The Redundancy of Liver Tests in the Diagnosis of Cirrhosis Estimated by Multivariate Statistics

K. RAMSÖE, N. TYGSTRUP & P. WINKEL
Medical Dept. A, Division of Hepatology, and
Medical Dept. B, Laboratory of Biomathematics, Rigshospitalet,
Copenhagen, Denmark

Ramsöe, K., Tygstrup, N. & Winkel, P. The Redundancy of Liver Tests in the Diagnosis of Cirrhosis Estimated by Multivariate Statistics. *Scand. J. clin. Lab. Invest. 26,* 307-312, 1970.

Multivariate discriminant analysis was applied to 9 laboratory tests, performed in 17 patients with cirrhosis and 19 control subjects. Bromsulfalein transport maximum was the most effective test in discriminating between the two groups. The combination of all 9 tests could not be shown to discriminate significantly better than the combination of bromsulfalein transport maximum and gamma-globulin. These 2 tests, however, could be replaced by the remaining 7 tests without a significant loss of discriminative information. The result of the discriminant analysis is greatly dependent on the size and selection of the materials. In larger materials more tests with significant contribution to the discrimination may be found.

Key-words: A priori probability; discriminant analysis; laboratory diagnosis; liver cirrhosis; liver function tests

K. Ramsöe, M. D., Medical Dept. A, Division of Hepatology, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen Ö, Denmark

No biochemical liver test is specific for cirrhosis, yet the tests play a great role in the diagnosis of this disease, largely by evaluation of the 'test profile'. It is sometimes inferred that the diagnostic precision is increased by increasing the number of tests. The diagnostic value of a single test can be assessed by comparing the test results in two materials, one with and the other without the diagnosis in question. The value of a combination of tests can be examined in a similar manner by discriminant analysis (1). By this statistical method the discriminative effectiveness of each test is combined to a discriminant function, which may be assessed, compared and visualized as single test results. By comparison of discriminant functions one can moreover study whether new tests add information to that already obtained by previous tests or can replace the latter (4).

Zieve and Hill (17, 18, 19, 20) used multivariate discriminant analysis of liver tests which, except for the bromsulfalein retention test, are no longer in common use. Tengström (12) was particularly interested in the galactose T/2 in differentiating between different types of liver disease. In the present study this statistical method was applied to some currently used routine liver function tests, and to two more complicated tolerance tests from control subjects and patients with cirrhosis.

MATERIAL AND METHODS

The material comprised 17 consecutive cases of biopsy proven cirrhosis (10 males and 7 females). The mean age of the patients was 55 years (40-70 years). The duration of the disease ranged from 2 months to 10 years. Alcoholism was a possible etiology in 8 patients and hepatitis in 1 patient. The etiology was unknown in 8 patients. At the examination icterus was present in 10, hepatomegaly in 13, vascular spiders and palmar erythema in 13,

Table I. Comparison of liver tests in control subjects and patients with cirrhosis

	Control		Cirrhosis					
	Mean	95 per cent range	Mean	95 per cent range	t-value	Overlap point	Overlap area per cent	Cumulative overlap area per cent
BSP-Tmax. (mg/min)	9.4	5.9 – 12.9	3.7	0.5 – 6.8	10.3	6.5	4.3	4.3
Gamma globulin* (g/100 ml)	0.73	0.43 – 1.22	1.97	0.78 – 4.90	8.2	1.19	8.5	1.9
Bilirubin* (mg/100 ml)	0.53	0.19 – 1.46	1.73	0.28 – 10.7	4.9	0.95	20.9	1.2
Prothrombin (per cent)	83	44 – 122	63	9 – 116	2.6	73	33.0	1.0
BSP–S (mg/mg %)	66	6 – 125	25	12 – 62	4.8	45.5	20.9	0.9
Alk. phosphatase* (μmol/min)	27	14 – 53	94	22 – 403	6.8	49.9	12.7	0.8
Albumin (g/100 ml)	4.60	4.00 – 5.20	3.50	2.10 – 4.90	6.4	4.10	14.5	0.8
GE (mg/min)	439	249 – 630	268	103 – 434	5.7	354	17.1	0.7
Transaminase* (μmol/h/ml)	0.76	0.29 – 1.99	2.92	0.6 – 14.2	6.2	1.49	14.9	0.7

* These variates are log. normal ditributed.

and ascites in 7 patients. In 3 patients there was a slight encephalopathy, but no patient was in hepatic coma. Patients with surgical porto-caval anastomosis and patients in steroid treatment were excluded. The control material comprised 12 males and 7 females from 20 to 56 years of age (mean 37 years). The group included 8 normal healthy volunteers and 11 hospital patients with minor diseases not affecting liver function.

