British Journal of Psychiatry (1984), **144**, 581–592

Psychiatric Diagnostic Discriminations with Combinations of Quantitative EEG Variables

CHARLES SHAGASS, RICHARD A. ROEMER, JOHN J. STRAUMANIS and RICHARD C. JOSIASSEN

Summary: The possible psychiatric diagnostic utility of certain quantitative EEG measures was evaluated by further analysis of previously reported data from 242 unmedicated patients and 94 non-patients. Time series of amplitude, frequency and wave symmetry measures for 12-lead EEGs (eyes closed and open) were factor analyzed across leads. Factor scores meeting specified criteria in multivariate analyses were entered into discriminant analyses comparing pairs of the following groups: non-patients, neurotics, personality disorders, overt schizophrenics, latent schizophrenics, major depressives and manics. The following discriminations were obtained with at least 50 per cent sensitivity, and diagnostic confidence rates from 69 to 92 per cent: (a) non-psychotic patients (neuroses, personality disorders) from overt schizophrenics, latent schizophrenics or manics; (b) major depressives from latent schizophrenics or manics; (c) non-patients from schizophrenics (overt and latent), depressives or manics. Most discriminations were replicable in split-half analyses. Possible utility of EEG measures in differential diagnosis is supported.

Quantitative features of the electroencephalogram (EEG) often have been found to differ from normal in 'functional' psychiatric disorders, especially in schizophrenia (Hurst *et al*, 1954; Goldstein *et al*, 1963; Volavka *et al*, 1966; Marjerrison *et al*, 1968; Rodin *et al*, 1968; Itil *et al*, 1972; Lifshitz and Gradijan, 1972, 1974; Giannitrapani and Kayton, 1974; Perris, 1975; Shagass *et al*, 1979, 1982; Étevenon *et al*, 1981; Kemali *et al*, 1981; Perris, 1981).

Recently, we presented the results of a relatively large scale quantitative EEG study, involving 242 unmedicated psychiatric patients and 94 non-patient controls, from whom recordings were made at 12 electrode locations during both eyes closed and eyes open conditions (Shagass *et al*, 1982). The procedure of time series analysis, used by Goldstein *et al* (1963), was applied to measures of amplitude, frequency, and wave symmetry; this procedure is a relatively simple approach to EEG quantification. We described the findings obtained in 43 comparisons of age and sex matched clinical groups, two or three at a time, with respect to these measurements. Our results demonstrated numerous EEG differences between the matched diagnostic groups, many of which confirmed previously reported findings. For example, the EEGs of overt schizophrenics showed less amplitude variability, more frequency variability and less reactivity to eye opening than those of non-patients. The EEG findings in overt schizophrenics were consistent with a higher than normal level of 'resting' activation and were similar in this respect to those obtained in manic and latent schizophrenic patients. In contrast to the results of schizophrenics, the EEGs of patients with personality disorders suggested lower than normal levels of activation.

Given the demonstration of many statistically significant quantitative EEG differences between psychiatric diagnostic groups, the question arises: to what extent could these EEG measurements aid psychiatric diagnosis? We have attempted to answer this question from the EEG data of the same 336 subjects. We present here the results of additional analyses of these data, which were conducted to determine how well several major psychiatric groups could be differentiated from one another by means of the EEG measures.

The strategy that we devised to assess the diagnostic potentialities of our EEG measures required decisions about a series of methodological questions. The following are the approaches that we used to address these questions; the descriptions also indicate how

present data analyses differed from those of our previous report.

(1) Which diagnostic criterion groups should be used? Thirty diagnostic categories or major subgroups were represented in our patient population, several with only one patient. In our previous analyses, we compared age and sex matched groups if each contained at least six subjects. However, we found few EEG differences between subgroups within major categories, e.g. between chronic paranoid and chronic undifferentiated schizophrenics or between different types of neuroses (Shagass *et al*, 1982). Consequently, since larger groups were desirable for present purposes, we decided to use the following relatively coarse diagnostic groupings: non-patient controls, neuroses, personality disorders, overt schizophrenia, latent schizophrenia, major depression, mania.

(2) Having verified that our EEG measures varied with age and sex, how should age and sex differences be handled? Our previous strategy of matching groups for age and sex resulted in omission of a number of subjects from data analysis because they could not be matched. For present purposes, we decided to adjust measures statistically for their relationship to age and sex, so that we could use all subjects in a category for comparisons.

(3) How many diagnostic groups should be compared simultaneously? We decided to compare only two diagnostic groups at a time. This decision was governed by: (a) statistical considerations related to the variable numbers of subjects in our seven groups, and (b) the nature of the usual differential diagnostic process in which laboratory data are brought to bear on the decision between two clinically formulated alternatives.

(4) How should a large number of EEG measures be reduced to a single value for each diagnostic decision? For this purpose, we used the following statistical methods: (a) the time series approach described about 11,000 original data values per subject in terms of 72 means and 72 standard deviations (SD); (b) principal components factor analysis reduced these 144 mean and SD values to 36 factor scores; (c) multivariate analyses were used to select, by specified criteria, a smaller number of factors that differentiated between groups; (d) discriminant function analyis was applied to these selected factors to yield one discriminant score that represented the optimal combination of factor scores for a two-group differentiation.

(5) How should diagnostic worth be assessed? Having reduced our measurements to a single discriminant score for each diagnostic decision, we could compute for each an estimate of sensitivity and specificity (Baldessarini *et al*, 1983).

(6) As the procedures used to develop discriminant scores optimize differentiations and may take advantage of chance, how replicable are the results? To assess replicability, we performed separate discriminant analyses of split-halves of our samples. We then applied these discriminant functions to the split-halves that were not used in each of the two discriminant analyses. In essence, the 'crossover' discriminant scores provided an estimate of the extent to which the diagnostic differentiations yielded by the original scores would be replicated in another population.

The findings provide an indication of the potential assistance that the present EEG measures could offer for several questions of psychiatric differential diagnosis. The word, 'potential', should be underscored, as this study represents an early step in the process of going beyond demonstration of statistically significant EEG differences between clinical groups to assess the possible clinical value of such differences.

Method

Subjects

Technically acceptable EEG data were obtained for 242 unmedicated psychiatric patients and 94 non-patient control subjects. All subjects gave informed consent after the nature of the procedure had been fully explained. Data were acquired while the second edition of the American Psychiatric Association Diagnostic and Statistical Manual (DSM II) system was in use. Of the 242 patients, 212 could be grouped into six global diagnostic categories. The remaining 30 could not be grouped, because of small numbers in each of several diverse categories, such as drug psychosis, geriatric clinic patients and alcoholism. Only the data for the 306 subjects described in Table I were used for group comparisons, but values for all 336 subjects were entered into the factor analyses so as to utilize the maximum range of variance in the data. The diagnostic groups in Table I result from DSM II hospital diagnoses made independently by two senior psychiatrists. In addition, the Feighner *et al* (1972) research diagnostic criteria (RDC) were applied where relevant and the RDC for schizo-affective disorder (Spitzer *et al*, 1978) were applied to patients diagnosed as schizo-affective or acute schizophrenia. Also, in order to relate certain DSM II categories not found in DSM III to the more recent system, DSM III diagnoses were retrospectively applied by chart review to patients classed as acute and latent schizophrenia. All of the acute schizophrenic patients met DSM III criteria for schizophreniform disorder, but none met RDC for schizophrenia or schizo-affective disorder.

