A STATISTICAL APPROACH TO THE DIAGNOSIS OF LIVER DISEASE ON THE BASIS OF SERUM BILIRUBIN AND ENZYME LEVELS

S. A. FELLINGHAM, *National Research Institute for Mathematical Sciences*, AND R. C. P. M. MEKEL, *National Nutrition Research Institute, Pretoria*

The clinician is often faced with the problem of coming to a definite diagnosis on the basis of a large variety of clinical and other observations. Any one finding taken on its own might point to a variety of conditions, and the problem is to identify a pattern in the observations made which is unique to a particular disease or condition. In practice it may happen that, where a large number of observations are made, the clinician tends to become confused and to base his diagnosis on the few which he considers to be most important. Such a selection is subjective, and the particular choice of one clinician may or may not be the choice of others. The choice may furthermore be biased, if the clinician has had to do only with a certain class of patient or type of disorder.

This problem is by no means unique to the clinician. A similar problem faces the nutritionist, who, after he has made a large number of biochemical, clinical and other observations on a population, must find some objective means of interpreting his observations in terms of nutrition status. Likewise, the taxonomist must endeavour to classify plants or other organisms on the basis of a variety of observations and the applied psychologist to classify prospective employees as suitable or unsuitable on the basis of a battery of tests.

It is clear that some kind of technique is needed which will deal simultaneously with a variety of observations and enable class allocations to be made with a minimum risk of error. The statistical technique known as a discriminatory analysis has been devised to deal with this type of situation. It was first introduced into statistics by Fisher[1] in 1936 to deal with taxonomical problems and has since been applied in a variety of fields, though never, as far as we know, to the problem of clinical diagnosis. The present paper describes the application of this technique to a clinical diagnostic problem.

THE EXPERIMENT

A total of 194 patients suffering from a variety of liver conditions, who were admitted to the Pretoria General Hospital, were studied over a period of 15 months. A specific diagnosis was made in each case on the basis of a liver biopsy.

Total and conjugated serum bilirubin were measured in each patient together with the serum activity of 8 different enzymes. Two of the enzymes determined, viz. glucose-6-phosphatase and glutamic lactic dehydrogenase (GLDH), had to be discarded for the purpose of the analysis because the observations for these variables were incomplete. The variables used were: (1) total bilirubin, (2) conjugated bilirubin, (3) aldolase, (4) isocitric dehydrogenase, (5) lactic dehydrogenase (LDH), (6) glutamic oxalacetate transaminase (GOT), (7) glutamate pyruvate transaminase (GPT), and (8) alkaline phosphatase.

No observations were made on a control group of normal persons and this, to some extent, detracts from the value of the analysis which was carried out.

For the purpose of the analysis, the patients were grouped into 5 groups or categories of liver disease (Table I). Five groups were chosen, since this was the maximum number that could be handled on the electronic computer by the available programme.

TABLE I. CLASSIFICATION OF LIVER DISEASES INTO GROUPS

Group	*Liver disease*
1	Hepatitis and necrosis
2	Siderosis, hepatofibrosis, cirrhosis and periportal infiltration
3	Hepatoma
4	Pneumonia with jaundice
5	Metastasis, liver abscess, amoebic abscess, tuberculosis, bilharzia, typhoid, sarcoidosis, amyloidosis, extrahepatic obstruction and cases with no histological abnormality

The means and 95% confidence limits of the values for each variable in the different groups are given in Table II. The calculation of confidence limits implies the assumption that the values have a normal distribution. It is unlikely that the pre-

TABLE II. MEANS AND 95% CONFIDENCE LIMITS* (BETWEEN BRACKETS) OF SERUM CONSTITUENT VALUES FOR THE FIVE GROUPS OF LIVER DISEASE

