Prediction of Fatality in Fulminant Hepatic Failure

E. CHRISTENSEN, A. BREMMELGAARD, M. BAHNSEN,
P. BUCH ANDREASEN & N. TYGSTRUP
Medical Dept. A and Dept. of Clinical Chemistry CL, Rigshospitalet;
Dept. of Hepatology, Hvidovre Hospital, University of Copenhagen;
and Medical Dept. F, Gentofte Hospital; Copenhagen, Denmark

Thirty-three consecutive patients admitted to the intensive care liver unit of Rigshospitalet with acute hepatic encephalopathy induced by viral hepatitis (17), drugs (14), or pregnancy (2) were studied. All were treated with a standard anticoma regime. The 20 patients (61%) who died had a higher bilirubin level and lower total cholic acid conjugation and glycine cholic acid conjugation ($p < 0.05$) than the surviving patients. Antipyrin clearance and galactose elimination capacity tended to be lower in the non-survival group than in the survival group ($p = 0.09$ and 0.11, respectively). Of single variables a bilirubin level of $>384 \mu mol/l$ gave the best prediction of non-survival (sensitivity, 0.80; specificity, 0.69; PVpos, 0.80; PVneg, 0.69; kappa, 0.49). However, a discriminant score based on combination of variables distinguished completely between non-survivors and survivors when validated by an unbiased method in which each patient is classified on the basis of the other patients' data. It is suggested that the discriminant score is used to select patients with very low probability of survival for liver transplantation or liver assistance procedures of unknown value.

Erik Christensen, M.D., Dept. of Hepatology 233, Hvidovre Hospital, University of Copenhagen, DK-2650 Copenhagen, Denmark

The fatality rate of fulminant hepatic failure is about 0.8 (1). The possibility of applying new resource-demanding and potentially hazardous liver assistance increases the importance of early identification of patients who do not benefit from standard treatment. Even though several prognostic indicators have been identified, such as age and etiology (2), hepatocyte volume, coagulation factors, bilirubin (3), ammonia levels (4), galactose elimination capacity (5), plasma phenazone clearance (6), bile acid conjugation (7), and $\alpha 2$-HS-glycoprotein (8), the general applicability of these measures for distinguishing between survivors and non-survivors in fulminant hepatic failure remains questionable.

The purpose of the present study was to evaluate the discriminative effectiveness of a set of observations in this situation.

PATIENTS AND METHODS

The patients, comprising 33 consecutive cases of fulminant hepatic failure admitted to the intensive care liver unit of Rigshospitalet, Medical Dept. A, have been described in detail elsewhere (9).

The hepatic failure was induced by hepatitis B (15 patients), non-A, non-B hepatitis (2), halothane (8), disulfiram (3), sulfamethoxazole with trimethoprim (1), paracetamol (1), radiotherapy plus vincristine (1), and acute fatty liver of pregnancy (2). Nineteen were females and 14 males. The median age was 30 years, the range being 17–72 years. On admission to the liver unit 11 patients had hepatic encephalopathy of grade I, 7 of grade II, 3 of grade III, and 12 of grade IV (10). The median duration of history of disease was 23 days, the range being 1–130 days.

In addition to standard hepatic tests determi-

nation of galactose elimination capacity (11), anti-pyrine clearance (12), and bile acid analyses (9) were performed on admission to the liver unit by the methods indicated.

All patients received standard treatment, including 20% glucose, 50 ml/h, as the only supply of calories; lactulose, 15–60 ml four times daily; and phytomenadione, 10 mg daily intravenously. Assisted respiration was administered before hypoxia or hypercapnia developed. Twenty patients died with signs of progressive, fulminant hepatic failure.

Clinical and laboratory data are given in Table I for the groups of survivors and non-survivors.

Statistical analysis

Comparison of data between the survivors and non-survivors was performed by the Mann–Whitney test (continuous variables) or Fisher's exact probability test (dichotomous variables) (13).