Excluding the nine acute schizophrenics, Table I indicates that 154 of 158 patients (97.5 per cent) with diagnoses of neurosis, overt schizophrenia, major depression or mania met RDC criteria to a definite or probable level; 38 RDC diagnoses were probable and 116 were definite. As the few non-RDC subjects had little effect on the results, they were retained in the analyses. The latent schizophrenics were originally diagnosed in accordance with criteria for pseudoneurotic schizophrenia (Hoch and Polatin, 1949), which was classed as latent schizophrenia in DSM II; retrospective application of DSM III criteria to these patients

TABLE I

Description of subject groups

Diagnosis	No.	No. male	Age, Years		No. with RDC diagnosis[a]	Days drug free[b]		Months since first ill	
			Range	Median		Range	Median	Range	Median
Non-patients	94	48	18–66	26	—	—	—	—	—
Neuroses[c]	37	21	19–57	31	36	4-none	13.6	1–264	60
Personality disorders[d]	25	21	17–51	24	9	5-none	8.2	18–240	78
Overt schizophrenics[e]	78	43	17–49	26	66	3—none	8.1	1–296	48
Latent schizophrenics[f]	20	8	17–46	24	—	6-none	7.4	9–216	77
Major depressives[g]	39	14	16–67	39	39	4-none	8.5	2–492	86
Manics	13	5	18–55	25	13	3-none	7.8	5–252	72

[a] RDC indicates Feighner *et al* or Spitzer *et al* Research Diagnostic Criteria, as applicable, definite or probable diagnosis.
[b] None indicates no history of drug use for at least 100 days.
[c] Anxiety (eight patients), depressive (12), obsessive-compulsive (15), anorexia nervosa (one) and depersonalization (one).
[d] Antisocial (seven); passive-aggressive (six), sexual deviation (six), hysterical (three), inadequate (two) and schizoid (one).
[e] Chronic paranoid (22), chronic undifferentiated (30), acute (schizophreniform) (nine), schizoaffective (depressed, eight; excited, four), catatonic (four), simple (one).
[f] Schizotypcal personality (15), borderline personality (five) by DSM III.
[g] Bipolar (19) and primary (29).

resulted in the following diagnoses: schizotypal personality disorder (15 cases); borderline personality disorder (5 cases). Because the EEG measures did not differ significantly between the schizotypal and borderline patients, they were kept together in one group. We attempted to test patients after at least seven days without psychoactive medications; however, 15 of the 212 patients were tested six days and eight patients were tested three to five days after drugs were withdrawn.

Recording procedures

EEGs were recorded with amplifiers set for upper frequency cut-off at 3 KHz and time constant of 0.45 sec in two eight-channel montages. Because of tape channel limitations, only EEGs from the three pairs of symmetrical lateral leads in each montage were stored on analogue tape for subsequent analysis; leads for one montage were: C3X, C4X, T3, T4, 03 and 04; for the other, they were F3X, F4X, T5, T6, 01 and 02. All leads were referenced to linked ears. Certain lead locations deviated from the 10-20 system because somatosensory, visual and auditory evoked potentials were recorded from the same leads following the EEG recording: F3X, F4X, C3X, and C4X were 2 cm posterior and 1 cm lateral to the usual F3, etc., locations, while 03 and 04 were midway between T5 and 01 and T6 and 02, respectively. A 50 uV (peak-to-peak) 10 Hz sine wave calibration signal of about 50 sec duration was also stored on each tape channel.

All recordings took place in a dark room in the same order for each montage, first with eyes closed and then while the subject fixated a central dot in a stationary checkerboard pattern, which was presented on a television monitor screen 101 cm from the subject's eyes. The pattern was composed of two 19 × 19 cm squares, separated by an 0.7 cm vertical dark strip with the fixation point in the center; each square contained 128 black and 128 white checks. The mean intensity of the checkerboard was 1.2 foot-lambert.

During recording, an artifact detector circuit was activated if: (a) the EEG at T3 exceeded for 50 msec the preset limits of 10 uV rms for signals of greater than 50 Hz frequency; (b) the absolute amplitude of the signal exceeded a voltage equivalent to the dynamic range of the analogue to digital converter for more than 50 msec. When these limits were exceeded, an artifact code was placed on the tape code channel. A PDP-12 computer monitored the number of EEG samples meeting artifact rejection criteria, and guided the acquisition of 20 EEG samples of 8 sec duration without artifact code for each montage and condition; if an artifact code occurred before a complete 8 sec sample free of such codes was obtained, the computer restarted the 8 sec sample.

Treatment of data

EEG measurements: Computer-implemented criteria for additional artifact screening before computer quantification have been described in detail previously (Shagass *et al*, 1982). Application of these criteria to EEGs played back from tape resulted in the acceptance of 25 to 153 one sec EEG segments (median, 149.6) from an arbitrary maximum of 153 segments per montage and condition. The final measurements were performed on filtered EEGs, bandpass from 2 to 30 Hz. Sampling rate was 500/sec, and the mean of each of the six EEGs in the 1 sec segment was set to zero. Each data point was then classified as to polarity (+, −) and as to whether its voltage was greater or smaller than that of the preceding data point from the same channel.

Table II outlines the EEG data and measurements, and Figure 1 schematizes the measurements. Measures computed for each 1 sec EEG sample were: mean amplitude, number of zero crosses, and the total duration of the rising phase of the waves. Mean amplitude was determined from the mean of the absolute values of the 500 data points, i.e., disregarding sign. This value was converted to microvolts with reference to an equivalent measurement performed upon the calibration signal. The number of zero crosses was determined from the number of changes in polarity between adjacent data points;

TABLE II
Outline of EEG data and measurements

A. EEG data per subject
1. Number of 1 sec samples after artifact screening—25 to 153 (median, 150)
2. Number of leads—12
3. Number of conditions—2 (eyes closed, eyes open)
Total average EEG—150 × 12 × 2 = 3,600 samples (1 sec)

B. Measures per sample (1 sec)
1. Mean amplitude (mean absolute deviation from zero, uV)
2. Frequency (zero-crosses/2)
3. Wave symmetry (total duration of rising phase, time rising, msec)

