

Research article

Open Access

A comparison of Child-Pugh, APACHE II and APACHE III scoring systems in predicting hospital mortality of patients with liver cirrhosis

Constantinos Chatzicostas¹, Maria Roussomoustakaki*¹, Georgios Notas², Ioannis G Vlachonikolis³, Demetrios Samonakis¹, John Romanos⁴, Emmanouel Vardas¹ and Elias A Kouroumalis^{1,2}

Address: ¹Department of Gastroenterology, University Hospital, Heraklion, Crete, Greece, ²Liver Research Laboratory, University of Crete Medical School, Greece, ³Biostatistics Laboratory, University of Crete Medical School, Greece and ⁴Department of Surgical Oncology, University Hospital, Heraklion, Crete, Greece

Email: Constantinos Chatzicostas - matrix01gr@yahoo.gr; Maria Roussomoustakaki* - mrouso1@hotmail.com; Georgios Notas - chief@med.uoc.gr; Ioannis G Vlachonikolis - i.vlachonikolis@surrey.ac.uk; Demetrios Samonakis - photein@in.gr; John Romanos - romanos@med.uoc.gr; Emmanouel Vardas - evardas@hotmail.com; Elias A Kouroumalis - kouroum@med.uoc.gr

* Corresponding author

Published: 8 May 2003

Received: 20 September 2002

BMC Gastroenterology 2003, 3:7

Accepted: 8 May 2003

This article is available from: <http://www.biomedcentral.com/1471-230X/3/7>

© 2003 Chatzicostas et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: The aim of this study was to assess the prognostic accuracy of Child-Pugh and APACHE II and III scoring systems in predicting short-term, hospital mortality of patients with liver cirrhosis.

Methods: 200 admissions of 147 cirrhotic patients (44% viral-associated liver cirrhosis, 33% alcoholic, 18.5% cryptogenic, 4.5% both viral and alcoholic) were studied prospectively. Clinical and laboratory data conforming to the Child-Pugh, APACHE II and III scores were recorded on day 1 for all patients. Discrimination was evaluated using receiver operating characteristic (ROC) curves and area under a ROC curve (AUC). Calibration was estimated using the Hosmer-Lemeshow goodness-of-fit test.

Results: Overall mortality was 11.5%. The mean Child-Pugh, APACHE II and III scores for survivors were found to be significantly lower than those of nonsurvivors. Discrimination was excellent for Child-Pugh (ROC AUC: 0.859) and APACHE III (ROC AUC: 0.816) scores, and acceptable for APACHE II score (ROC AUC: 0.759). Although the Hosmer-Lemeshow statistic revealed adequate goodness-of-fit for Child-Pugh score ($P = 0.192$), this was not the case for APACHE II and III scores ($P = 0.004$ and 0.003 respectively)

Conclusion: Our results indicate that, of the three models, Child-Pugh score had the least statistically significant discrepancy between predicted and observed mortality across the strata of increasing predicting mortality. This supports the hypothesis that APACHE scores do not work accurately outside ICU settings.

Background

The recognition of risk factors that can stratify a population of cirrhotic patients into subgroups with different survival is of great prognostic value for the clinician. Numerous attempts have been made to develop a reliable prognostic survival model for cirrhosis. The target population of the different scoring systems in the literature covers patients with liver cirrhosis [1–7], alcoholic liver disease [8,9], variceal bleeding [10–17], and upper gastrointestinal bleeding including variceal bleeding [18–20]. The Child-Turcotte classification [1] and its subsequent modification by Pugh [10] are old empiric methods to assess hepatocellular functional reserve in candidates for portosystemic shunting. Although Child-Turcotte and Child-Pugh scores (CPS) have not been formally evaluated for their statistical accuracy, they have been useful for risk-stratifying groups of patients with cirrhosis [21–23], for assessing the efficacy of interventional procedures such as transjugular intrahepatic portosystemic shunting [24,25] or sclerotherapy [17,26], and for evaluating therapy for complications of cirrhosis [27–29]. Although CPS score is considered an adequate method to establish the degree of liver failure and the survival probability [30], two of its elements are very subjective (ascites and encephalopathy), and a further limitation is its limited discriminatory ability [7]. In some studies, the prognostic value of CPS is described as incomplete, and other variables are demonstrated to have prognostic significance [31]. In addition, prognostic factors unrelated to hepatic function (cardiac, renal, pulmonary, acid-base and electrolyte status, other important associated comorbid conditions and factors) are not included.