The discriminative effectiveness of each single variable was investigated by estimation of the sensitivity (incidence of a positive observation in non-survivors), specificity (incidence of a negative

Table I. Variables in patients surviving or not surviving acute liver failure

Variable	Non-survivors (no. = 20)		Survivors (no. = 13)		P value
	Median or %	Range	Median or %	Range	
Age (years)	31.5	20–72	26	17–49	0.11
Females (%)	60	—	54	—	0.99
Viral hepatitis B (%)	50	—	38	—	0.77
Non-A non-B hepatitis (%)	0	—	15	—	0.30
Halothane hepatitis (%)	25	—	23	—	0.77
Disulfiram hepatitis (%)	15	—	0	—	0.42
Fatty liver of pregnancy (%)	0	—	15	—	0.30
Body weight (kg)	65.0	47–120	65.0	47–81	0.68
Grade of encephalopathy (1–4)	4	1–4	2	1–4	0.07
Duration of history (days)	24	4–130	15	1–43	0.13
Blood type O (%)	35	—	45	—	0.85
Leukocytes ($\times 10^3/\mu l$)	11.9	4.4–32.0	7.8	1.0–19.1	0.06
Serum potassium (mmol/l)	3.75	2.5–7.2	3.5	2.3–4.2	0.20
Serum creatinine (mmol/l)	0.11	0.03–0.67	0.09	0.03–0.35	0.36
Serum carbamide (mmol/l)	8.5	1.6–40.6	5.1	1.4–20.6	0.28
Blood glucose (mmol/l)	6.3	1.3–11.0	5.2	2.3–10.0	0.16
Serum albumin (550–830 µmol/l)	428	253–520	421	352–579	0.90
Prothrombin index (0.7–1.3)	0.16	0.04–0.51	0.24	0.07–0.46	0.17
Serum bilirubin (4–17 µmol/l)	545	83–946	336	69–734	0.02
Alanine aminotransferase (10–40 U/l)	525	48–5800	600	52–7520	0.80
Alkaline phosphatase (80–275 U/l)	347	72–760	368	196–720	0.94
Galactose elimination capacity (21.4–54.5 µmol/min·kg)	12.1	9.7–19.7	14.0	10.9–19.7	0.07
Plasma phenazone clearance (356–1150 µl/min·kg)	71.5	43.3–242.5	104.0	48.1–209.0	0.09
Total (24-^{14}C) cholic acid conjugation (%/kg)	0.85	0.29–1.47	1.21	0.07–1.98	0.02
Glycine (24-^{14}C) cholic acid conjugation (%/kg)	0.47	0.12–1.19	0.70	0.18–1.74	0.01
Taurine (24-^{14}C) cholic acid conjugation (%/kg)	0.24	0.00–0.68	0.39	0.003–0.70	0.13
Sulphate (24-^{14}C) cholic acid conjugation (%/kg)	0.011	0.00–0.12	0.009	0.0–0.066	0.77
Glycolithocholic acid sulphate (µg/ml)	2.90	0.60–16.0	3.2	1.00–16.8	0.66
Glycocholic acid (µg/ml)	26.4	3.25–420	20.3	0.40–51.2	0.37

observation in survivors), predictive value of a positive observation (PVpos) (incidence of non-survival in patients with a positive observation), predictive value of a negative observation (PVneg) (incidence of survival in patients with a negative observation) (13), and kappa (the degree of agreement between a variable and the outcome (survival or non-survival) corrected for chance agreement) (14). For continuous variables the above estimations were performed after dichotomization, using the discrimination point that maximized chi-square (13). Observations with a higher frequency among non-survivors were considered positive.

