BRIEF ARTICLES

Application of a biochemical and clinical model to predict individual survival in patients with end-stage liver disease

Eduardo Vilar Gomez, Luis Calzadilla Bertot, Bienvenido Gra Oramas, Enrique Arus Soler, Raimundo Llanio Navarro, Javier Diaz Elias, Oscar Villa Jiménez, Maria del Rosario Abreu Vazquez

Eduardo Vilar Gomez, Luis Calzadilla Bertot, Enrique Arus Soler, Department of Hepatology, National Institute of Gastroenterology, Havana 10400, Cuba
Bienvenido Gra Oramas, Department of Pathology, National Institute of Gastroenterology, Havana 10400, Cuba
Raimundo Llanio Navarro, Department of Gastroenterology, National Institute of Gastroenterology, Havana 10400, Cuba
Javier Diaz Elias, Deparment of Gastroenterology, The "Calixto Garcia" Hospital, Havana 10400, Cuba
Oscar Villa Jiménez, Department of Gastroenterology, National Institute of Gastroenterology, Havana 10400, Cuba
Maria del Rosario Abreu Vazquez, Department of Biostatistics, National Institute of Gastroenterology, Havana 10400, Cuba
Author contributions: Gomez EV and Bertot LC contributed equally to this work; They performed the study, acquisition of data, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript and statistical analysis; Oramas BG, Soler EA, Elias JD, Jiménez OV performed the design, acquisition of data, and analysis and interpretation of data; Navarro RL performed critical revision of the manuscript; Abreu Vazquez MR performed the statistical analysis.
Correspondence to: Eduardo Vilar Gomez, National Institute of Gastroenterology, 25th Avenue, 503, Vedado, Havana 10400, Cuba. vilar@infomed.sld.cu
Telephone: +53-7-8325067 Fax: +53-7-8333253
Received: February 11, 2009 Revised: May 1, 2009
Accepted: May 8, 2009
Published online: June 14, 2009

Abstract

AIM: To investigate the capability of a biochemical and clinical model, BioCliM, in predicting the survival of cirrhotic patients.

METHODS: We prospectively evaluated the survival of 172 cirrhotic patients. The model was constructed using clinical (ascites, encephalopathy and variceal bleeding) and biochemical (serum creatinine and serum total bilirubin) variables that were selected from a Cox proportional hazards model. It was applied to estimate 12-, 52- and 104-wk survival. The model's calibration using the Hosmer-Lemeshow statistic was computed at 104 wk in a validation dataset. Finally, the model's validity was tested among an independent set of 85 patients who were stratified into 2 risk groups (low risk ≤ 8 and high risk > 8).

RESULTS: In the validation cohort, all measures of fit, discrimination and calibration were improved when the biochemical and clinical model was used. The proposed model had better predictive values (c-statistic: 0.90, 0.91, 0.91) than the Model for End-stage Liver Disease (MELD) and Child-Pugh (CP) scores for 12-, 52- and 104-wk mortality, respectively. In addition, the Hosmer-Lemeshow (H-L) statistic revealed that the biochemical and clinical model (H-L, 4.69) is better calibrated than MELD (H-L, 17.06) and CP (H-L, 14.23). There were no significant differences between the observed and expected survival curves in the stratified risk groups (low risk, $P = 0.61$; high risk, $P = 0.77$).

CONCLUSION: Our data suggest that the proposed model is able to accurately predict survival in cirrhotic patients.