Acute Physiology, Age and Chronic Health Evaluation (APACHE) II and III scores were developed by Knaus *et al* in 1985 and 1991, respectively [32,33] are being used mainly for critically ill patients of all disease categories admitted to the intensive care units (ICUs). They differ in how chronic health status is assessed, in the number of physiologic variables included (12 vs. 17), and in the total score. Specific parameters of liver function (i.e. serum bilirubin and albumin) are included only in the APACHE III scoring system. Some prognostic variables (e.g., prothrombin time) and other indicators of responses to therapy (e.g., blood units transfused) which are known to be important outcome predictors in cirrhotic patients are not measured by the acute physiology scores [17,21,22,26]. APACHE II and III scores have been successfully used to risk stratify cirrhotic patients admitted to medical ICUs [34–38]. APACHE II has been previously used to risk stratify a mixed population of both ICU and non-ICU cirrhotic patients with upper gastrointestinal bleeding [39], while recently, an incomplete APACHE III score (i.e. a score in which data for blood gas analysis were omitted) has been

reported to be superior to CPS in risk stratifying cirrhotic patients outside ICU settings [40].

The aim of the present study was to compare the prognostic accuracy of Child-Pugh, 24 hour APACHE II and complete 24 hour APACHE III scoring systems in predicting hospital mortality of patients with liver cirrhosis admitted to a gastroenterological medical ward.

Methods

This prospective study included two hundred consecutive hospitalizations of 147 patients with liver cirrhosis admitted to the Department of Gastroenterology of the University Hospital of Heraklion, from February 1999 through January 2001. For the purpose of the study, each admission was considered as one patient. The criterion for inclusion was the presence at admission or in the past history of any of the major complications of cirrhosis (ascites, encephalopathy, variceal bleeding or spontaneous bacterial peritonitis). Patients transferred from elsewhere were included in the study only if the transfer occurred within 24 hours after initial admission. Patients with hepatocellular carcinoma and patients admitted for less than one day were excluded. Patients admitted to a medical ICU during the first 24 hours of their presentation were also excluded. The diagnosis of cirrhosis was based on liver biopsy in 93 out of 200 patient admissions (46.5%). For the remaining 107 patient admissions, the diagnosis of cirrhosis was based on clinical, laboratory and radiological criteria: history of portal hypertension excluding other etiologies, evidence of esophageal varices confirmed by endoscopy, splenomegaly, ascites confirmed by abdominal ultrasound and physical examination, impaired liver function tests and clotting profile, ultrasound or computer tomography criteria [39,41].

To calculate the APACHE II score [32], twelve common physiological and laboratory values (temperature, mean arterial pressure, heart rate, respiratory rate, oxygenation (PaO_2 or A-aDo_2), arterial pH, serum sodium, serum potassium, serum creatinine, haematocrit, white blood cell count and Glasgow coma score) are marked from 0 to 4, with 0 being the normal, and 4 being the most abnormal. The sum of these values is added to a mark adjusting for patient age and a mark adjusting for chronic health problems (severe organ insufficiency or immunocompromised patients) to arrive at the APACHE II score.

APACHE III scores range from 0 to 299 and are derived from marks for the extent of abnormality of 17 physiologic measurements (the acute physiology score), adjusts for age, and adjusts for seven comorbidities that reduce immune function and influence hospital survival [33]. The 17 physiological variables include eleven laboratory parameters (haematocrit, white blood cell count, serum

creatinine, serum BUN, serum sodium, serum albumin, serum bilirubin, blood glucose, PaO₂, A-aDO₂, and a scoring for acid-base abnormalities), five vital signs (pulse, mean blood pressure, temperature, respiratory rate, urine output) and a modified Glasgow coma score.

Clinical and laboratory data necessary to the CPS and APACHE systems and prothrombin time (PT) values were recorded on the first day for all patients. Physiological data (temperature, heart rate, mean blood pressure and respiratory rate) were recorded 3-hourly during the first 24 hours of admission. The calculation of APACHE II and III scores was based on the worst values taken during the first 24 hours after admission.

