" figured among the 35 items finally chosen for analysis by Kiloh and Garside; both show significant correlations with a diagnosis of endogenous depression and a high loading on factor 1 in this enquiry.

The statistical analyses revealed correlations between features and diagnosis which broadly conformed with the classical descriptions of

endogenous and neurotic depression, and also corresponded at many points with the findings of other workers who have attempted to apply these techniques to the problem of diagnosis in this field (Hamilton and White, 1959; Roberts, 1959; Foulds, 1960). In some respects, however, the findings were perhaps unexpected. Thus, neither family history of endogenous depression nor morning aggravation of the symptoms had appreciable correlations with diagnosis. All the phenomena positively correlated with a diagnosis of endogenous depression also showed positive correlations with a favourable outcome following E.C.T. Of the 35 features listed in Table IV, 31 were highly significantly correlated with diagnosis and 23 with E.C.T. outcome at 6 months (that is, correlation coefficients of 0·245 or more).

In evaluating these features in relation to diagnosis and prognosis, a clear distinction has to be maintained between correlations and multiple regression weights. It will have been noted, for example, that the weights calculated for the prediction of diagnosis and E.C.T. response from the different features do not always appear to be proportionate to the importance generally awarded to these features in clinical practice. This paradox has been explained by Kiloh *et al.*, who point out that whereas estimates of the relative importance of clinical symptoms are normally based on intuitive assessments of their correlation with diagnosis, weights derived from multiple regression depend not only upon correlations with diagnosis but also upon correlations with other symptoms. Thus many ubiquitous symptoms tend to have small or even negative weights if their correlations with other features in the total cluster are as great or greater than their correlations with diagnosis (Kiloh *et al.*, 1962). The reason for this is that the weights derived from the regression analysis reflect the *independent discriminating power* of the feature in question. Thus, in the present study, both depressive psychomotor activity and nihilistic delusions have negative weights for E.C.T. response, although they have positive correlations with it. When it is realized that their contributions to the predicted variance were more than covered by the contributions of other features, the apparent contradiction is resolved.

The findings reported cannot be satisfactorily explained in terms of a single depressive disorder. They lend confirmation to previous observations and in particular those of Kiloh and Garside which could be accounted for only in terms of two relatively separate conditions. As in these latter studies, the second factor loadings indicate that the two disorders have some features in common, but the distinction between them is clearly indicated by the loadings on the bipolar first factor. The possibility has to be considered that, to some extent at any rate, the observations reflect preconceptions about the classification of depressive illness and the halo effects these might have engendered. It would, however, be difficult to explain along these lines the close correspondence between the correlation of features with diagnosis, on the one hand, and their loadings for the bipolar factor on the other. Moreover, findings based on the factorial analysis of symptoms are supported in considerable measure by an independent source of evidence, namely the outcome of treatment with E.C.T. as judged both by social and symptomatological criteria after intervals of 3 and 6 months. It seems likely, therefore, that the dichotomy that emerges from the statistical analysis reflects the existence of a dimension of depressive illness independent of severity, or in other words that the distinction between endogenous and neurotic forms of depression is a qualitative rather than a quantitative one. This would resolve an apparent contradiction to which attention has been drawn by a number of workers (Roth, 1959; Kalinowsky, Hoch and Grant, 1961; Kiloh and Garside, 1963). For if all variation in depression were merely quantitative it would be difficult if not impossible to explain why the usually milder or neurotic varieties showed a far less favourable response to almost every form of treatment than the more severe (endogenous) varieties. If the two conditions are qualitatively distinct, the difference in response to treatment is readily explained.

It is perhaps worthy of note at this stage, however, that there are a number of features in respect of which endogenous and neurotic

depressions resemble one another. This is made plain by the fact that many of the features showing positive loadings on the second or general factor are common to both groups of disorders. There are, in other words, variations in degree as well as variations in kind within the depressive group of conditions. But the dimensions that describe these two types of variation are independent of one another. This is in accord with observations in clinical practice, since of two patients, one suffering from an endogenous and the other a neurotic depressive illness, who are comparable in all other relevant respects, the former will occasionally be the less severely depressed as far as can be judged by the criteria at the disposal of the psychiatrist. In other words, the kind of depression and its severity require to be specified separately for any adequate description of the illness in question.