C. Variables per condition per lead (150 samples)
1. Mean amplitude
2. SD amplitude
3. Mean frequency
4. SD frequency
5. Mean time rising
6. SD time rising

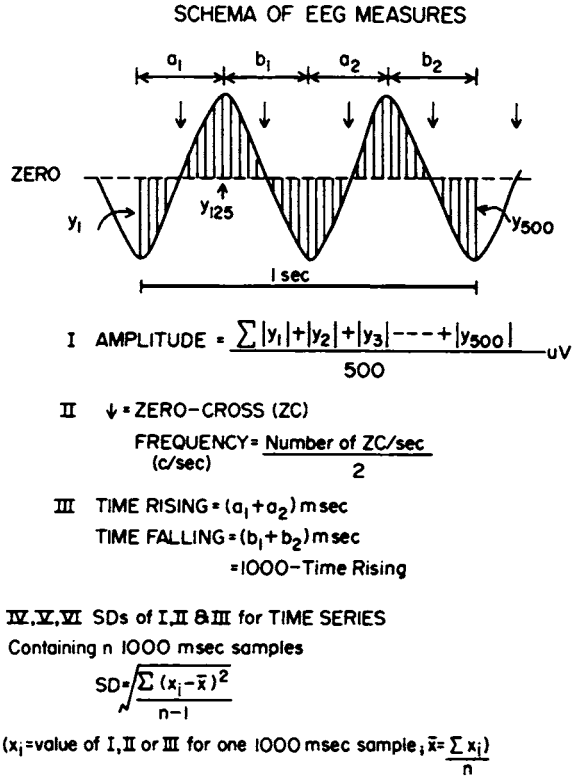

FIG 1.—Schematic description of EEG measurement procedures.

frequency was obtained by dividing zero crosses by two. The time occupied by the rising phase of the waves in the sample was given by the number of data point values that were greater than the one preceding, multiplied by 2 msec (time represented by each data point); 'time rising' values above or below 500 msec would indicate that the positive-going phases of the waves occupied more or less time than their negative-going phases, and that the average EEG wave was not symmetrical. For each measure, the mean and SD were determined for the time series of approximately 150 1 sec samples per lead per condition.

Sex and age: As we wished to use all available subjects, regardless of sex or age, we first ascertained whether or not adjustment for these factors was necessary. Split-plot analyses of variance (ANOVA), with three "between" factors (diagnosis, sex and age) and one repeated measure (leads), were performed for several groups matched for sex and age (Kirk, 1968). The ANOVAs showed at least one main effect for sex and age for each EEG variable, indicating that it was necessary to adjust the EEG measures for sex and age. The validity of using a linear regression adjustment for age was assessed by inspecting scatterplots relating age to measurement, lead by lead; major departures from linearity were not observed.

TABLE III
Outline of steps in data analysis

1. Basic variables—6 measures with eyes closed
   6 difference values (eyes closed minus eyes open)
2. Adjustment for regression upon age and sex—all 12 variables, one lead at a time
3. Adjustment of difference values for regression upon eyes closed values (6 difference values, one lead at a time)
4. Factor analyses across 12 leads (12 factor analyses for 6 eyes closed and 6 adjusted difference values)
5. Factor scores computed
6. Multivariate comparisons of all factor scores between pairs of clinical groups; MANOVA; parallelism (profile) test; conservative ANOVA. 3 sets of comparisons per pair of groups: (a) 15 eyes closed factors; (b) 11 mean difference value factors; (c) 10 SD difference value factors
7. Selection of factor scores for discriminant analyses if: (a) multivariate test $P < 0.10$, and also (b) univariate test $P < 0.05$
8. Two-group discriminant analyses, using selected factors. Discriminant scores computed.
9. Discriminant scores tabulated in approximate decile steps. Sensitivity and specificity determined for decile steps of discriminant scores.
10. Assessment of replicability. Steps 8 and 9 repeated for split-half samples (alternate subjects). "Crossover" discriminant scores computed by applying discriminant functions from one half to the other half.

https://doi.org/10.1192/bjp.144.6.581 Published online by Cambridge University Press

Data analysis steps. Table III outlines the data transformations and the steps in data analysis.

1. The basic data: For each of the recording sites, the basic variables used were the six EEG measures (three means, three SDs) for the eyes closed recordings and the set of six difference values resulting from subtracting the eyes open values from the eyes closed values.

2. Regression adjustment for age and sex: These sets of 12 measures for each lead were each adjusted for their linear regression upon age and sex, using all 336 subjects to estimate the slope and intercept.

3. Regression adjustment of difference values by eyes closed values: Because the eyes closed minus the eyes open difference values were correlated with the eyes closed values, an additional regression adjustment was performed. The difference values were adjusted for their regression upon the corresponding eyes closed values. Thus, the resulting adjusted difference measures were independent of the eyes closed values.

4. Factor analysis: Each of the three adjusted mean and three adjusted SD measures for the eyes closed condition and the corresponding six adjusted difference values (eyes closed values–eyes open) was subjected to factor analysis across the 12 leads using all 336 subjects. That is, for a given measurement (e.g. age and sex adjusted amplitude), the subject values at each of the 12 leads constituted the data for a factor analysis. The factor analysis procedure was a principal component analysis, utilizing the correlation matrix with 1's in the diagnonal, followed by orthogonal varimax rotation to Kaiser's criteria (Cooley and Lohnes, 1971). The resulting factor structure could be described in terms of the relative contribution of each lead to a particular factor by examination of the magnitude of the respective factor loadings. Consequently, the resulting factors for each analysis could be labelled topographically (Shagass *et al*, 1982). The 12 factor analyses yielded a total of 36 such factors.

5. Computation of factor scores: After varimax rotation, each subject's data for a particular measure was transformed into factor scores (Cooley and Lohnes, 1971).

6. Identification of between-group factor score differences: Initial statistical comparisons were performed between all pairs of the seven groups, and also separate groups composed of: (a) those 59 overt schizophrenics who met Feighner RDC; (b) the combined groups of neurotics and personality disorders (non-psychotic patients). The 36 factor scores were subjected to the following tests: multi-variate analysis of variance (MANOVA), multivariate profile analysis (Morrison, 1967), and ANOVA using conservative degrees of freedom (Greenhouse and Geisser, 1959). As our program could not handle all 36 factor scores at one time, and because of limitations in degrees of freedom related to sample size, three separate MANOVAs were done for: (a) the 15 eyes closed factors; (b) the 11 factors for change in mean values; and (c) the 10 factors for change in SD values. Multivariate analysis was used in order to diminish the probability of chance-determined effects for single variables.

7. Selection of factor scores for discriminant analysis: If a two-group MANOVA, profile, or ANOVA, indicated the existence of an overall difference or 'effect' at $P < 0.10$, then those variables yielding a univariate effect at $P < 0.05$ were selected for entry into a subsequent two-group discriminant analysis. At least two such factor scores were required for any discriminant analysis.