Statistical analysis

Chi-square test was used to assess the differences of mortality within Child-Pugh classes A, B, and C. Individual relationship of each score (CPS, APACHE II, APACHE III) and PT values to the risk of death was assessed by t-test. For the assessment of the magnitude of correlation of length of stay (LOS) with CPS, APACHE II and APACHE III, Pearson correlation was used. Descriptive statistics were expressed as mean \pm SD unless otherwise stated. Discrimination was tested using the receiver operating characteristic (ROC) curves and by comparing areas under the curve (AUCs) [42]. AUCs between 0.7 and 0.8 were classified as "acceptable" and between 0.8 and 0.9 as "excellent" discrimination [43]. For the different scoring systems tested, the sensitivity, specificity, overall correctness of prediction, positive and negative predictive values were calculated, and the cutoff point giving the best Youden index was determined [44]. This cutoff point was also used to calculate the predicted and observed outcome for patients. In order to test the overall classification accuracy of APACHE III score in association with PT, we applied discriminant analysis (backward stepwise method). A *P* value less than 0.05 was considered statistically significant for all above analyses. Calibration was assessed using the Hosmer-Lemeshow goodness of fit statistic which divides subjects into deciles based on predicted probabilities of death and then computes a chi-square from observed and expected frequencies [45]. Lower chi-square values and higher *P* values are associated with a better fit. A good fit was defined as *P* > 0.05.

Results

Of the 200 patient admissions, 137 (68.5%) were men and 63 were women. The mean age was 62.3 years (range, 33–86 years). Eighty eight (44%) admissions were for viral-associated cirrhosis (HBV-associated 40 cases, HCV-associated 48 cases), 66 (33%) for alcoholic cirrhosis and 37 (18.5%) for cryptogenic cirrhosis. In 9 cases (4.5%) there was both viral and alcoholic etiology. The reasons for admission were ascites for 127 patients (63.5%), encephalo-

lopathy for 37 (18.5%), variceal bleeding for 54 (27%), spontaneous bacterial peritonitis for eight (4%), and another infection (respiratory, biliary and urinary tract infections, cellulitis) for 36 patients (18%). Five cases were transferred to a medical ICU. During this study period, 23 patients (11.5%) died. Three patients died in the ICU and 20 patients in the medical ward. The causes of death were liver failure in seven cases (30.4%), kidney failure in two cases (8.7%), hepatorenal syndrome in seven cases (30.4%), variceal bleeding in one case (4.5%), and infection in six cases (26%).

Forty nine cases (24.5%) were classified as Child-Pugh class A, 88 cases (44%) as class B and 63 cases (31.5%) as class C. No deaths were recorded among patients with Child-Pugh class A. Two patients with Child-Pugh class B and 21 with class C died. Mortality increased significantly with increasing Child-pugh classes (*P* < 0.001). Table 1 shows that there were significant differences in CPS, APACHE II score, APACHE III score, and PT between survivors and non-survivors. Table 2 reports predictive values of the various scoring systems calculated at the cutoff point giving the best Youden index. ROC curves are shown in Figure 1. Discrimination power of CPS AUC and APACHE III AUC was excellent, while that of APACHE II AUC was acceptable. When information regarding PT values were combined with APACHE III score into a new discriminant function, the overall classification accuracy of APACHE III was not improved, thus PT was deleted from the full model (non-significant at the 5% level). The results of Hosmer-Lemeshow goodness-of-fit tests are shown in Table 3, while deciles risk are shown in Tables 4, 5 and 6. The Hosmer-Lemeshow statistic was best for CPS. However, for the two APACHE scores, calibration was poor.

The median LOS for survivors was 9 days (range 2–85 days), 7 days (range 2–17 days) for patients with Child-Pugh class A, 9 days (range 2–48 days) for Class B, and 15 days (range 2–85 days) for those with class C. CPS and APACHE III score correlated strongly with the duration of hospitalization (*P* < 0.001), while APACHE II score had a weak and non significant correlation.

Discussion

The performance of the prognostic models is evaluated by their discrimination and calibration. Discrimination (i.e the ability of a prognostic score to classify patients correctly as survivors or non-survivors) is measured by AUC [42,43]. Calibration evaluates the degree of correspondence between the estimated probabilities of mortality produced by a model and the actual mortality experience of patients and can be tested using Hosmer-Lemeshow goodness-of-fit statistic [45].