Serum constituents	*Groups*				
	1	*2*	*3*	*4*	*5*
1 (Tot. bil.)	11·7 (0–25·4)	0·8 (0–1·8)	1·7 (0–6·0)	6·1 (0–15·6)	3·6 (0–15·0)
2 (Conj. bil.)	9·7 (0–20·8)	0·4 (0–1·1)	1·2 (0–5·0)	5·3 (0–14·6)	2·8 (0–12·5)
3 (Aldolase)	33·6 (0–87·0)	14·2 (6·4–21·9)	20·9 (0–46·8)	22·0 (3·3–40·7)	16·4 (0–35·8)
4 (Isoc. DH.)	922 (0–2,770)	169 (22–317)	847 (0–3,330)	265 (0–594)	379 (0–1,430)
5 (LDH)	699 (0–1,200)	357 (104–610)	710 (0–1,780)	604 (0–1,310)	575 (0–1,530)
6 (GOT)	634 (0–1,900)	45 (0–97)	168 (0–466)	117 (0–357)	87 (0–295)
7 (GPT)	748 (0–2,460)	32 (0–69)	68 (0–194)	64 (0–154)	72 (0–325)
8 (Alk. phos.)	5·7 (0–15·0)	2·9 (0·1–5·6)	9·5 (0–22·6)	1·7 (0·2–3·2)	8·5 (0–28·6)

* Since the 95% limits are based on the assumption that the values are normally distributed, the calculated value for the lower limit was sometimes found to be below zero. Such lower limits have been tabulated as 0.

sent values were normally distributed, but the confidence limits are nevertheless given in order to demonstrate the degree of scatter of the values. It can be seen that there is considerable overlap in the confidence limits found for any one variable in the different groups, and it is clear that in most cases very little confidence could be placed in a diagnosis based on a single variable.

We shall seek to determine by means of a discriminatory analysis whether distinct patterns exist in the different groups in respect of the observed variables, such that we could allocate a patient to his particular group solely on the basis of the variables measured. The statistical techniques used to test for the existence of patterns are rather involved, and the calculations are best carried out on an electronic computer. The application of the results to individual cases is not too difficult, however, and well within the reach of anyone armed with a desk calculating machine.

The rationale of the statistical technique applied to our problem is briefly as follows:

For every possible pair of groups (Table I) we seek a discriminant function of the type

$$X = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7 + a_8x_8 \quad \ldots\ldots (1)$$

where x_1, ... x_8 are the serum constituent values, numbered 1 to 8 in the order listed above, and a_1, ... a_8 are constants.

The constants, $a_1, \ldots a_8$ are so determined that the resulting values for X will differ maximally for each pair of groups, i.e. the ratio of the distance between the two groups to the variance within them will be maximal.[2-4]

We shall use the notation X_{ik} to denote the discriminant function between groups i and k (X_{13}, for example, would denote the function for comparing groups 1 and 3). A set of functions X_i (i=1, . . . 5) typifying group i can be found for each group. The coefficients of X_{ik} are then obtained from X_i and X_k by taking the difference between the equivalent coefficients of X_i and X_k. Thus:

$$X_{ik} = \Sigma\,(a_{ij} - a_{kj})\,x_j \quad \ldots\ldots (2)$$

where Σ directs us to sum over-all values for j = 1, ... 8.

The analysis was applied to the data and the following values for X_i (i=1, . . . 5) were obtained:

$$\left.\begin{aligned}
X_1 &= -\ 7{\cdot}7863 + 1{\cdot}0126x_1 - 0{\cdot}6539x_2 + 0{\cdot}0904x_3 \\
&\quad - 0{\cdot}0006x_4 + 0{\cdot}0015x_5 + 0{\cdot}0046x_6 + 0{\cdot}0047x_7 \\
&\quad + 0{\cdot}0235x_8 \\
X_2 &= -\ 1{\cdot}0079 + 0{\cdot}0923x_1 - 0{\cdot}1994x_2 + 0{\cdot}0899x_3 \\
&\quad - 0{\cdot}0007x_4 + 0{\cdot}0020x_5 - 0{\cdot}0004x_6 - 0{\cdot}0006x_7 \\
&\quad + 0{\cdot}0704x_8 \\
X_3 &= -\ 3{\cdot}1671 - 0{\cdot}0111x_1 - 0{\cdot}1914x_2 + 0{\cdot}1052x_3 \\
&\quad - 0{\cdot}0003x_4 + 0{\cdot}0033x_5 + 0{\cdot}0015x_6 - 0{\cdot}0014x_7 \\
&\quad + 0{\cdot}2280x_8 \\
X_4 &= -\ 2{\cdot}9427 - 0{\cdot}4383x_1 + 0{\cdot}8592x_2 + 0{\cdot}1311x_3 \\
&\quad - 0{\cdot}0009x_4 + 0{\cdot}0029x_5 - 0{\cdot}0020x_6 - 0{\cdot}0014x_7 \\
&\quad - 0{\cdot}0365x_8 \\
X_5 &= -\ 2{\cdot}2052 + 0{\cdot}0753x_1 - 0{\cdot}0683x_2 + 0{\cdot}0803x_3 \\
&\quad - 0{\cdot}0010x_4 + 0{\cdot}0031x_5 - 0{\cdot}0004x_6 - 0{\cdot}0003x_7 \\
&\quad + 0{\cdot}1958x_8
\end{aligned}\right\} \quad (3)$$