8. Discriminant analysis and discriminant scores: For each two-group discriminant analysis, a discriminant score was computed for all subjects in the analysis.

9. Evaluation of sensitivity and specificity: The discriminant scores were tabulated in pre-determined ranges that yielded approximate decile steps, allowing the distributions of discriminant scores for the two groups to be compared. It was then possible to compute, at any decile level, sensitivity (percentage of target group correctly identified), specificity (percentage of non-target group correctly classified), positive predictive value or diagnostic confidence (percentage of those 'test positive' subjects identified as having the 'correct' diagnosis).

10. Evaluation of replicability: To assess replicability of the results, the same factor scores selected for each two-group analysis were entered into two separate discriminant analyses for the two halves of the group, using alternate subjects. The discriminant function for each split-half was used to obtain discriminant scores for that split-half and also for the other split-half. Thus, two discriminant scores were computed for each subject, one derived from the split-half of the sample to which he belonged (same-half) and one from the other split-half to which he did not belong (other-half). The differentiations provided by discriminant scores based on same- and other-half functions could then be evaluated. The differentiations yielded by other-half discriminant scores could also be compared with those obtained by using the discriminant scores based on the entire population (step 9). Since the other-half scores were based on a different group of subjects, they yielded an estimate of discriminability that was uncontaminated by the factors that could have enhanced chance results in the original discriminant analysis.

Results

Factor score differences

A total of 31 sets of multivariate analyses were performed for two-group comparisons: 21 sets comparing pairs of the seven main groups; seven sets comparing these seven groups with the 59 overt schizophrenics who met Feighner RDC; and three sets comparing the combined groups of neurotic and personality disorder patients (designated non-psychotic) with overt schizophrenics, latent schizophrenics, and major depressives. The specific factor score differences will not be presented, as they were in general agreement with previous results for age and sex matched groups (Shagass *et al*, 1982). Their nature can be illustrated by two examples from the comparisons between non-patients and personality disorders and overt schizophrenics. Compared with those of non-patients, the eyes closed factor scores of personality disorder patients reflected higher mean amplitude, lower mean frequency, lower mean time rising (more asymmetry), and higher time rising SD values; the eye opening factor scores of the patients indicated greater changes in mean amplitude and frequency SD. Compared with those of non-patients, the eyes closed factor scores of schizophrenics reflected lower amplitude SD, higher frequency SD, and lower time rising SD, while the eye opening scores of schizophrenics indicated less change in mean and SD of amplitude, mean frequency, and mean time rising. These results were essentially the same as

https://doi.org/10.1192/bjp.144.6.581 Published online by Cambridge University Press

those in our previous report, which led us to interpret the EEGs of personality disorder subjects as less activated and those of schizophrenics as more activated than normal.

For the 21 pairwise comparisons of the first seven groups, a total of 92 variables met selection criteria. Of these, 44 involved comparisons between subjects with personality disorders and those of other groups and 35 resulted from comparing non-patients or neurotics with subjects other than personality disorders. For the three comparisons between overt schizophrenics, latent schizophrenics and manics, a total of only two variables met selection criteria, reflecting the previously observed similarities between these groups. The results for the total group of 78 overt schizophrenics were generally similar to those for the 59 schizophrenics meeting Feighner RDC.

Discriminant analyses were performed only when two or more variables met the selection criteria; thus, these analyses were conducted for only 15 of the 21 possible comparisons between pairs of the seven major groups, and for two of the three comparisons involving the non-psychotic patient group. Some features of the factor scores used in discriminant analyses merit description. (1) Of 101 scores used, 73 were for eyes closed and only 28 for eye opening measures, even though originally there were more eye opening factors (21 of 36). (2) Posterior lead factors contributed 65 per cent of the scores used, the proportion being equal for both eyes closed and eye opening factors. (3) Comparisons involving the overt schizophrenic group contributed 17 of the 28 eye opening factors. These observations suggest that eyes closed recordings from posterior leads could provide about two-thirds of the EEG information relevant to diagnostic differences not involving overt schizophrenics; however, the eye opening data contributed heavily to differences between overt schizophrenics and other groups. A smaller number of eye opening differences between groups was found here than in our previous report (Shagass *et al*, 1982); this appears due to the fact that the measures indicating change with eye opening were rendered independent of the eyes closed values. Apparently, the eyes closed values contained much of the relevant information.

Discriminant analyses, total samples

The statistical significance of the discriminant analyses performed for all subjects of the two-group samples was evaluated by chi square (Cooley and Lohnes, 1971). Of the 15 discriminant functions for the seven major groups, nine attained at least the 0.01 level, three the 0.05 level and three the 0.10 level of significance. Results for comparisons involving the 59 RDC schizophrenics were like those for the 78 overt schizophrenics. Non-psychotic patients were discriminated from both overt and latent schizophrenics ($P < 0.01$).

Discriminant analyses, split-half samples

To assess replicability of the results, discriminant analyses were performed for split-halves of all group pairs, except for those involving the small group of 13 manic patients. The discriminant function obtained for each split-half was used to compute discriminant scores for the subjects of both that half (same) and for those of the other half (other). The split-half analyses for each pair of groups thus yielded four sets of discriminant scores for the two half-samples. The scores were on a normalized scale, with mean of 0.0 and SD of 1.0. Two by two tables were constructed for the distributions of the discriminant scores of each group above and below a value of zero. Chi square tests (1 df, Yates correction) were performed for the two by two tables obtained by adding together the corresponding cells obtained in the two tables from the 'same-half' functions and the two tables obtained from 'other-half' functions. Not counting the comparisons involving RDC schizophrenics, which gave results like those involving the total overt schizophrenia group, there were 13 same-half comparisons, all of which were significant ($P < 0.01$ for 10 and < 0.05 for three).

Since the same-half results represented the addition of two optimal discriminations, whereas neither of the other-half scores was optimal, it would be expected that the differences between groups should be greater with same-half than with other-half scores, and the results generally bore out this expectation. However, 11 of 13 comparisons involving other-half scores did yield differences ($P < 0.01$ for five and < 0.05 for six); only the comparisons between personality disorders and nonpatients or neurotics did not give differences. Thus, the results showed that discriminant functions, representing optimal combinations of EEG variables for one sample, generally yielded similar differentiations when applied to another sample.

Note: Because of space limitations, Tables describing the factors used in discriminant analyses, the results of these analyses and the split-half chi square comparisons are not included here. These can be obtained from the authors upon request.

Sensitivity, specificity and diagnostic confidence

Sensitivity, specificity and diagnostic confidence (positive predictive value) for discriminating one group from the other were assessed at each approximate decile step of the discriminant scores. The diagnostic confidence estimates were made with the assumption that each group comprised half of the total pair sample; although this asumption is subject to criticism (Baldessarini *et al*, 1983), it seemed the one most applicable to distinctions between two groups within the constraints of our data.