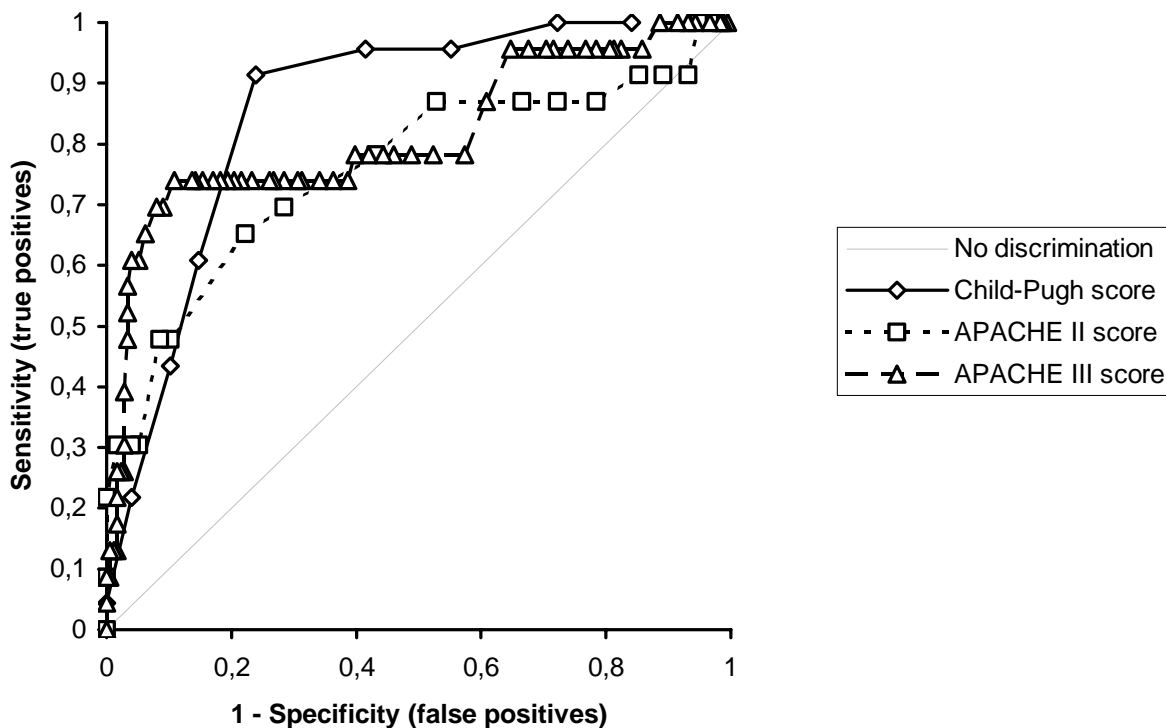


Figure 1
ROC curves for CPS, APACHE II and APACHE III scoring systems

Table 1: Average values on prothrombin time and average scores on Child-Pugh, APACHE II and APACHE III

	Survivors (N = 177)	Nonsurvivors (N = 23)	P value
Child-Pugh score	8.05 ± 2.24	11.13 ± 1.60	< 0.001
APACHE II score	11.45 ± 3.86	15.73 ± 5.24	< 0.001
APACHE III score	41.67 ± 15.13	66.04 ± 21.46	< 0.001
Prothrombin time*	16.95 ± 3.45	21.20 ± 5.15	< 0.001

All data reported as mean ± SD All data were recorded during the first 24 hours after admission to the ward N = number of patients *in seconds

In our series, discrimination was acceptable to excellent for Child-Pugh and APACHE scores, however both APACHE prognostic systems had inadequate goodness-of-fit for death. Our results for APACHE II and CPS discrimination compare well with those published by Afessa et al [39]. In their study, the prognostic value of APACHE II

(AUC 0.78) was as good as that of Child-Pugh score (AUC 0.76) in predicting short-term outcome of 111 cirrhotic patients hospitalized for upper gastrointestinal bleeding, although no informations regarding correct classification rates, sensitivity, specificity, cutoff values and goodness-of-fit have been assessed. The reported APACHE II mean

Table 2: Comparison of the predictive values of the scoring systems

Parameter	Child-Pugh	APACHE II	APACHE III
Cutoff point ^a	10	15	62
Youden index	0.67	0.43	0.63
Sensitivity (%)	91	65	73
Specificity (%)	76	77	89
Correct classification (%)	78	76	87
Positive predictive value (%)	33	27	47
Negative predictive value (%)	98	94	96
AUC (95% CI) ^b	0.859 (0.793 to 0.924)	0.759 (0.704 to 0.928)	0.816 (0.634 to 0.882)

^aValue giving the best Youden index ^bAUC; area under the curve

Table 3: Hosmer-Lemeshow goodness-of-fit tests

	Chi-Square	DF	P value
Child-Pugh score	8.693	6	0.192
APACHE II score	19.157	6	0.004
APACHE III score	23.622	8	0.003

values were higher for both survivors and non-survivors (17.2 ± 6 and 25.6 ± 10.1 respectively). However, 71% of their patients were ICU admissions and 57.6% had active variceal bleeding, thus they have studied much sicker patients than those included in our sample.