The following liver function tests were in each subject performed within three days by the methods indicated: bromsulfalein transport maximum (BSP-Tmax) and storage capacity (BSP-S) (15), the galactose elimination capacity (GE) (14), serum bilirubin (6), alkaline phosphatase (2), alanine transaminase (7), prothrombin (8), serum albumin and gammaglobulin (5).

Statistical analysis

The statistical tests used in this study require that the variates are normally distributed (1). The distributions of the variates from control subjects and patients with cirrhosis were examined separately by inspection of graphs of the cumulative distributions of the data and the logarithms of the data on probability paper. When the logarithms of the data were closer to a straight line on these graphs, the logarithms of the test values were used in the subsequent analysis. A stepwise discriminant analysis (3, 4) was performed, using all the liver tests. At the first step the single test giving the best discrimination was selected. At each of the following steps the variate which gave the greatest increase in the discrimination entered the analysis. The variates were ranked in this order. To study redundancy the discrimination function of all tests was compared to the discriminant functions derived by the stepwise addition of the tests according to their rank. When a discriminant function was reached which did not discriminate significantly less than the discriminant function of all tests, the remaining tests were considered redundant.

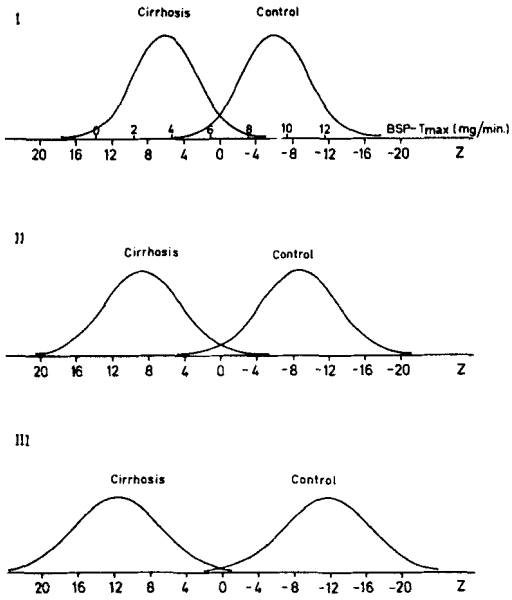

Fig. 1. The estimated probability distributions of z.
I. $z = -0.0605$ BSP-Tmax + 0.3960.
II. $z = -0.0559$ BSP-Tmax + 0.1912 log gamma-globulin + 0.3320.
III. $z = -0.0439$ BSP-Tmax + 0.1602 log gamma-globulin + 0.0982 log bilirubin + 0.0015 prothrombin − 0.0006 BSP-S + 0.0060 log alkaline phosphatase − 0.0544 albumin − 0.0001 GE + 0.0134 log alanine transaminase + 0.2107.

RESULTS

The values of serum albumin, BSP-Tmax, BSP-S, and GE showed a normal distribution. Logarithmic transformation gave an approximately normal distribution of alkaline phosphatase, bilirubin, alanine transaminase, and gamma-globulin in both groups.

In Table I the mean values of the tests in control subjects and patients with cirrhosis are given together with the theoretical 95 per cent range, calculated from the normal distribution curves. The t-value of the difference between mean values of controls and patients shows that this difference in all the tests examined is significant at the 5 per cent level. It is seen that the ranking is independent of this t-value except for the first two tests. The overlap point indicates the test value which gives the best separation between controls and patients. In most cases this value corresponds reasonably well to the accepted normal limit of the respective tests, but less so with the 95 per cent limit, calculated from the present control material. This also appears from the overlap area, i.e. the relative area of the distribution curve on the wrong side of the overlap point. This relative area indicates the proportion of controls and patients which are misclassified by the test.