Figures 2 and 3 give examples of sensitivity, specificity and diagnostic confidence as a function of discriminant score decile level (computed for the total sample of each pair of groups). Figure 2 plots the results for the 20 latent schizophrenics and 37 neurotic patients. The diagnostic confidence and specificity curves follow the same course, which is opposite to that of the sensitivity curve, i.e., as sensitivity increases, and more target subjects are encompassed in the range of scores, specificity and diagnostic confidence diminish. Discriminant scores of -0.5 or less selected 60 per cent of the latent schizophrenics and only 8.1 per cent of the neurotics (sensitivity, 60 per cent; specificity, 91.9 per cent); diagnostic confidence was 88.1 per cent for equal size groups. Figure 3 shows results for the discrimination between 78 overt schizophrenics and 62 non-psychotic (neurotic, personality disorder) patients. Discriminant scores of -0.5 or less were obtained by 51.3 per cent of the schizophrenics and 11.3 per cent of the non-psychotic subjects; sensitivity, specificity, and diagnostic confidence

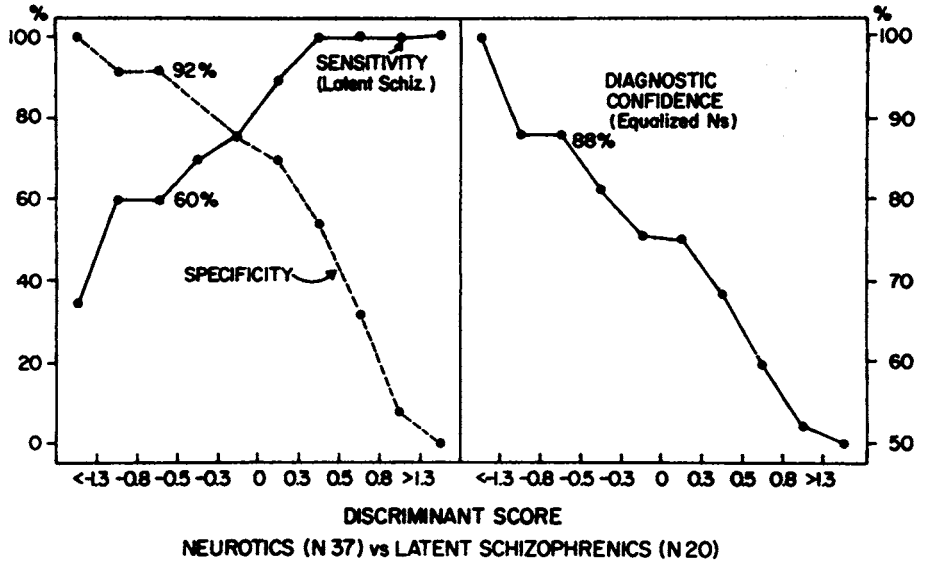

FIG 2.—Sensitivity, specificity and diagnostic confidence percentages for distinguishing latent schizophrenics from neurotics at different approximate decile levels of discriminant scores. Diagnostic confidence calculated by assuming that half of the 57 subjects were in each group.

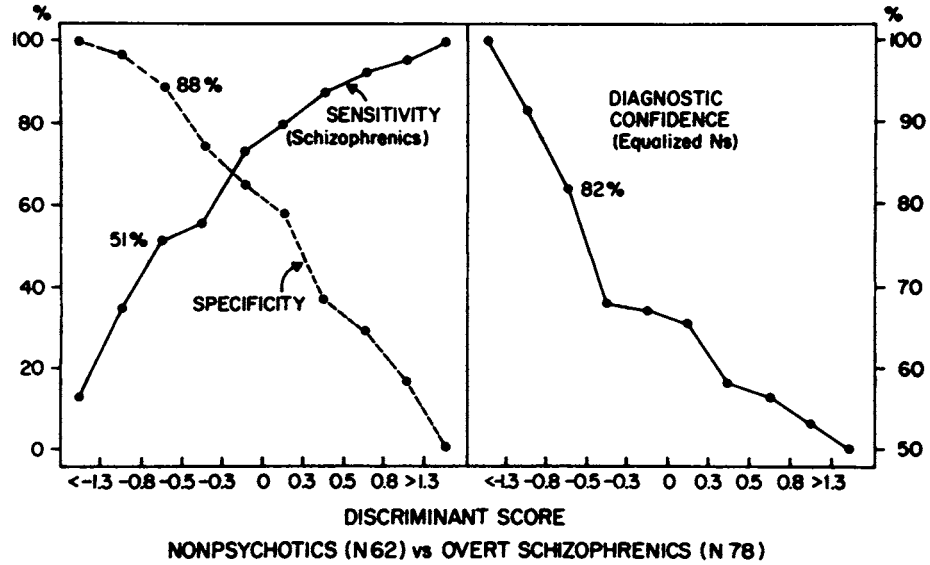

FIG 3.—Sensitivity, specificity and diagnostic confidence percentages for distinguishing overt schizophrenics from non-psychotic patients (neuroses and personality discorders) as a function of discriminant score approximate decile level.

were 51.3 per cent, 88.7 per cent, and 82.0 per cent, respectively.

Table IV gives the sensitivity, specificity, and diagnostic confidence percentages at one discriminant score level for each two-group comparison; these levels were selected from decile distributions of the scores, derived from the discriminant analysis of the total sample, to give at least 50 per cent sensitivity and optimal diagnostic confidence for one of the two groups. Table IV also shows sensitivity, specificity, and diagnostic confidence percentages obtained by applying the same 'cut' decile derived from the total sample to the scores computed by applying the two other-half, or crossover, discriminant functions (manics excluded). The percentages from the total sample indicate approximately the optimal

https://doi.org/10.1192/bjp.144.6.581 Published online by Cambridge University Press

levels of differentiation afforded by the data, while the crossover percentages indicate the levels of differentiation that could be expected from applying the same discriminant functions and 'cut' levels to new samples.

For the total sample, sensitivites for discriminating different patient groups from non-patients ranged from 51.3 per cent for major depressives to 69.2 per cent for manics, with corresponding diagnostic confidence values of 75.1 per cent and 91.6 per cent. Only the neurotic subjects, for whom discriminant functions were not computed because only one variable met selection criteria, could not be differentiated from non-patients. Discriminations between the neurotics and patient groups other than major depressives gave sensitivies ranging from 52.0 per cent to 69.2 per cent, with diagnostic confidence estimates ranging from 69.5 per cent to 88.1 per cent. Similar results were obtained in comparisons between the personality disorder group and patient groups other than neurotics, between the major depressives and latent schizophrenics and manics, and between non-psychotics and overt and latent schizophrenics.