Butt et al reported that by using discriminant analysis, APACHE III score correctly classified 75% of cases vs. 67% of cases for Child-Pugh score [40]. No cutoff values were reported, the overall model calibration was not tested and data from blood gas analysis were not included in the calculation of the APACHE III score, thus resulting in an incomplete score. The APACHE III mean values were found high for both survivors and non-survivors (58.9 ± 35.1 and 87.4 ± 30.3 respectively). This might be related to the high percentage of patients admitted with upper gastrointestinal tract bleeding (i.e. 57%). Since four out of five vital signs (pulse, mean blood pressure, respiratory rate, and possibly urine output) and some of the laboratory parameters (i.e. haematocrit, serum BUN, and possibly creatinine) which need to arrive at the APACHE III score are markedly affected by bleeding, this might be the reason of the observed higher scores. Furthermore, the authors did not specify if they have included patients admitted to a medical ICU during the first 24 hours of their admission, whereas patients with hepatocellular carcinoma were also included in the study. The reported mortality on day 1 was 26% and 68% in patients with an APACHE III score of 51 to 75 points and greater than 75 respectively. It is note-

worthy that in our series 17 out of 67 patients (25.3%) with an APACHE III score of 51 to 75 and 6 out of 11 patients (54.5%) with an APACHE III score greater than 75 also died. This suggests that at least in this sub-group of sicker patients our results compare well.

There are many potential reasons for insufficient calibration of APACHE scores. Clinically useful predictive models should demonstrate ease of use, accuracy, reproducibility and acceptance by data collecting staff [46]. Some variables of the APACHE scores (i.e heart rate) depend on continuous monitoring. In addition, it has been shown that the inter-observer variability is high when these scoring systems are not used on a regular basis (like in most non-ICU wards), thus affecting the accuracy and reproducibility of the data [47,48]. This is potentially relevant in our study, since physiological data collection was performed by several physicians and over a long period of time (24 months). As previously suggested [49] we tried to minimize variability by having one person to coordinate the process of data collection and having a written reference of definitions based on the original articles of APACHE scores.

Another potential reason for the inadequate calibration is the differences in level of disease severity between our database and the development databases of the mortality prediction systems [50]. Statistically derived prediction models like the APACHE systems are calibrated to the

Table 4: Tables of deciles risk for Child-Pugh score

Decile	N	PD	AD	PS	AS
1	28	0.2	0	27.8	28
2	21	0.3	0	20.7	21
3	31	0.7	1	30.3	30
4	25	1.1	0	23.9	25
5	32	2.5	1	29.5	31
6	23	3.2	7	19.8	16
7	28	8.7	9	19.3	19
8	12	6.5	5	5.5	7

N = number of patients; PD = predicted to die; AD = actually died; PS = predicted to survive; AS = actually survived

Table 5: Tables of deciles risk for APACHE II score

Decile	N	PD	AD	PS	AS
1	22	0.3	2	21.7	20
2	20	0.5	1	19.5	19
3	21	0.9	0	20.1	21
4	24	1.5	0	22.5	24
5	47	4.3	4	42.7	43
6	37	5.5	5	31.5	32
7	24	7.2	6	16.8	18
8	5	2.9	5	2.1	0

N = number of patients; PD = predicted to die; AD = actually died; PS = predicted to survive; AS = actually survived

Table 6: Tables of deciles risk for APACHE III score

Decile	N	PD	AD	PS	AS
1	20	0.2	0	19.8	20
2	22	0.4	1	21.6	21
3	22	0.6	0	21.4	22
4	26	0.9	4	25.1	22
5	22	1.0	0	21.0	22
6	21	1.5	1	19.5	20
7	20	2.4	0	17.6	20
8	21	4.3	2	16.7	19
9	20	7.5	12	12.5	8
10	6	4.2	3	1.8	3

N = number of patients; PD = predicted to die; AD = actually died; PS = predicted to survive; AS = actually survived

overall outcome prevalence in the development sample. Although APACHE II and III have been shown to work well in cirrhotic patients with a high severity of illness admitted to ICUs [34–38], it is a well known fact that the mortality prediction models performance usually deteriorates when models are applied to different population samples (i.e. less sick patients) [51]. In studies conducted in cirrhotic patients admitted to ICU, cutoff values for

APACHE II have ranged from 17 (AUC 0.69; ICU mortality rate 52%) to 22 (AUC 0.79; ICU mortality rate 36%, hospital mortality 46%) [34,37], while those reported for APACHE III have ranged from 75 (AUC 0.78; ICU mortality rate 43%, hospital mortality 57%) to 80 (AUC 0.75) [34,35]. In our series, APACHE II scores equal or greater than 17 and 22, and APACHE III scores equal or greater than 75 and 80 were recorded in only 25 (12.5%), 5

(2.5%), 11 (5.5%) and 8 (4 %) patients respectively, thus emphasizing the much lower level of disease severity in our patients. It should be also recognized that the wide 95% CI of our AUCs (Table 2) suggests sample size problem, especially when only 23 patients died.