The efficacy of correct allocation by sets of observations was studied by multivariate discriminant analysis (15), using stepwise addition of the variable with the most significant F-value for inclusion (15). The discriminative power of the estimated discriminant function was studied by allocating each patient on the basis of all the other patients ('leaving current patient out' method) (16), except when specifically stated. This means that for each of the 33 possible sets of 32 patients a discriminant function was derived and that the excluded patient was classified by that function. To include all patients in the multivariate analysis, a few missing observations (on average 3%) were replaced by the grand mean of the appropriate variable (15). The discriminant function reduced the variables of each patient to a single number (his/her discriminant score). To allow expression of the strength of evidence for any patient falling into each of the two groups, the distribution of the discriminant scores in each group was fitted by a normal distribution (13). For a particular value of the discriminant score the ratio of the ordinates of the two normal distribution curves (the likelihood ratio) is an estimate of the relative

likelihood of belonging to one or the other group (13). From the likelihood ratio and the prior probabilities of non-survival of 0.5, 0.6 (estimated from the present series), and 0.8 (a pooled estimate based on the literature (1)) the posterior probability of non-survival was estimated by Bayes' theorem (13).

RESULTS

Table I gives the observations in survivors and non-survivors. Only 3—that is, bilirubin, total cholic acid conjugation, and glycine cholic acid conjugation—are significantly different, with P values of 0.05 or smaller. Galactose elimination capacity, antipyrin clearance, grade of encephalopathy, and leukocytes are close to statistical significance.

The discriminative effectiveness of the three significant variables is shown in Table II. The highest kappa value was 0.49 (bilirubin, >384 μmol/l, and total cholic acid conjugation, <0.52%/kg), reflecting relatively poor discriminative effectiveness of single variables.

The result of the discriminant analysis based on the 33 patients is shown in Table III (the 'full score'), where variables are ranked by the order in which they contributed to discrimination. It appears that factors other than the P values given in Table I influence this order.

The distribution of discriminant scores for all patients (obtained by the 'leaving current patient out' method) and the best-fitting normal distribution curve for survivors and non-survivors are shown in Fig. 1. Using 0.5 as the discrimination point, there is complete separation between survivors and non-survivors; that is, sensitivity, specificity, PVpos, PVneg, and kappa are all 1.0.

The likelihood ratio of not surviving to surviv-

Table II. Discriminative effectiveness of significant variables from Table I (prediction of non-survival)

Variable	Sensitivity	Specificity	PVpos	PVneg	Kappa
Serum bilirubin >384 μmol/l	0.80	0.69	0.80	0.69	0.49
Total (24-^{14}C)cholic acid conjugation <1.09%/kg	0.80	0.69	0.80	0.69	0.49
Glycine (24-^{14}C)cholic acid conjugation <0.52%/kg	0.60	0.92	0.92	0.60	0.47

Table III. Significant discriminant function coefficients* for prediction of fatality in patients with fulminant hepatic failure. (All variables subjected to analysis ('full score'))

Variable	Scoring	Coefficient	P value
Taurine (24-^{14}C)cholic acid conjugation	%/kg	−1.15	3×10^{-5}
Sex	Female: 0 Male: 1	−0.42	7×10^{-5}
Pregnancy	Present: 1 Absent: 0	−0.85	2×10^{-4}
Carbamide	Mmol/l	0.024	2×10^{-4}
Alkaline phosphatase	U/l	0.0015	3×10^{-4}
Glycolithocholic acid sulphate	μg/ml	−0.040	6×10^{-4}
Non-A non-B viral hepatitis	Present: 1 Absent: 0	−0.74	0.003
Age	Years	0.012	0.003
Glycine (24-^{14}C)cholic acid conjugation	%/kg	−0.53	0.008
Duration of history	Days	0.0044	0.01
Halothane hepatitis	Present: 1 Absent: 0	−0.27	0.02
Grade of encephalopathy	1, 2, 3, or 4	0.087	0.02
Leukocyte count	$\times 10^3/\mu l$	0.013	0.05
Constant		0.18	

* The discriminant score of a patient is obtained by multiplying each variable by the corresponding coefficient and adding all the products to the constant.

ing and the probablility of non-survival as a function of the score are shown in the lower parts of Fig. 1. If the score is 1.0, non-survival is almost certain, and if it is 0.0, survival is almost certain for the selected values of prior probability of non-survival. For scores around 0.5 outcome is uncertain, and in this area the posterior probability of non-survival is influenced markedly by the prior probability of non-survival.