If the values of X_{ik} obtained for two groups did not differ markedly, it would imply that there was no essential difference between those two groups of liver disease. A test to determine whether the differences between 2 (or more) groups are meaningful has been suggested by Mahalanobis.[5-7] When this test (known as the D^2-test) was applied, significant differences ($P<.001$) were found to exist between the 5 groups.

Use of the Discriminant Functions

In practice the functions (3) above are used as follows: The first patient diagnosed by liver biopsy as belonging to group 1, for example, had the following values for variables 1 to 8: $x_1=10{\cdot}0$, $x_2=7{\cdot}8$, $x_3=13{\cdot}0$, $x_4=1701{\cdot}0$, $x_5=1150{\cdot}0$, $x_6=352{\cdot}0$, $x_7=970{\cdot}0$, $x_8=5{\cdot}1$. If we substitute these values for $x_1, \ldots x_8$ in each of the equations (3) we find that X_1 is greater than any of the remaining Xs. Our conclusion on the basis of the serum values is then that the patient belongs to group 1.

If we assume that the X-values are normally distributed, we can further calculate the probability that our allocation of a patient to a particular group is correct. For this purpose the following formula[8] is used:

$$\frac{e(X_k - \max X_i)}{\Sigma\, e(X_k - \max X_i)} \quad \ldots\ldots (4)$$

where e is the base of the natural logarithms, X_k is the k-th discriminant function, max X_i is the discriminant function which gives the highest value, and the Σ sign directs one to sum over-all groups for k= 1, . . . 5.

This procedure was carried out for each of the 194 cases. The results are too lengthy to be given in full, but as an example the results for the cases belonging to group 1 (according to diagnosis by liver biopsy) are given in Table III. We can

TABLE III. PROBABILITIES ATTACHED TO THE ALLOCATION OF THE CASES FROM GROUP 1 TO GROUPS 1–5*

Case	*Group No.*					*Largest probability*	*Group selected*
	1	*2*	*3*	*4*	*5*		
1	0·96	0·01	0·00	0·01	0·02	0·96	1
2	1·00	0·00	0·00	0·00	0·00	1·00	1
3	0·98	0·00	0·00	0·01	0·00	0·98	1
4	0·99	0·00	0·00	0·00	0·00	0·99	1
5	1·00	0·00	0·00	0·00	0·00	1·00	1
6	0·99	0·00	0·00	0·01	0·00	0·99	1
7	1·00	0·00	0·00	0·00	0·00	1·00	1
8	1·00	0·00	0·00	0·00	0·00	1·00	1
9	0·94	0·02	0·02	0·00	0·02	0·94	1
10	0·99	0·00	0·00	0·01	0·00	0·99	1
11	1·00	0·00	0·00	0·00	0·00	1·00	1
12	1·00	0·00	0·00	0·00	0·00	1·00	1
13	0·88	0·01	0·00	0·09	0·02	0·88	1
14	1·00	0·00	0·00	0·00	0·00	1·00	1
15	0·01	0·36	0·14	0·25	0·25	0·36	2
16	0·36	0·06	0·07	0·39	0·12	0·39	4
17	0·00	0·33	0·11	0·27	0·29	0·33	2
18	0·02	0·01	0·00	0·94	0·03	0·94	4
19	0·12	0·15	0·27	0·37	0·10	0·37	4
20	0·03	0·00	0·00	0·94	0·02	0·94	4
21	0·27	0·03	0·01	0·60	0·09	0·60	4
22	0·06	0·15	0·12	0·38	0·30	0·38	4
23	0·01	0·30	0·21	0·10	0·38	0·38	5
24	0·00	0·10	0·31	0·01	0·58	0·58	5