The crossover sensitivity percentages (Table IV) are generally similar to those for the total sample; exceptions to this statement are seen in the two comparisons between personality disorders and nonpatients or neurotics, sensitivities being much lower. Most of the specificity and diagnostic confidence percentages for the crossover discriminant scores are lower than those for the total sample. Whereas 11 of the 13 diagnostic confidence estimates for the total sample exceeded 75 per cent, this was true for only 3 of 13 crossover score estimates. The median difference in diagnostic confidence between the 13 total sample and 13 comparable crossover estimates was 6.9 per cent. However, it should be noted that the crossover estimates in Table IV were based on two discriminant functions, each performed on populations half the size of the total sample. Consequently, these crossover estimates are likely to be lower than those that one might expect from replication populations of the same size as that of the total sample.

Discussion

The present study and our previous report (Shagass *et al*, 1982), although based on the same quantitative EEG data, differed considerably in focus. Having

TABLE IV

Sensitivity, specificity and diagnostic confidence (equalized Ns). Percentages for selected discriminant score levels. Scores from discriminant analysis of total sample (S) compared with scores from crossover split-half discriminant functions (C)

Group[a]	"Cut"[b] Decile	Sensitivity S	Sensitivity C	Specificity S	Specificity C	Diagnostic confidence S	Diagnostic confidence C
A. Comparisons with non-patients (N94)							
Personality disorders(N25)	4	52.0	28.0	81.0	79.8	73.1	58.1
Overt schizophrenics (N78)	4	53.8	50.0	89.4	86.2	83.5	78.3
Latent schizophrenics (N20)	5	65.0	65.0	81.8	64.9	78.2	64.9
Major depressives (N39)	4	51.3	51.3	83.0	81.9	75.1	73.9
Manics (N13)	3	69.2	—	93.6	—	91.6	—
B. Comparisons with neurotics (N37)							
Personality disorders (N25)	4	52.0	16.0	86.5	59.5	79.4	28.3
Overt schizophrenics (N78)	5	61.5	59.0	73.0	73.0	69.5	68.6
Latent schizophrenics (N20)	4	60.0	70.0	91.9	86.5	88.1	83.8
Manics (N13)	4	69.2	—	83.8	—	81.0	—
C. Comparisons with personality disorders (N25)							
Overt schizophrenics (N78)	6	55.1	59.0	84.0	68.0	77.5	64.8
Latent schizophrenics (N20)	4	50.0	55.0	92.0	80.0	81.2	73.3
Major depressives (N39)	5	61.5	64.1	80.0	84.0	75.5	80.0
Manics (N13)	5	84.6	—	88.0	—	87.6	—
D. Comparisons with major depressives (N39)							
Latent schizophrenics (N20)	4	50.0	40.0	89.7	82.1	83.0	69.0
Manics (N13)	5	69.2	—	76.9	—	75.0	—
E. Comparisons with non-psychotics (N62)							
Overt schizophrenics (N78)	4	51.3	52.6	88.7	80.6	82.0	73.1
Latent schizophrenics (N20)	4	70.0	55.0	79.0	79.0	77.0	72.4

[a] Group for which sensitivity was calculated.
[b] Discriminant score decile below which incidence of specified group gives sensitivity.

https://doi.org/10.1192/bjp.144.6.581 Published online by Cambridge University Press

already demonstrated numerous EEG differences between clinical groups, we sought here to evaluate the extent to which optimal combinations of the EEG measurements could provide diagnostic information. Thus, the emphasis here was not so much on whether groups differed, but on the magnitude of the discriminations afforded by combining the differing EEG measures, and on the replicability of the differentiations. The results indicate that our EEG measurements should be able to contribute information relevant to certain diagnostic decisions (Table IV). The positive nature of these results also suggests that the strategy for data analysis used here, which is possibly unique for psychiatric EEG studies, may have general applicability for investigations involving a large number of physiological measurements.

The findings indicated that the following clinical groups could be differentiated from one another by the EEG measures: (a) non-patients from patients with overt schizophrenia, latent schizophrenia, major depression, or mania; (b) patients with neuroses or personality disorders from those with overt schizophrenia, latent schizophrenia or mania; (c) patients with major depressions from those with latent schizophrenia or mania. Split-half analyses, performed for all comparisons except those involving the small manic group, showed that the foregoing discriminations were reasonably replicable (Table IV). Several methodological issues, pertinent to the validity of these findings, should be considered before discussion of their diagnostic implications.

Methodological issues

Chance effects: In embarking upon determination of the degree of diagnostic differentiation that could be obtained from optimal combinations of variables selected from many measurements, the issue of possible chance effects was a major concern. A point to be noted first in considering this issue is that there is little reason to do discriminant analysis unless the variables entered into the analysis have already been found to give differences (Overall and Klett, 1972). Second, although we selected differing single variables, we stipulated that they must come from multivariate analyses that achieved a minimal statistical criterion; this placed some restriction upon taking advantage of chance variations. Third, and most important, we performed split-half analyses with crossover application of discriminant functions to subjects not involved in computing the functions; the fact that these analyses demonstrated replicability for most of the discriminations is reassuring. The evidence supports the expectation that future studies should yield similar results. Furthermore, because our replication data involved half samples, replicability should be better with samples of equal or greater size than the present total samples.

Rationale for two-group comparisons: Our rationale for comparing two groups at a time, rather than three or more groups, should be made explicit. We did perform pilot analyses with three or more groups. However, this meant that we had to select variables for entry into discriminant analysis on the basis of the outcome of simultaneous comparisons between several groups. Thus, some variables could meet the selection criteria because they differed strongly only between two of the larger groups and not between others, while other variables would not be selected because they differed only between two of the smaller groups. This biased the selection of variables toward those that differentiated the larger groups from one another, with the consequence that all but the first one or two discriminant functions were difficult to interpret. The interpretive problem was avoided by restricting the comparisons to two groups; each comparison could yield only one discriminant score, for which sensitivity and specificity could be evaluated. Also, the final product of the two-group comparisons, i.e., a series of probability statements concerning the subject's membership in one of two possible groups seemed compatible with the clinical diagnostic process. A common way of using laboratory test data is to provide support for one of two diagnostic alternatives suggested by clinical observations. When several tests are used, they can bear on more than one set of alternatives. Our two-group discriminant functions can be regarded as a series of tests, each of which provides information bearing on a specific decision between two possible clinically formulated diagnoses.

Limitations of EEG measures: It should be noted that our EEG measures, which were selected for computational simplicity, largely ignore EEG frequency details. There is reason to believe that frequency information, such as provided by power spectral analysis of the EEG, may provide psychiatric correlates not available with our measures (Volavka *et al*, 1981). Present positive results with relatively simple EEG quantitative methods suggest the probability that more complex techniques could be even more productive of psychiatrically relevant information.