The high correlation between the first factor loadings and response to E.C.T. accords with the observations of other workers that the outcome of depressive disorders treated with E.C.T. is markedly influenced by diagnosis. That this association is far from perfect, however, is shown by the fact that a better prediction of the outcome of E.C.T. can be made from the features directly than can be made from diagnosis; the correlation between diagnostic weights and E.C.T. outcome at 3 and 6 months were 0·61 and 0·62, whereas the multiple correlations based on 18 features were 0·72 and 0·74 respectively. There was, in fact, rather more overlap in relation to pattern of outcome than to distribution of diagnostic scores, and some patients with neurotic depression achieved very good social recoveries from their illness following treatment. This is hardly surprising in view of the fact that the outcome of treatment must, to some extent, be affected by social and environmental factors to which no form of depressive disorder would be wholly impervious. There is, moreover, a good deal of evidence to indicate that E.C.T. is a non-specific form of treatment, exerting more or less favourable effects in a wide range of conditions of differing aetiology (Roth and Rosie, 1953). In the light of these facts, the high correlation between the diagnostic and E.C.T. outcome weights may be interpreted as further evidence in favour of the relative independence of the two main groups of depressive disorders under consideration, even though the pattern of outcome does not exhibit the same unequivocally bimodal distribution as do diagnostic scores. However, some measure of uncertainty must at times be attached to the diagnosis in an individual case, and it seems justifiable to attempt predictions about the effects of treatment from clinical features directly, rather than from diagnosis alone.

In respect of the diagnostic significance of items relating to guilt, paranoid and nihilistic delusions, agitation and early wakening, the present findings accord with those of Foulds (1960). However, discriminant function analysis in this enquiry shows that these items should have differing weights if the diagnostic index is to be valid; in Foulds's symptom inventory, equal weight was awarded to each feature. In relation to the effects of E.C.T., Hobson's finding that self-reproach and retardation were correlated with a good response to E.C.T., whereas neuroticism, hypochondriasis and hysterical attitude symptoms were associated with a poor one, was upheld. Making use of Hobson's scale, Roberts (1959) reported correct prediction of outcome of E.C.T., at 1 and 3 months after treatment in 78 per cent. of cases, compared with 79 per cent. in Hobson's original enquiry. In the present investigation it has been shown that, with the aid of the index derived from the result, outcome after the longer period of observation of 6 months could be correctly forecast in 87 per cent. of patients, which compares favourably with the results previously recorded. In future investigations better predictions might be achieved if attention were devoted to more detailed definition of clinical features. In retrospect, it is clear to the authors that some of the terms used to describe the features may have been allowed to cover too wide a range of clinical phenomena. For example, a distinction needs to be made between the continual scrutiny and over-interpretation of the bodily accompaniments of anxiety on the one hand and, on the other, the excavation of bizarre, somatic sensations, together with a near-delusional conviction of

disease. "Hypochondriasis" crudely covers both types of case, but the distinction between them may be of some significance. For the present the derivation of scales for diagnosis and prediction of the results of E.C.T. that can be readily applied in the clinic may prove to be of practical value and help to promote consistency and uniformity in diagnostic procedure in different psychiatric centres. If a fairly wide consensus of opinion could be achieved in relation to the most clear and useful lines of demarcation in the depressive disorders, which are the commonest conditions encountered by the modern psychiatrist, progress in defining their aetiological basis and in the search for effective remedies would be hastened.

To this end a quantitative approach to diagnosis needs to be applied in therapy and research. The value of new remedies would be more promptly established if psychiatrists used diagnostic scales that were comparable and capable of translation into quantitative terms. Again, in physiological and biochemical investigations into mental illness clarification of many issues might well be achieved if the cases of depression under scrutiny could be more precisely defined and classified into relatively homogenous groups. It would, for example, be interesting to discover how the two groups defined in the present studies would compare in regard to the changes in the activity of the adrenal cortex described in the studies of Gibbons (1964) among others. It would also be of some interest to compare the two groups for thiopentone-induced delta activity in the EEG, which has been shown by Roth *et al.* (1957) to have a bearing on the prognosis of endogenous depression.

Summary

An investigation of 129 in-patient depressives treated with E.C.T. is presented. All were followed up to 3 months, and 108 to 6 months. Initially all were scored for the presence or absence of 35 features considered to discriminate between endogenous and neurotic depressions. Diagnoses were allocated before or early during treatment. Patients were rated at the termination of E.C.T. at 3 months and at 6 months, on a four-point scale. At 3 months, only 12 out of 63 neurotic depressives were found to have responded well to E.C.T., whereas 44 out of 53 endogenous depressives had done so; the corresponding figures for 6 months were 8 out of 53 and 34 out of 44.