* A value of 1·00 in the table indicates that the probability differed from 1·00 by less than 0·01.

see from Table III that case 1, for example, was correctly allocated to group 1 with a probability of 0·96. Case 2 was correctly allocated with a probability of nearly 1·00. Case 15 was incorrectly allocated to group 2, but with a probability of only 0·36. In general it appears that when a case was incorrectly allocated, the probability tended to be low. There were, however, exceptions to this (e.g. case 18).

The results of the classification for each group are summarized in a classification matrix in Table IV.

TABLE IV. SUMMARY OF NEW ALLOCATION FOR GROUPS 1–5

Group to which case belonged	*No. of cases*	*Group in which case was classified*				
		1	*2*	*3*	*4*	*5*
1	24	**14**	2	0	6	2
2	52	0	**52**	0	0	0
3	50	1	13	**31**	0	5
4	19	0	7	0	**11**	1
5	49	2	25	8	2	**12**
	Total	17	99	39	19	20

The diagonal of the matrix (figures in bold type) indicates those cases which were correctly classified. It can be seen that, of the 24 cases in group 1, 14 were correctly classified. Of the remaining 10, 2 were (incorrectly) placed in group 2, 6 in group 4 and 2 in group 5. All 52 of the cases in group 2 were correctly classified, but the probabilities (not shown) were low, ranging from 0·33 to 0·56. For the remaining groups the probabilities associated with the correctly classified cases were usually higher than those for group 2 but lower than those for group 1.

Effect of Mixed Group

Of the 49 cases in group 5 (the mixed group) only 12 were correctly classified, while 25 were incorrectly placed in group 2. This suggests that group 5 was not homogenous in respect of the variables tested, but consisted of subgroups with divergent patterns.

If the cases taken together as group 5 had differing enzyme patterns, the effect would be to decrease the precision of the discriminant functions for differentiating between the 5 main groups. To ascertain whether this actually was the case, it was decided to carry out a further discriminatory analysis on group 5 taken by itself. This group was divided into 4 subgroups as shown in Table V. The Mahalanobis D^2-test showed that significant differences existed between the subgroups. The classification matrix found for these groups is shown in Table VI. It can be seen that 29 out of the 49 cases were correctly classified.

TABLE V. CLASSIFICATION OF GROUP 5 INTO SUBGROUPS OF LIVER DISEASE

Subgroup	*Liver disease*
1	Metastasis without obstruction
2	Obstruction (extrahepatic)
3	No histological abnormality
4	Liver abscess, amoebic abscess, tuberculosis, bilharzia, typhoid, sarcoidosis, amyloidosis

TABLE VI. SUMMARY OF NEW ALLOCATION FOR SUBGROUPS 1–4 OF GROUP 5

Group to which case belonged	*No. of cases*	*Group in which case was classified*			
		1	*2*	*3*	*4*
1	7	**2**	0	4	1
2	7	0	**5**	2	0
3	11	0	0	**11**	0
4	24	6	1	6	**11**
	Total	8	6	23	12

Effect of Fewer Variables

In view of the fact that the estimation of the 8 variables listed entails a considerable amount of work, it was decided to see to what extent the accuracy of the discriminatory analysis would be affected if fewer variables were used. The following four variables were selected as the most important on the basis of clinical considerations: (1) total bilirubin, (2) isocitric DH, (3) GPT, and (4) alkaline phosphatase.