The analysis was repeated omitting the special liver function tests (galactose elimination capacity, antipyrine clearance, and bile acid analyses). Table IV shows the most effective combination of the remaining variables (the 'reduced score'). The distribution of the 'reduced scores' is shown in Fig. 2. Using the same discrimination point as in the overall analysis (0.5), sensitivity was 0.90, specificity 0.92, PVpos 0.95, PVneg 0.86, and kappa 0.81. It can be seen that the shape of the likelihood ratio curves differ from those in Fig. 1 in being less steep, reflecting a somewhat poorer discriminative effectiveness of the 'reduced score' compared with the 'full score'. This is also reflected in the probability curves.

DISCUSSION

The number of patients studied is small, but most of the commoner causes of hepatic failure are represented. The fatality rate of 60% is significantly lower than the pooled estimate of about 80% from the literature (1), but many of the smaller series previously published have similar fatality rates (1).

In this study no single variable could in itself distinguish between non-survivors and survivors in patients with fulminant hepatic failure, but combination of factors in discriminant functions markedly improved differentiation.

The discriminant functions also included variables differing insignificantly between survivors and non-survivors by univariate statistical testing. This phenomenon commonly occurs because univariate—in contrast to multivariate—statistical methods do not take into account the influence of other variables on the variable in question.

On the other hand, tests which by themselves are relatively good predictors may not be included in the discriminant function if they are highly

Table IV. Significant discriminant function coefficients* for prediction of fatality in patients with fulminant hepatic failure. (Only clinical and simple laboratory variables subjected to analysis ('reduced score'))

Variable	Scoring	Coefficient	P value
Disulfiram hepatitis	Present: 1 Absent: 0	0.83	3×10^{-4}
Blood glucose concentration	Mmol/l	0.089	5×10^{-4}
Duration of history	Days	0.0081	8×10^{-4}
Leukocyte count	$\times 10^3/\mu l$	0.032	0.001
Age	Years	0.015	0.005
Prothrombin index	Arb. units	−1.28	0.008
Sex	Female: 0 Male 1	−0.27	0.02
Viral hepatitis B	Present: 1 Absent: 0	0.31	0.02
Serum potassium	Mmol/l	0.12	0.03
Serum albumin	μmol/l	−0.0019	0.05
Blood type O	Present: 1 Absent: 0	−0.25	0.05
Constant		−0.45	

* The discriminant score of a patient is obtained by multiplying each variable by the corresponding coefficient and adding all the products to the constant.

correlated with and therefore containing nearly the same information as another test already included.

The unfavorable prognostic factors in the discriminant functions are indicators of severer disease (for example, severer encephalopathy, history of longer duration, and more pronounced cholestasis (high alkaline phosphatase (full score) but not high bilirubin, since this factor was closely correlated to the duration of history), severer renal insufficiency, or more reduced hepatic functional capacity (reduced taurine and glycine cholic acid conjugation and reduced glycolithocholic acid sulphate concentration (full score), lower prothombin index, and lower serum albumin (reduced score)). High leukocyte count indicative of inflammation (due to endotoxins or liver necroses?) is an unfavorable prognostic factor. The finding that high plasma glucose concentration seems to have an unfavorable prognostic influence was surprising, but it reflects the situation under treatment with glucose, which was started at the first sign of encephalopathy and in many cases before referral to the liver unit. If this finding is not accidental, it may reflect changed utilization pattern of insulin and glucagon (claimed to be hepatotrophic factors) in liver or extrahepatic tissues (17, 18).