Comparability of diagnostic categories: The diagnostic categories used for present purposes can give rise to two kinds of problems. First, by grouping several, possibly different kinds of patients into global categories, such as overt schizophrenia, error variance could be increased and important differences between subtypes could be obscured. We can say only that we previously attempted to determine EEG differences between subgroups within major categories, such as overt schizophrenia or the neuroses, with little success

(Shagass *et al*, 1982). Consequently, since larger sample sizes were desirable, we combined the subtypes for present purposes. Present results show that our global categories were discriminable by EEG measures. However, it is quite possible that better EEG discriminations could be obtained by comparing larger samples of subtyped populations, or by classifying patients by other criteria, such as symptom patterns, rather than by diagnosis. Certainly, since a substantial portion of each clinical group was not discriminated by EEG, it would be useful if one could identify the clinical characteristics that distinguish these patients from those who were discriminated.

The second diagnostic classification problem in present data results from the fact that we tested our patients during the era of the DSM II system, and we are reporting the data during the DSM III era. Being aware of problems of diagnostic reliability, we demanded independent agreement between two senior psychiatrists and also applied the Feighner *et al* (1972) RDC to most of our patients. It seems relevant to note that the Feighner RDC are conservative and come very close to DSM III criteria for the categories encompassed by both systems (Helzer *et al*, 1981); with very few exceptions, our patients met the available RDC. Furthermore, we have also applied Spitzer *et al* (1978) and DSM III criteria to certain DSM II classifications, and described our patients in these terms. The results obtained by applying these various diagnostic criteria suggest that psychiatrists using DSM III criteria, providing they classed borderline and schizotypal personality disorders as latent schizophrenia, would assign nearly all of our subjects to the same global categories in which we placed them. Consequently, the criterion variables used here should be replicable elsewhere.

Diagnostic implications

Although no combination of EEG measures differentiated completely between any two groups, sensitivity rates in the total samples (Table IV) of at least 50 per cent (median, 60 per cent) were associated with specificity rates ranging from 73 per cent to 94 per cent (median, 84 per cent) and adjusted diagnostic confidence levels ranging from 69 per cent to 92 per cent (median, 81 per cent). The comparable diagnostic confidence levels for crossover discriminant scores were lower, but still generally respectable, considering the fact that they were attenuated by being based on half samples.

Given further verification of the discriminations found in this study, to what extent could they provide information of practical value? Such test information would probably find application mainly for diagnostic differentiations that are clinically difficult and encountered with some frequency. From the standpoint of diagnostic difficulty and relative frequency of the problem, the results involving latent schizophrenics may be of special interest. Pope *et al* (1983) have drawn attention to the difficulty involved in differentiating between borderline and other personality disorders. Half of our latent schizophrenic group (50 per cent sensitivity), which consisted of patients with borderline and schizotypal personality disorders, was discriminable by EEG from other personality disorders with 81.2 per cent diagnostic confidence (Table IV). Our EEG measurements did not differ between borderline and schizotypal patients, which is in agreement with the clinical observations of Gunderson *et al* (1983). Latent schizophrenics often are also difficult to differentiate from neurotics with severe anxiety, depressive and compulsive disorders like those in our sample; the relatively good sensitivity and specificity achieved by our EEG measures for this discrimination (Figure 2, Table IV) seems noteworthy. Clearly, present results with a non-homogeneous latent schizophrenic sample of only 20 patients require replication; however, it is encouraging that the findings obtained by comparing the latent group with personality disorders and neurotics were fairly well replicated in split-half samples (Table IV).

Clinical discrimination of overt schizophrenic from non-psychotic patients is generally a lesser problem than distinction between latent schizophrenics and nonpsychotics, but there are many instances in which objective EEG test information would be contributory. Present data suggest that about half of overt schizophrenics can be distinguished from non-psychotics with diagnostic confidence somewhere between 73 per cent and 82 per cent (Table II, Figure 3). The finding that 69 to 85 per cent of manics were discriminated from non-schizophrenic subjects at confidence levels ranging from 75 to 92 per cent (Table IV) may be of greater scientific than clinical interest. However, under some circumstances, a means for distinguishing bewtween hypomanic patients and those with conduct disorders could be quite contributory.

Unfortunately, the data failed to reveal differences that could aid three very important clinical distinctions, those between manics and overt schizophrenics, major depressives and overt schizophrenics, and between major depressives and neurotics.

The discriminations between non-patients and overt schizophrenics and major depressives were both substantial and replicable (Table IV). Manics also differed from normal. These findings provide strong evidence of functional CNS deviations from normal in these disorders. However, from a practical clinical standpoint, the need to make a differential diagnosis

between normality and the three major psychoses is probably infrequent and, consequently, of limited utility. The distinction between latent schizophrenics and normals could find more frequent application, but our split-half replication results suggest caution in accepting the validity of this discrimination.

Issues for practical application

The thrust of this study was to go beyond demonstration of statistical significance to evaluate the possible clinical diagnostic significance of present EEG measurements. In our view, the generally positive results are encouraging, but additional replication of the findings is essential before practical application should be attempted. However, two issues, related to possible future application of EEG measurements to psychiatric diagnosis, seem worthy of consideration at this stage of development: (a) the problems introduced by psychiatric medications; (b) possibilities for simplifying the EEG measurements.

Effects of medication: Given replicable findings, their practical application to diagnosis will depend upon showing that the discriminating measures are relatively insensitive to commonly used psychoactive drugs. In the usual clinical situation, it is unlikely that patients can be kept unmedicated for about one week, as was the case with our subjects. We now have evidence, from retests of 21 initially unmedicated patients while they were receiving antipsychotic drugs, that these medications exert almost no effects on those EEG factor scores selected for present discriminant analyses. On the other hand, retests of 7 patients while receiving tricyclic antidepressants revealed a number of effects on the factor scores. Our antipsychotic drug results, but not the findings with antidepressants, agree with other evidence that the effects of these agents on the quantitative EEG may be relatively minor (Matousek *et al*, 1981; Perris *et al*, 1981). More information is required concerning drug effects on diagnostically discriminating EEG measures, but the situation seems reassuring with respect to antipsychotic agents.

Simplification of procedures: Ideally, one would like to obtain useful diagnostic information from EEGs recorded from very few electrodes under a single condition and with the use of minimal computational facilities. While present data suggest that much of the EEG information required for the diagnostic discriminations could come from the eyes closed condition, using fewer electrodes, it would be premature to accept this possibility. We intend to evaluate empirically the actual loss of discriminative power produced by such simplification. It may also be possible that simplification of procedures, with attendent loss of information, is not really cost-beneficial under modern technological conditions, and that effort directed toward refining the procedures may be more productive. The recording of at least 16 channels of EEG is now commonplace. The cost of computer equipment is constantly becoming less. Even though the development of present discriminant scores required a complex series of statistical manipulations, the generation of individual scores, and their interpretation as probability statements in two-group discriminations, is a relatively straightforward programming task. Given a set of empirically validated equations, they could be put to use without inordinate difficulty or expense.