Potential limitations of our study should also be mentioned. Our study was performed in an academic referral hospital; therefore our results may not be applicable to institutions with different patient populations. Because mathematical equations for APACHE III have not been published and for APACHE II this equation is available only for admission, these equations have not been used to calculate the relative risk of death. In agreement with other studies [34,35,37,39,40], we wanted to test the accuracy of single-score values. Patients admitted to a medical ICU during the first 24 hours of their presentation were excluded from our study, thus resulting in a mortality rate of only 11.5%. It could be stated that the rationale of excluding these patients weakens our study, since sicker patients at presentation are more likely to die. However, physiological data included in APACHE III score are recorded 3-hourly during the first 24 hours of admission and the worst value at this time interval is taken into account to calculate the total score [33]. Furthermore, we aimed to define within a 24 hour interval patients not sick enough to be admitted in a medical ICU, but who are likely not to benefit from the standard therapy and for whom a more intensive monitoring and treatment might be tried.

Conclusions

In conclusion, we cannot recommend the use of APACHE II and III scores in non-ICU patients. The present study showed that the discrimination power of CPS AUC and APACHE III AUC was excellent, while that of APACHE II AUC was acceptable. Although the Hosmer-Lemeshow statistic revealed adequate goodness-of-fit for CPS, this was not the case for APACHE II and III scores. Our results indicate that between the three scores, CPS had the least statistically significant discrepancy between predicted and observed mortality across the strata of increasing predicting mortality. This supports the hypothesis that APACHE scores do not work accurately outside ICU settings.

Competing interests

none declared.