Discriminatory functions were calculated for the 5 main groups in the same way as before. The following functions were found:

$$\left.\begin{aligned} X_1 &= -6{\cdot}3760 + 0{\cdot}6049x_1 + 0{\cdot}0005x_2 + 0{\cdot}0070x_3 + 0{\cdot}0118x_4 \\ X_2 &= -0{\cdot}1162 + 0{\cdot}0083x_1 + 0{\cdot}0001x_2 + 0{\cdot}0003x_3 + 0{\cdot}0675x_4 \\ X_3 &= -1{\cdot}4618 - 0{\cdot}0339x_1 + 0{\cdot}0011x_2 + 0{\cdot}0002x_3 + 0{\cdot}2160x_4 \\ X_4 &= -1{\cdot}1338 + 0{\cdot}3750x_1 + 0{\cdot}0002x_2 - 0{\cdot}0002x_3 - 0{\cdot}0382x_4 \\ X_5 &= -1{\cdot}0477 + 0{\cdot}1145x_1 + 0{\cdot}0002x_2 + 0{\cdot}0005x_3 + 0{\cdot}1874x_4 \end{aligned}\right\} \quad (5)$$

The results are summarized in a classification matrix in Table VII. A comparison of Table VII with Table IV shows that very little, if any, accuracy was lost by omitting the 4 variables. In groups 1 and 2 the same number of cases were correctly classified, in group 3 one case less, in group 4 one case more and in group 5 five cases less.

TABLE VII. SUMMARY OF NEW ALLOCATION FOR GROUPS 1–5

Group to which case belonged	*No. of cases*	*Group in which case was classified*				
		1	*2*	*3*	*4*	*5*
1	24	**14**	2	0	6	2
2	52	0	**52**	0	0	0
3	50	0	10	**30**	2	8
4	19	0	7	0	**12**	0
5	49	4	26	9	3	**7**
	Total	18	97	39	23	17

There are two possible explanations for this finding: In the first place, the variables omitted may have been highly correlated with the variables which were retained. This was probably true in the cases of conjugated bilirubin, LDH and GOT, which are closely associated with total bilirubin, isocitric DH and GPT respectively. A second possibility is that the omitted variables might not be connected with liver disease, this being probably true of aldolase.

DISCUSSION AND CONCLUSIONS

For the purpose of the above analyses it was assumed that the diagnoses according to liver biopsy were correct, but this was not necessarily the case. The probability that the diagnosis is correct must obviously vary from disease to disease. A diagnosis of hepatitis, cirrhosis or siderosis on the basis of a liver biopsy would indicate that the patient almost certainly actually suffered from the disease in question. A *positive* diagnosis of a localized disease such as hepatoma or liver abscess would likewise have a high degree of certainty, but the value of a negative diagnosis would be limited, for when the liver is not affected throughout by the disease in question, the biopsy needle might enter an unaffected area. If this were to happen in a case of hepatoma, for instance, the patient might be classified as suffering from cirrhosis and siderosis, and the effect of this misclassification would be to reduce the precision of the discriminant functions.

The other circumstance which is likely to have impaired the precision of the discriminant functions is the fact that in group 5 we had a mixed group of diseases. The majority of this group were incorrectly classified, and tended to cloud the picture by turning up in other groups, most of them in group 2 (Table IV). If patients belonging to group 5 had been excluded, those patients belonging to groups 1-4 could have been identified with a much higher degree of accuracy.

It is of interest that so many more cases were allocated to group 2 than to any of the other groups by the discriminatory analysis (Table IV), for this would appear to suggest that the basic pattern represented by this group is common to more than one variety of disease. It is also of interest that the probabilities associated with the cases allocated to group 2 were on the whole low, indicating that the cases in this group were less clearly demarcated from the rest than were, for instance, the first 14 cases in group 1 (Table IV). Since, then, group 2 differed less from the other groups than did these groups from one another, and since it contained the greatest number of cases, the possibility is suggested that group 2 represents a 'normal' pattern of which the others are modifications. This possibility would have been confirmed or excluded if a control group of normal patients could have been studied along with the others.

Let us now consider what value the investigations described above might have for the practising clinician.

It has been shown by means of the Mahalanobis D^2-test that different patterns exist in the serum constituents dealt with above which are associated with (and may be causally related to) specific groups of liver disease. The existence of definite patterns in the serum values of groups distinguished by means of liver biopsy indicates that an alternative method of diagnosis might hereby be provided which would dispense with the need for liver biopsy. It would, however, be impossible for the clinician to make a diagnosis with any degree of assurance merely by studying a patient's serum values (Table II). The discriminatory analysis technique places at his disposal a means of interpreting the values in an objective way. In so far as the constants in the discriminant functions are mathematically determined so as to give the most selective linear functions that can be found, the clinician may rest assured that he is making the 'best' diagnosis that the observed serum values (which may be subject to fluctuation and error due to a variety of causes) permit.