The cumulative overlap area (see also Fig. 1) given in the last column of Table I illustrates the reduction in misclassification obtained by increasing the number of tests in the order in which they are ranked by the analysis. Each figure reflects the relative misclassification obtained by adding the corresponding test to all tests with a higher ranking.

The single test giving the best discrimination was BSP-Tmax, but it was significantly smaller than the discrimination based on all tests ($p < 0.05$). The discriminative effectiveness obtained by combining all nine tests, however, is not significantly greater (i.e. the relative misclassification is not significantly smaller) than that resulting from the two tests with the highest ranking, the BSP-Tmax and serum gamma-globulin. The discrimination between cirrhosis and control by BSP-Tmax and gamma-globulin is illustrated in Fig. 2. The discriminative effectiveness of the remaining seven tests was not significantly smaller than that of all nine tests. Likewise the more complicated 'quantitative' tolerance tests, BSP-Tmax, S

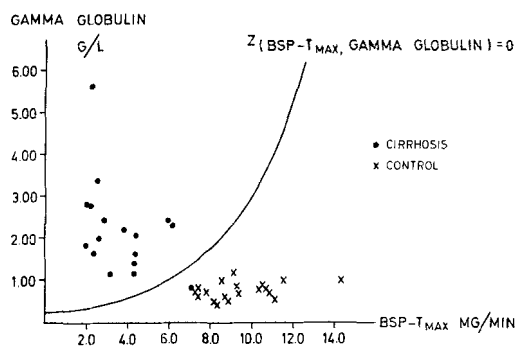

Fig. 2. The discrimination between cirrhosis and control by two variables when $q_B = q_A$ (see Appendix).

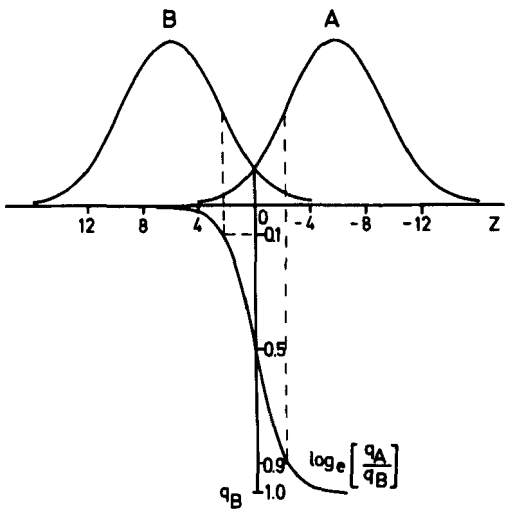

Fig. 3. The estimated probability distribution of $z_{(BSP\text{-}Tmax)}$ and the discrimination point ($\log_e q_A/q_B = \log_e 1\text{-}q_B/q_B$) depicted as a function of the a priori probability q_B.

and GE could be withdrawn from the nine tests without any significant reduction in the discriminative effectiveness.

DISCUSSION

In the present study it could be shown that no single liver function test could substitute the combination of all nine tests in discriminating between control subjects and patients with cirrhosis. The discriminative effectiveness of all nine tests in combination could not be demonstrated to be greater than that of the two highest ranked tests (BSP-Tmax and gamma-globulin). The remaining seven tests were therefore considered redundant in this connection. BSP-Tmax and gamma-globulin could, however, be deleted from the total combination without significantly reducing the discrimination.

Zieve and Hill (19) found that among nine liver function tests only four gave a significant independent contribution to the discrimination between control subjects and patients with cirrhosis. The highest ranking test was the BSP retention test followed by the zinc sulphate turbidity, intravenous hippuric acid, and urine coproporphyrine. The difference in the tests employed does not allow direct comparison with the present results, but it is noted that tests reflecting impaired BSP elimination and hypergamma-globulinemia in both materials gave the largest independent contribution to the discrimination.

The intravenous galactose tolerance test (60-minute retention) used in Zieve and Hill's material did not contribute independently to the discrimination, owing to a high correlation to the BSP retention test. The galactose elimination capacity used in the present study is highly correlated to BSP-Tmax (9). In Tengström's study it was found that the galactose T/2 was the highest ranking discriminator between cirrhosis and a control group. No BSP studies were done, however, in that material; furthermore it is not comparable with the present material, because the control group consisted of patients with suspected, but not confirmed, liver disease.