Acknowledgements

Research supported in part by U.S. Public Health Service Grant MH12507. We thank I. Chung Hung, Jeffrey Kline, Michael Kormanik, Ann W. McGrath, William Nixon and Stephen Slepner for assistance.

References

BALDESSARINI, R. J., FINKLESTEIN, S. & ARANA, G. W. (1983) The predictive power of diagnostic tests and the effects of prevalence of illness. *Archives of General Psychiatry*, **40**, 569–73.

COOLEY, W. W. & LOHNES, P. R. (1971) *Multivariate Data Analysis*. New York: Wiley.

ÉTEVENON, P., PERON-MAGNAN, P., RIOUX, P., PIDOUX, B., BISSERBE, C., VERDEAUX, G. & DENIKER, P. (1981) Schizophrenia assessed by computerised EEG. *Advances in Biological Psychiatry*, **6**, 29–34.

FEIGHNER, J. P., ROBINS, E., GUZE, S. B., WOODRUFF, R. A., WINOKUR, G. & MUNOZ, R. (1972) Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry*, **26**, 57–63.

GIANNITRAPANI, D. & KAYTON, L. (1974) Schizophrenia and EEG spectral analysis. *Electroencephalography and Clinical Neurophysiology*, **36**, 377–86.

GOLDSTEIN, L., MURPHREE, H. B., SUGERMAN, A. A., PFEIFFER, C. C. & JENNEY, É. H. (1963) Quantitative electroencephalographic analysis of naturally occurring (schizophrenic) and drug-induced psychotic states in human males. *Clinical and Pharmacologic Therapies*, **4**, 10–21.

GREENHOUSE, S. W. & GEISSER, S. (1959) On methods in the analysis of profile data. *Psychometrika*, **24**, 95–112.

GUNDERSON, J. G., SIEVER, L. J. & SPAULDING, E. (1983) The search for a schizotype — Crossing the border again. *Archives of General Psychiatry*, **40**, 15–22.

HELZER, J. E., BROCKINGTON, I. F. & KENDELL, R. E. (1981) Predictive validity of DSM III and Feighner definitions of schizophrenia. *Archives of General Psychiatry*, **38**, 791–797.

HOCH, P. & POLATIN, P. (1949) Pseudoneurotic forms of schizophrenia. *Psychiatric Quarterly*, **23**, 248–76.

HURST, L. A., MUNDY-CASTLE, A. C. & BEERSTECHER, D. M. (1954) The electroencephalogram in manic-depressive psychosis. *Journal of Mental Science*, **100**, 220–40.

ITIL, T. M., SALETU, B. & DAVIS, S. (1972) EEG findings in chronic schizophrenics based on digital computer period analysis and analog power spectra. *Biological Psychiatry*, **5**, 1–13.

KEMALI, D., VACCA, L., MARCIANO, F., NOLFE, G. & IORIO, G. (1981) CEEG findings in schizophrenics, depressives, obsessives, heroin addicts and normals. *Advances in Biological Psychiatry*, **6**, 17–28.

KIRK, R. E. (1968) *Experimental Design: Procedures for the Behavioral Sciences*. Belmont, Massachusetts: Brooks/Cole.

LIFSHITZ, K. & GRADIJAN, J. (1972) Relationships between measures of the coefficient of variation of the mean absolute EEG voltage and spectral intensities in schizophrenic and control subjects. *Biological Psychiatry*, **5**, 149–163.

—— & —— (1974) Spectral evaluation of the electroencephalogram: Power and variability in chronic schizophrenics and control subjects. *Psychophysiology*, **11**, 479–90.

MARJERRISON, G., KRAUSE, A. E. & KEOGH, R. P. (1968) Variability of the EEG in schizophrenia: Quantitative analysis with a modulus voltage integrator. *Electroencephalography and Clinical Neurophysiology*, **24**, 35–41.

MATOUSEK, M., CAPONE, C. & OKAWA, M. (1981) Measurement of the inter-hemispheral differences as a diagnostic tool in psychiatry. *Advances in Biological Psychiatry*, **6**, 76–80.

MORRISON, D. F. (1967) *Multivariate Statistical Methods*. New York: McGraw Hill.

OVERALL, J. E. & KLETT, C. J. (1972) *Applied Multivariate Analysis*. New York: McGraw Hill.

PERRIS, C. (1975) EEG techiques in the measurement of the severity of depressive syndromes. *Neuropsychobiology*, **1**, 16–25.

—— VON KNORRING, G. L., CUMBERBATCH, J. & MARCIANO, F. (1981) Further studies of depressed patients by means of computerized EEG. *Advances in Biological Psychiatry*, **6**, 41–9.

POPE, H. G., JONAS, J. M., HUDSON, J. I., COHEN, B. M. & GUNDERSON, J. G. (1983) The validity of DSM III borderline personality disorder. A phenomenologic, family history, treatment response, and long-term follow-up study. *Archives of General Psychiatry*, **50**, 23–30.

RODIN, E., GRISELL, J. & GOTTLIEB, J. (1968) Some electrographic differences between chronic schizophrenic patients and normal subjects. In *Recent Advances in Biological Psychiatry*, 194–204, Vol. 10 (ed J. Wortis). New York: Plenum Press.

SHAGASS, C., STRUMANIS, J. J. & OVERTON, D. A. (1979) Correlations between psychiatric diagnosis and some quantitative EEG variables. *Neuropsychobiology* **5**, 16–26.

—— ROEMER, R. A. & STRAUMANIS, J. J. (1982) Relationships between psychiatric diagnosis and some quantitative EEG variables. *Archives of General Psychiatry*, **39**, 1423–35.

SPITZER, R. L., ENDICOTT, J. & ROBINS, E. (1978) Research diagnostic criteria: Rationale and reliability. *Archives of General Psychiatry*, **36**, 773–82.

VOLAVKA, J., MATOUSEK, M. & ROUBICEK, J. (1966) EEG frequency analysis in schizophrenia, *Acta Psychiatrica Scandinavica*, **42**, 237–45.

—— ABRAMS, R., TAYLOR, M. A. & REKER, D. (1981) Hemispheric lateralization of fast EEG activity in schizophrenia and endogenous depression. *Advances in Biological Psychiatry*, **6**, 72–5.

* Charles Shagass, M.D., *Professsor of Psychiatry*

Richard A. Roemer, Ph.D., *Associate Professor of Psychiatry*

John J. Straumanis, M.D., *Professor of Psychiatry*

Richard C. Josiassen, Ph.D., *Research Assistant Professor of Psychiatry*

Temple University Medical School and Philadelphia Psychiatric Center, Philadelphia, Pennsylvania 19131, USA

* Correspondence.

(Received 10 October; revised 21 November 1983)