The value of the present results for clinical diagnostic purposes would appear to be somewhat limited. In group 1 (hepatitis and necrosis) only 14 of the 24 cases would be correctly diagnosed on the basis of the discriminant functions, and this is not particularly helpful, since the diagnosis of these diseases on clinical grounds is reasonably straightforward. In group 2 (siderosis, hepatofibrosis and cirrhosis) 52 out of 52 cases would be correctly diagnosed, but the value of this result would be greatly reduced by the fact that an additional 45 cases would be incorrectly placed in this group. In group 3 (hepatoma) 31 of the 50 cases would be correctly diagnosed, and this could be of real assistance, since hepatoma is not always readily diagnosed on clinical grounds. Even if pneumonia with jaundice (group 4) were not clinically obvious, however, a correct diagnosis in only 11 out of 19 cases on the basis of the discriminatory analysis would not mean a great deal to the clinician. In group 5 the findings would be rather misleading than otherwise, but this is only to be expected in a heterogeneous group.

The fact that highly significant differences were found between the subgroups of group 5 indicates that if we had been able to handle a discriminatory analysis operating on 8 individual groups, better results would have been obtained. Even on the basis of the present data alone, therefore, discriminant functions might have been found which would be of real use to the clinician. If data were to become available in which all diagnostic errors had been excluded (e.g. through postmortem examination) there could be little doubt that the discriminant functions found would constitute a valuable diagnostic aid. It remains to be established what combination of variables would be optimal from this point of view.

In conclusion, it should be pointed out that the discriminatory analysis technique has been shown by the present investigation to be applicable to clinical findings. The limited practical value of the present results is due to limitations in the data and not to shortcomings in the technique, which should have a valuable application wherever groups of individuals classified with difficulty by any other means are found to be distinguished by characteristic patterns in a combination of easily accessible variables.

SUMMARY

A group of 194 patients suffering from liver disease were classified into 5 types or categories of liver disease on the basis of liver biopsies.

For each patient the activity of 6 different serum enzymes was determined, together with total and conjugated serum bilirubin.

A statistical technique known as a discriminatory analysis was applied to the data to determine whether definite patterns existed in the different groups of patients in respect of the observed serum values which would enable the clinician to make a reliable diagnosis on the basis of these values.

For this purpose a linear equation in 8 variables was calculated for each category of liver disease. The constants of these equations were determined in such a way that maximal distinction would be made between the 5 groups.

These equations were used to classify each of the 194 cases. It was found that in 120 cases a correct diagnosis could be made solely on the basis of the observed serum values.

It is concluded that the discriminatory analysis technique might prove to have a valuable application in the field of clinical diagnosis.

The chemical determination of the serum constituents was carried out by Miss H. Wolff, Mrs. E. Praekelt and Mrs. E. Freier of the Nutrition Clinic for Adults of the NNRI.

The statistical calculations were done on the IBM 704 electronic computer of the National Research Institute for Mathematical Sciences, using a Biomedical Computer Program adapted to the 704 computer by Mr. F. E. Steffens of the Division of Statistics of this Institute. Dr. M. L. Neser, of the NNRI, assisted in the preparation of the report.

REFERENCES

1. Fisher, R. A. (1936): Ann. Eugen. (Lond.), **7**, 179.
2. Kendall, M. G. (1960): *The Advanced Theory of Statistics*, 4th ed., vol. 2, pp. 341-348. London: Charles Griffin & Co.
3. Tintner, G. (1952): *Econometrics*, pp. 96-102. New York: John Wiley & Sons.
4. Anderson, T. W. (1958): *An Introduction to Multivariate Statistical Analysis*, pp. 147-153. New York: John Wiley & Sons.
5. Mahalanobis, P. C. (1936): Proceedings of the National Institute of Science of India, **12**, 49.
6. Bose, R. C. and Roy, S. N. (1938): Sankhya, **4**, 19.
7. Mahalanobis, P. C., Bose, R. C. and Roy, S. N. (1937): *Ibid.*, **3**, 1.
8. Dixon, W. J. (1964): *BMD Biomedical Computer Programs*, p. 202. Los Angeles: University of California.