A factor analysis of the features produced three significant factors, a bipolar factor corresponding to the distinction between endogenous and neurotic depression, a general factor with high loadings for many features common to all the depressive cases studied and, probably, a paranoid psychotic factor. The bipolar factor closely resembles that extracted in a previous investigation (Kiloh and Garside, 1963) and supports the hypothesis that these are two distinct depressive populations. Among features with high positive loadings on the first factor, and thus corresponding to a diagnosis of endogenous depression, were adequate premorbid personality, absence of adequate psychogenic factors in relation to illness, a distinct quality to the depression, weight loss, pyknic body build, occurrence of previous depressive episode, early morning awakening, depressive psychomotor activity, nihilistic, somatic and paranoid delusions, and ideas of guilt. Among features with a negative loading corresponding to a diagnosis of neurotic depression were anxiety, aggravation of symptoms in the evening, self-pity, a tendency to blame others, and hysterical features.

By means of multiple regression analysis, three series of 18 weighted coefficients for the differential diagnosis between the two varieties of depression and for the prediction of E.C.T. response at 3 and 6 months were calculated. The multiple correlations between the summed features on the one hand and diagnosis and outcome at 3 and 6 months on the other were 0·91, 0·72 and 0·74 respectively. It was thought that E.C.T. response could be better predicted by the direct use of the weights for E.C.T. response than from the diagnostic weights alone. The weights based on the 18 features were complex, and therefore a further table was constructed giving simplified weights based on ten features of diagnosis and ten features for prediction of outcome of treatment with E.C.T. after 6 months.

ACKNOWLEDGMENTS

We are grateful to the consultant psychiatrists at St. Nicholas' Hospital, Newcastle upon Tyne, and at Newcastle General Hospital, for their generous help and co-operation. We are also grateful to Dr. E. S. Page, Director of the Computing Laboratory, University of Newcastle upon Tyne, Miss E. D. Barroclough, computer operator, and Miss S. O. Allison for their help in processing the data.

REFERENCES

BURT, C. (1952). *Brit. J. Psychol. Stat. Sec.*, 5, 109.
FISHER, R. A. (1941). *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.
FOULDS, G. A. (1960). *J. ment. Sci.*, 106, 1394.
GIBBONS, J. L. (1964). *Arch. Gen. Psychiat.*, 10, 572.
HAMILTON, M. (1960). *J. Neurol. Neurosurg. and Psychiat.*, 23, 56.
—— and WHITE, J. (1959). *J. ment. Sci.*, 105, 985.
—— (1960). *Ibid.*, 106, 1031.
HOBSON, R. F. (1953). *J. Neurol. Neurosurg. and Psychiat.*, 16, 275.
KALINOWSKY, L. B., HOCH, P. H., and GRANT, B. (1961). *Somatic Treatments in Psychiatry*. New York: Grune and Stratton.
KILOH, L. G., BALL, J. R. B., and GARSIDE, R. F. (1962). *Brit. med. J.*, 1, 1225.
—— and GARSIDE, R. F. (1963). *Brit. J. Psychiat.*, 109, 451.
KRETSCHMER, E. (1926). *Körperbau und Charakter*. Berlin: Springer.
PLACKETT, R. L. (1962). *J. roy. Stat. Soc.* B, 24, 162.
REES, W. L. (1944). *Proc. roy. Soc. Med.*, 37, 635.
ROBERTS, J. M. (1959). *J. ment. Sci.*, 105, 693.
ROSE, J. T. (1962). *Brit. J. Psychiat.*, 108, 624.
ROTH, M. (1959). *Canad. psychiat. Ass. J.* (*Suppl.*), 4, 532.
—— and ROSIE, J. (1953). *J. ment. Sci.*, 99, 103.
——, KAY, D. W. K., SHAW, J., and GREEN, J. (1957). *EEG clin. Neurophysiol.*, 9, 225.
SANDIFER, M. G., PETTUS, C., and QUADE, D. (1965). *J. nerv. ment. Dis.* In press.
SNEDECOR, G. W. (1956). *Statistical Methods*. Iowa: Iowa State College Press.
THOMSON, G. H. (1956). *The Factorial Analysis of Human Ability*. London: University of London Press.

M. W. P. Carney, M.D., M.R.C.P.I., D.P.M., *Consultant Psychiatrist, Blackpool and Fylde Hospital Group and Lancaster Moor Hospital; lately Senior Registrar, Department of Psychological Medicine, University of Newcastle upon Tyne*

Martin Roth, M.D., F.R.C.P., D.P.M., *Professor of Psychological Medicine, University of Newcastle upon Tyne*

R. F. Garside, B.Sc., *Senior Lecturer in Applied Psychology, University of Newcastle upon Tyne*

(*Received 24 December, 1964*)