The discriminant analysis requires that the observations originate from one of two normal distributed populations, with identical covariance matrices. Furthermore, the a priori probabilities must be taken into account (Fig. 3). If the probability (q_A), that an observation belongs to population A, is equal to the probability ($q_B = 1\text{-}q_A$) that it belongs to population B, maximum discrimination will correspond to the point where Z, the discriminant function + a constant (see Appendix) is zero, i.e. to the intercept of the probability distribution curves. If $q_A > q_B$, maximum discrimination will correspond to $z > O$. Fig. 3 depicts the probability distribution of z, when the observations come from group A or group B, and in the lower part of the Figure the discrimination point as a function of the probability that an observation originates from group B. It appears that the influence of the probability q_B on the discrimination point is relatively small in the interval $0.1 < q_A < 0.9$.

The choice of limit between two populations should be influenced by the consequences implied by a misclassification. If the consequences of misclassifying a patient who belongs to population A into population B are more serious than vice versa, then a higher

discrimination point should be used for the discrimination.

Discriminant analysis may be used as a means of rationalizing a test program by indicating redundant tests. In the present material the number of tests could be reduced to two without significantly decreasing the discrimination between cirrhosis and control subjects. However, since redundancy is defined by statistical non-significant contribution to the discrimination, and conclusions based on lack of significance are more uncertain the smaller the material, this result cannot be applied to practical clinical work without some reservation. It is conceivable that fewer tests prove redundant if larger materials are examined.

An important clinical purpose of a statistical analysis of laboratory tests is to establish a test program which makes it possible to select patients with a certain disease (in casu cirrhosis) among all subjects suspected to suffer from this disease. This requires that the patients examined with and without the disease, respectively, are representative of the groups in question. It is doubtful whether the present group of patients with cirrhosis does represent cirrhosis in general, and the control material, mainly consisting of young and healthy subjects, does not represent the group of patients that will ordinarily be suspected of suffering from cirrhosis. The latter group will probably be less separated from the group with cirrhosis than the present normal control group will. The separation (i.e. the total discrimination) does not necessarily influence ranking and redundancy of the variables, which will depend only on their relative contribution to the discrimination.

The power with which laboratory tests discriminate between groups representing normality and a certain disease may give information about pathogenetic mechanisms of the disease if the physiological meaning of the tests is known. From animal experiments (15) and human studies (11) it is known that the BSP transport maximum is a measure of the hepatic excretory capacity. Increased serum gamma-globulin is usually considered to be a sign of chronic inflammation. The large diagnostic weight of BSP-Tmax and serum gamma-globulin found in the present work is thus in agreement with the concept that decreased excretory capacity and chronic inflammation are important pathophysiological features of liver cirrhosis.

The diagnostic situation presented is somewhat hypothetical. In clinical practice mostly several diagnostic possibilities exist, and a number of factors other than the laboratory tests are involved. A diagnostic problem may, however, be split up into a number of sequential choices between two diagnoses which may be subject to discriminant analysis. The information obtained from the laboratory tests regarding the prognosis of the disease, the activity, etc. also requires separate analysis.

APPENDIX

Given two p-variate normal populations (π_A and π_B) with a common covariance matrix (Σ) and different mean vectors (μ_A and μ_B). Let x be a p-dimensional vector of observations coming from either π_A or π_B and let q_A be the probability that x belongs to π_A, and q_B ($= 1-q_A$) be the probability that x belongs to π_B.

The probability of misclassifying x is minimized by the following rule (1):

If $z \geqq \log_e \left\{\frac{q_A}{q_B}\right\}$, then x is assigned to π_B, and

if $z < \log_e \left\{\frac{q_A}{q_B}\right\}$, then x is assigned to π_A

$z = D(x) - \frac{1}{2}\alpha$, where

$D(x)$ (the discriminant function) $= x' \Sigma^{-1}(\mu_B - \mu_A)$,

and $\alpha = (\mu_B + \mu_A)' \Sigma^{-1} (\mu_B - \mu_A)$.