References

- Child CG and Turcotte JG **Surgery and portal hypertension** In: *The liver and portal hypertension* (Edited by: Child CG) Philadelphia, Saunders 1964, 50-64
- Gines P, Quintero E, Arroyo V, Teres J, Bruguera M, Rimola A, Caballeri J, Rodes J and Rozman C **Compensated cirrhosis: natural history and prognostic factors** *Hepatology* 1987, **7**:122-128
- Schlichting P, Christensen E, Andersen PK, Fauerholdt L, Juhl E, Poulsen H and Tygstrup N **Prognostic factors in cirrhosis identified by Cox's regression model** *Hepatology* 1983, **3**:889-895
- Adler M, Van Laethem J, Glibert A, Gelin M, Bourgeois N, Vereerstraten P and Cremer M **Factors influencing survival at one year in patients with non biliary hepatic parenchymal cirrhosis** *Dig Dis Sci* 1990, **35**:1-5
- Adler M, Verset D, Bouhdid H, Bourgeois B, Gulbis B, Le Moine O, Vanderstadt J, Gelin M and Thiry P **Prognostic evaluation of patients with parenchymal cirrhosis: Proposal of a new simple score** *J Hepatol* 1997, **26**:642-649
- Merkel C, Bolognesi M, Bellow S, Bianco S, Honisen B, Lampe H, Angeli P and Gatta A **Aminopyrine breath test in the prognostic evaluation of patients with cirrhosis** *Gut* 1992, **33**:836-842
- Kamath PS, Wiesner RH, Malinchoc M, Kremers W, Therneau TM, Kosberg CL, D'Amico G, Dickson ER and Kim WR **A model to predict survival in patients with end-stage liver disease** *Hepatology* 2001, **33**:464-470
- Orrego H, Israel Y, Blake JE and Medline A **Assessment of prognostic factors in alcoholic liver disease: toward a global quantitative expression of severity** *Hepatology* 1983, **3**:896-905
- Pingon JP, Poynard T, Naveau S, Marreau P, Zourabien-Vilio and Chapat JC **Analyse multidimensionnelle selon le modele de Cox dela survie de patients atteints de cirrhose alcoolique** *Gastroenterol Clin Biol* 1986, **10**:461-467
- Pugh RNH, Murray-Lyon IM, Dawson JL, Pietroni MC and Williams R **Transection of the esophagus in bleeding oesophageal varices** *Br J Surg* 1973, **60**:648-652
- Garden OJ, Motyl H, Gilmour WH, Utley RJ and Carter DC **Prediction of outcome following acute variceal hemorrhage** *Br J Surg* 1985, **72**:91-95
- Teres J, Baroni R, Bordas JW, Visa J, Pera C and Rodes J **Randomized trial of portocaval shunt stapling transection and endoscopic sclerotherapy in uncontrolled variceal bleeding** *J Hepatol* 1987, **4**:159-167
- Terblanche J, Northover JM, Bomman P, Kahn D, Silber W, Barbezat GO, Sellars S, Campbell JA and Saunders SJ **A prospective controlled trial of sclerotherapy in the long term management of patients after esophageal variceal bleeding** *Surg Gynecol Obstet* 1979, **148**:323-333
- Rikkers LF, Burnett DA, Volentine GD, Buchi KN and Cormier RA **Shunt surgery versus endoscopic sclerotherapy for long-term treatment of variceal bleeding. Early results of a randomized trial** *Ann Surg* 1987, **206**:201-271
- Sauerbruch T, Ansari H, Wotzka R, Soehendra N and Kopcke W **Prognostic factors in cirrhosis of the liver, variceal bleeding and sclerotherapy: comparison of prognosis systems obtained by discriminant analysis with the Child-classification** *Dtsch Wed Wochenschr* 1988, **113**:11-14
- Harley HAJ, Morgan T, Redeker AG, Reynolds TB, Villamil F, Weiner JM and Yellin A **Results of a randomized trial of end-to-side portocaval shunt and distal splenorenal shunt in alcoholic liver disease and variceal bleeding** *Gastroenterology* 1986, **91**:802-809
- LeMoine O, Adler M, Bourgeois N, Delhayne M, Deviere J, Gelin M, Vandermeeren A, Van Gossom A and Vereerstraten P **Factors related to early mortality in cirrhotic patients bleeding from varices and treated by urgent sclerotherapy** *Gut* 1992, **33**:1381-1385
- Provenzale D, Sandler RS, Wood DR, Levinson SL, Frakes JT, Sartor RB, Jackson AL, Kinard HB, Wagner EH and Powell DW **Development of a scoring system to predict mortality from upper gastrointestinal bleeding** *Am J Med Sci* 1987, **294**:26-32
- Borsch G, Matuk ZE and Leverkus F **Prognosis in upper gastrointestinal bleeding: univariate and multivariate statistical analysis in 447 bleeding episodes** *Med Klin* 1987, **82**:774-780
- Gatta A, Merkel C, Amodio P, Bellon S, Bellumat A, Bolognesi M, Borzato L, Butto M, Casson FF and Cavallarin G **Development and validation of a prognostic index predicting death after upper gastrointestinal bleeding in patients with liver cirrhosis: A multicenter study** *Am J Gastroenterol* 1994, **89**:1528-1536
- Infante-Rivard C, Esnaola S and Villeneuve JP **Clinical and statistical validity of conventional prognostic factors in predicting short-term survival among cirrhotics** *Hepatology* 1987, **7**:660-664
- Ferro D, Saliola M, Quintarelli C, Alessandri C, Basili S and Violi F **I-Year survey of patients with advanced liver cirrhosis: prognostic value of clinical and laboratory indexes identified by the Cox regression model** *Scand J Gastroenterol* 1992, **27**:852-856