In the multivariate analyses male sex seems to be associated with better prognosis. The reason for this is unknown, but the effect of male sex hormone on liver regeneration may be important (19, 20). The finding that blood type O is associated with better prognosis is difficult to explain biologically and may be accidental, since the level of significance is not very high. One must bear in mind that the number of patients analyzed in the present study is relatively small, and when many statistical tests are performed, the risk of committing type 1 errors increases. Thus the prognostic significance of some variables may be accidental, but most of the factors identified are biologically meaningful.

The discriminant score was evaluated by an unbiased method in which each patient is classified on the basis of the other patients' data. With this method the reduced score distinguished quite well (kappa, 0.81) and the full score distinguished completely (kappa, 1.00) between survivors and non-survivors. If the unbiased method had not been used, the following results would have been obtained for the reduced score: kappa, 0.94; sensitivity, 0.92; specificity, 0.92; PVpos, 0.95; and PVneg, 1.00.

The confidence of the results may to a certain degree be expressed by the 95% confidence limits
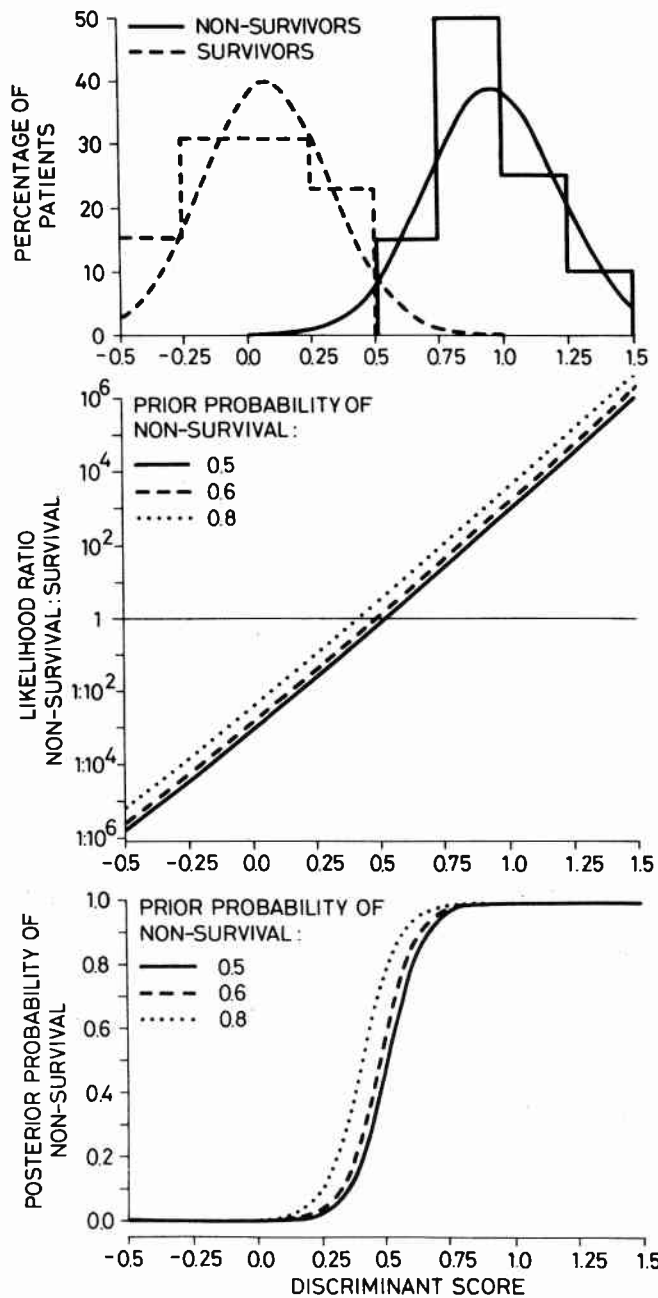
Fig. 1. Distribution of discriminant scores (obtained by the 'leaving current patient out' method) for survivors and non-survivors of fulminant hepatic failure, the best-fitting normal distribution curves, the ratio of their ordinates (the likelihood ratio), and the posterior probability of non-survival for the discriminant function in Table III ('full score').
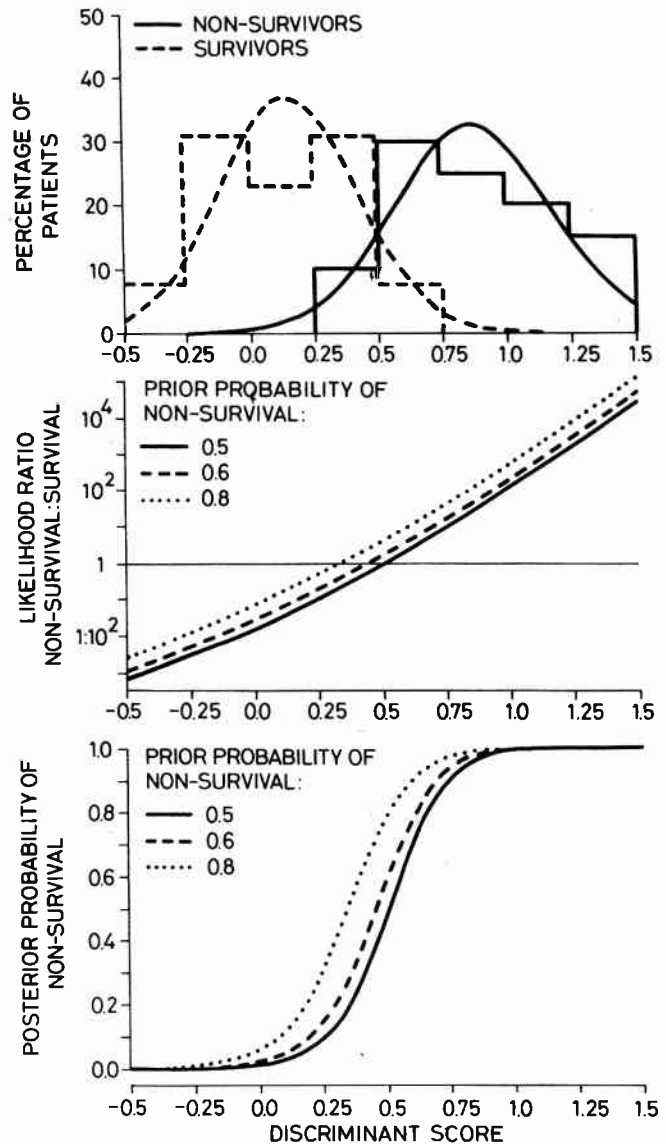
Fig. 2. Distribution of discriminant scores (obtained by the 'leaving current patient out' method) for survivors and non-survivors of fulminant hepatic failure, the best-fitting normal distribution curves, the ratio of their ordinates (the likelihood ratio), and the posterior probability of non-survival for the discriminant function in Table IV ('reduced score').