$\text{Log}_e \left\{\frac{q_A}{q_B}\right\}$ may be designated the discrimination point for z. If x comes from π_A, z follows a normal distribution with mean $-\frac{1}{2}\alpha$ and variance α. If x comes from π_B, z follows a normal distribution with mean $\frac{1}{2}\alpha$ and variance α.

REFERENCES

1. Anderson, T. W. *An Introduction to Multivariate Statistical Analysis.* Wiley and Sons, New York, 1966.

2. Bessey, O. A., Lowry, O. H. & Brock, M. J. Method for rapid determination of alkaline phosphatase with 5 cubic millimeters of serum. *J. biol. Chem. 164,* 321, 1946.
3. Dixon, W. J. *BMD Biometrical Computer Programs, BMD 07 M.* University of California Press, 1967.
4. Efroymsen, M. A. Multiple regression analysis. p. 191 in Ralston, A. & Wilf, H. S. (eds.) *Mathematical Methods for Digital Computers,* part V. Wiley & Sons, New York, 1960.
5. Hjelm, M. A methodological study of the enzymatic determination of galactose in human whole blood, plasma and erythrocytes with galactose oxidase. *Clin. chim. Acta 15,* 87, 1967.
6. Jendrassik, L. & Grof, P. Vereinfachte photometrische Methoden zur Bestimmung des Blutbilirubins. *Biochem. Z. 279,* 81, 1938.
7. Laursen, T. & From Hansen, P. A fluorimetric method for measuring the activity in serum of the enzyme glutamic pyruvic transaminase. *Scand. J. clin. Lab. Invest. 10, 53,* 1958.
8. Owren, P. A. & Aas, K. The control of dicumarol therapy and the quantitative determination of prothrombin and proconvertin. *Scand. J. clin. Lab. Invest. 3,* 201, 1951.
9. Ramsöe, K. To be published in *Scand. J. clin. Lab. Invest.*
10. Rao, C. R. *Advanced Statistical Methods in Biometric Research,* p. 253. Wiley and Sons, New York, 1952.
11. Schoenfield, L. J., McGill, D. B. & Foulk, W. T. Studies of sulfobromophtalein sodium (BSP) metabolism in man. III. Demonstration of a transport maximum (Tm) for biliary excretion of BSP. *J. clin. Invest. 43,* 1424, 1964.
12. Tengström, B. The discriminatory ability of a galactose tolerance test and some other tests in the diagnosis of cirrhosis of the liver, hepatitis and biliary obstruction. *Scand. J. clin. Lab. Invest. 23,* 159, 1969.
13. Tiselius, A. & Flodin, P. Zone electrophoresis. *Advanc. Protein Chem. 8,* 461, 1953.
14. Tygstrup, N. Determination of the hepatic galactose elimination capacity after a single intravenous injection in man. *Acta physiol. scand. 58,* 162, 1963.
15. Wheeler, H. O., Meltzer, J. I. & Bradley, S. E. Biliary transport and hepatic storage of sulfobromophtalein in the unanesthetized dog, in normal man, and in patients with hepatic disease. *J. clin. Invest. 39,* 1131, 1960.
16. Winkler, K., Tygstrup, N. & Munkner, T. A study of Gaebler's method for determination of bromsulfalein in plasma. *Scand. J. clin. Lab. Invest. 12,* 357, 1960.
17. Zieve, L. & Hill, E. An evaluation of factors influencing the discriminative effectiveness of a group of liver function tests. I. The utilization of multiple measurements in medicine. *Gastroenterology 28,* 759, 1955.
18. Zieve, L. & Hill, E. An evaluation of factors influencing the discriminative effectivenes of a group of liver function tests. II. Normal limits of eleven representative hepatic tests. *Gastroenterology 28,* 766, 1955.
19. Zieve, L. & Hill, E. An evaluation of factors influencing the discriminative effectiveness of a group of liver function tests. III. Relative effectiveness of hepatic tests in cirrhosis. *Gastroenterology 28,* 785, 1955.
20. Zieve, L. & Hill, E. An evaluation of factors influencing the discriminative effectiveness of a group of liver function tests. IV. Nature of the interrelationships among hepatic tests in cirrhosis. *Gastroenterology 28,* 914, 1955.

Received 2 January 1970
Accepted 13 June 1970