23. Hartmann AH, Bircher J and Creutzfeldt W **Superiority of the Child-Pugh classification to quantitative liver function tests for assessing prognosis of liver cirrhosis** *Scand J Gastroenterol* 1989, **24**:269-276
24. Conn HO **Transjugular intrahepatic shunts: the state of the art** *Hepatology* 1993, **17**:148-158
25. Van Buuren HR, Pieterman CH, Schalm SW and Lameris JS **Transjugular intrahepatic portosystemic shunt: requiem for the surgical portosystemic shunt?** *Scand J Gastroenterol* 1993, **28**(suppl 200):48-52
26. Christensen E, Krintel JJ, Hansen SM, Johansen JK and Juhl E **Prognosis after the first episode of gastrointestinal bleeding or coma in cirrhosis** *Scand J Gastroenterol* 1989, **24**:999-1006
27. Dasarathy S, Dwivedi M, Bhargava DK, Sundaram KR and Ramachandran K **A prospective randomized trial comparing repeated endoscopic sclerotherapy and propranolol in decompensated (Child class B and C) cirrhotic patients** *Hepatology* 1992, **16**:89-94
28. Toledo C, Salmeron JM, Rimola A, Navasa M, Arroyo V, Llach J, Gines A, Gines P and Rodes J **Spontaneous bacterial peritonitis cirrhosis: predictive factors of infection resolution and survival in patients treated with cefotaxime** *Hepatology* 1993, **17**:251-257
29. Mathuna PM, Westaby D and Williams R **Taking the tension out of the portal system: an approach to the management of portal hypertension in the 1990s** *Scand J Gastroenterol* 1990, **25**(suppl 175):131-145
30. Bircher J **Assessment of prognosis in advanced liver disease: to score or to measure, that's the question** *Hepatology* 1986, **6**:1036-1037
31. Cello JP, Deveney KE, Trunkey DD, Heilbron DC, Stoney RJ, Ehrenfeld WK and Way LW **Factors influencing survival after therapeutic shunts. Results of a discriminant function and liner logistic regression analysis** *Am J Surg* 1981, **141**:257-265
32. Knaus WA, Draper EA, Wagner DP and Zimmerman JE **APACHE II: a severity of disease classification system** *Crit Care Med* 1985, **13**:818-829
33. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, Sirio CA, Murphy DJ, Lotring T and Damiano A **The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults** *Chest* 1991, **100**:1619-1636
34. Zauner CA, Apsner RC, Kranz A, Kramer L, Madl C, Schneider B, Schneeweiss B, Ratheiser K, Stockenhuber F and Lenz K **Outcome prediction for patients with cirrhosis of the liver in a medical ICU: a comparison of the APACHE scores and liver-specific scoring systems** *Intensive Care Med* 1996, **22**:559-563
35. Zauner C, Schneeweiss B, Schneider B, Madl C, Klos H, Kranz A, Ratheiser K, Kramer L and Lenz K **Short-term prognosis in critically ill patients with liver cirrhosis: an evaluation of a new scoring system** *Eur J Gastroenterol Hepatol* 2000, **12**:517-522
36. Zimmerman JE, Douglas P, Wagner P, Seneff MG, Becker RB, Sun X and Knaus WA **Intensive care unit admissions with cirrhosis: Risk-stratifying patient groups and predicting individual survival** *Hepatology* 1996, **23**:1393-1401
37. Wehler M, Kokoska J, Reulbach U, Hahn EG and Strauss R **Short-term prognosis in critically ill patients with cirrhosis assessed by prognostic scoring systems** *Hepatology* 2001, **34**:255-261
38. Aggarwal A, Ong JP, Younossi ZM, Nelson DR, Hoffman-Hogg L and Arroliga AC **Predictors of mortality and resource utilization in cirrhotic patients admitted to the medical ICU** *Chest* 2001, **119**:1489-97
39. Afessa B and Kubilis PS **Upper gastrointestinal bleeding in patients with hepatic cirrhosis: clinical course and mortality prediction** *Am J Gastroenterol* 2000, **95**:484-489
40. Butt AK, Khan AA, Alam A, Shah SWH, Shafiqat F and Naqvi AB **Predicting hospital mortality in cirrhotic patients: Comparison of Child-Pugh and Acute Physiology, Age and Chronic Health Evaluation (APACHE III) scoring systems** *Am J Gastroenterol* 1998, **93**:2469-2475
41. Brown JJ, Naylor MJ and Yagan N **Imaging of hepatic cirrhosis** *Radiology* 1997, **202**:1-16
42. Hanley JA and McNeil BJ **The meaning and use of the area under a receiver operating characteristic (ROC) curve** *Radiology* 1982, **143**:29-36
43. Hosmer DW and Lemeshow S **Assessing the Fit of the Model** *In Applied Logistic Regression* NY, John Wiley & sons 2000, 160-164
44. Youden WJ **Index for rating diagnostic tests** *Cancer* 1950, **3**:32-35
45. Lemeshow S and Hosmer DW **A review of goodness of fit statistics for use in the development of logistic regression models** *Am J Epidemiol* 1982, **115**:92-106
46. Holt AW, Bury LK, Bersten AD, Skowronski GA and Vedig AE **Prospective evaluation of residents and nurses as severity score data collectors** *Crit Care Med* 1992, **20**:1688-1691
47. Polderman KH, Thijs LG and Girbes AR **Interobserver variability in the use of APACHE II scores** *Lancet* 1999, **353**:380
48. Polderman KH, Joma EM and Girbes AR **Inter-observer variability in APACHE II scoring: effect of strict guidelines and training** *Intens Care Med* 2001, **27**:1465-1369
49. Rue M, Valero C, Quintana S, Artigas A and Alvarez M **Interobserver variability of the measurement of the mortality probability models (MMP II) in the assessment of severity of illness** *Intens Care Med* 2000, **26**:286-291
50. Justice AC, Covinsky KE and Berlin JA **Assessing the generalizability of prognostic information** *Ann Intern Med* 1999, **130**:515-524
51. Hadorn DC, Draper D, Rogers WH, Keeler EB and Brook RH **Cross-validation performance of mortality prediction models** *Stat Med* 1992, **28**:475-489

Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1471-230X/3/7/prepub>

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