of the discrimination point, 0.5 ± 2 SE(d), where the standard error of the discrimination point SE(d) is estimated as $1/2\sqrt{SEM_1^2 + SEM_2^2}$, SEM_1 and SEM_2 being the standard error of mean of the scores in the two groups of patients. In the present study the 95% confidence interval of the discrimination point amounted to 0.41–0.59 for the full score (corresponding to the limits of 1.00

to 0.95 for sensitivity and 1.00 to 1.00 for specificity) and to 0.40–0.60 for the reduced score (corresponding to the limits of 0.90–0.75 for sensitivity and 0.77–0.92 for specificity).

For the results to be utilized in new patients with fulminant hepatic failure, they should be expressed as 'predictive values' (PV). This demands knowledge of the *a priori* probability of non-survival. PVpos and PVneg have been estimated for the studied patients. Furthermore, the posterior probability of non-survival in relation to the score has been estimated for three different

prior probabilities, using Bayes' theorem (13). Corresponding to the curve closest to the prior probability of non-survival in a particular setting, the posterior probability of non-survival can be read against the value of the discriminant score (Figs. 1 and 2). The reduced score, including only clinical and simple laboratory tests, may be obtained in a short time. If this score has a value corresponding to either very high or very low probability of non-survival, the full score may not be necessary. If the reduced score is intermediate (that is, 0.0–0.7), the full score demanding more special and time-consuming laboratory tests may be required for a more precise prognosis.

The differentiation may have practical implications. If, for example, the facilities of an intensive care liver unit are not available for all patients with fulminant hepatic failure, the scores may serve to exclude patients who are most unlikely to survive under any circumstance. Identification of the latter group also serves to select patients for 'liver assistance' procedures of unknown value. Thereby patients with a chance to survive with conventional treatment are not exposed to a potentially hazardous procedure, and if the procedure has certain benefits, this is more likely to be recognized if it is only applied to patients with a high probability of dying with conventional treatment alone. Similar considerations may apply when liver transplantation is a possibility.

Our results should be regarded as provisional. They may be improved in the future if better variables can be found, if more patients can be analyzed, and particularly if not only transectional observations, as used here, but also longitudinal observations can be utilized.

ACKNOWLEDGEMENTS

REFERENCES

1. Tygstrup N, Ranek L. Clin Gastroenterol 1981, 10, 191–208
2. Trey C. Can Med Ass J 1972, 106, 525–7
3. Nusinovici V, Crubille C, Opolon P, Touboul J-P, Darnis F, Caroli J. Gastroenterol Clin Biol 1977, 1, 875–86
4. Demedts M, de Groote J, van Damme B, Desmet VJ. Digestion 1974, 11, 105–14
5. Ranek L, Andreasen PB, Tygstrup N. Gut 1976, 17, 959–64
6. Ramsøe K, Andreasen PB, Ranek L. Scand J Gastroenterol 1980, 15, 65–72
7. Horak W, Waldram R, Murray-Lyon IM, Schuster E, Williams R. Gastroenterology 1976, 71, 809–13
8. Kobayashi K, Kameda S, Sugimoto T, Suzuki K, Nishimura K, Kato Y, Hattori N, Nakaya Y, Takase S, Takade A, Takeuchi J. Acta Hepatogastroenterol 1978, 25, 287–91
9. Bremmelgaard A, Ranek L, Bahnsen M, Buch Andreasen P, Christensen E. Scand J Gastroenterol 1983, 18, 797–802
10. Trey C, Davidson CS. In: Popper H, Schaffner F, eds. Progress in liver diseases. Grune & Stratton, New York, 1970, Vol. 3, 282–98
11. Tygstrup N. Scand J Clin Lab Invest 1966, 18 (suppl 92), 118–25
12. Andreasen PB, Ranek L, Statland BE, Tygstrup N. Eur J Clin Invest 1974, 4, 129–34
13. Armitage P. Statistical methods in medical research. Blackwell, Oxford, 1971
14. Koran LM. N Engl J Med 1975, 293, 642–6
15. Nie NH, Hull CG, Jenkins JG, Steinbrenner K, Bent DH. Statistical package for the social sciences. 2nd ed. McGraw Hill, New York, 1975
16. Lachenbruch PA, Mickey MR. Technometrics 1968, 10, 1–11
17. Farivar M, Wands JR, Isselbacher KJ, Bucher. N Engl J Med 1976, 295, 1517–9
18. Baker AL, Jaspan JB, Haines NW, Hatfield GE, Krager PS, Schneider JF. Gastroenterology 1981, 80, 1410–4
19. Sengupta P, Sarkar SK. J Indian Med Assoc 1977, 68, 47–9
20. Vic P, Saint-Aubert B, Astre C, Bories P, Bonardet A, Descomps B, Humeau C, Joyeux H. Hepatology 1982, 2, 247–8