INTRODUCTION

The Model for End-stage Liver Disease (MELD) and Child-Pugh (CP) scores have been the most widely applied prognostic markers for organ allocation in liver transplantation, mainly due to their simplicity of use in daily clinical practice[1-5]. The MELD score has gained wide acceptance for predicting survival in patients undergoing liver transplantation. It has been suggested

that it provides more accurate prognosis than the Child-Pugh (CP) score in patients with decompensated cirrhosis and that it therefore improves the evaluation of priority for liver graft allocation[4,5]. It is not surprising, however, that the magnitude of superiority of the MELD score over the CP score is modest and is primarily limited to the population at the highest risk of renal failure[6]. Additionally, changes in some objective laboratory parameters of the MELD score may be directly related to the extensive use of diuretics, volume status, albumin infusion and the patient's nutritional status. Finally, clinical complications of portal hypertension such as ascites, encephalopathy, spontaneous bacterial peritonitis (SBP) and gastrointestinal bleeding are not considered in the MELD score, probably underestimating any direct association with the severity of liver disease[7]. However, the model has been shown to predict mortality independent of the occurrence of complications of portal hypertension[3,4]. The classification applied to the clinical complications of portal hypertension (ascites, encephalopathy, variceal bleeding and SBP) in the MELD score does not clearly reveal the different grades of severity of liver disease and its clinical response to medical treatment. Therefore, its utility as a prognostic model could be limited. In this regard, several recent studies have shown that clinical manifestations secondary to portal hypertension (encephalopathy, ascites) are good prognostic markers in cirrhotic patients[8,9]. According to the results of these studies, the use of clinical markers in prognostic models may be recommended.

The aim of this study was to evaluate the short-, medium- and long-term prognosis of a series of cirrhotic patients by means of the BioCliM score using biochemical (creatinine and bilirubin) and clinical (encephalopathy, bleeding esophageal varices and ascites) variables, to compare BioCliM with the MELD and CP scores, and to identify those variables with liver-related mortality. Our model was developed to improve accuracy in predicting survival and consequently improve the further evaluation of priority for liver graft allocation in cirrhotic patients.

MATERIALS AND METHODS

Study design, setting and participants
We prospectively evaluated 180 consecutive cirrhotic patients who were admitted at the National Institute of Gastroenterology of Havana during the period May 2003 to January 2006. Inclusion criteria were histological, laparoscopic or clinical diagnosis of cirrhosis and presence of compensated or decompensated disease (stages A, B or C according to the CP classification). Patients with hepatocellular carcinoma, severe infection, severe primary cardiopulmonary failure, alcohol use within one month before initial evaluation, and intrinsic kidney disease were excluded from the study. Among 180 patients who had complete medical profiles and an established diagnosis of hepatic cirrhosis, 172 patients fulfilled the above selection criteria.

The model validation was performed by applying it to

Table 1 Baseline characteristics of the patient population		
Variables	**Derivation set** $n = 172$	**Validation set** $n = 85$
Follow-up period	56 (4-104)	58 (8-104)
Age (yr)	56 (20-79)	59 (23-78)
Sex, *n* (%)		
Male	106 (62)	58 (68)
Female	66 (38)	27 (32)
Cause of cirrhosis, *n* (%)		
Alcohol	30 (17)	16 (19)
Alcohol plus viral infection	15 (9)	4 (5)
HBV	20 (12)	12 (14)
HCV	92 (53)	50 (59)
Viral co-infection (HBV/HCV)	1 (1)	1 (1)
Unknown	13 (7)	1 (1)
NAFL	1 (1)	1 (1)
Complications on admission, *n* (%)		
Ascites		
Absent or controlled	147 (85)	72 (85)
Uncontrolled	25 (15)	13 (15)
BEV		
Absent or present without relapses	167 (97)	79 (94)
Present with relapses	5 (3)	5 (6)
Encephalopathy		
Absent or controlled	160 (93)	79 (93)
Uncontrolled	12 (7)	6 (7)
SBP		
Absent or present without relapses	168 (98)	82 (96)
Present with relapses	4 (2)	3 (4)
Hepatorenal syndrome, *n* (%)	3 (2)	1 (1)
Prothrombin time (s)	19 (13-55)	17 (13-53)
Partial thromboplastin time (s)	38 (26-165)	39 (26-167)
INR for prothrombin time	1.7 (1-7.5)	1.5 (1-6.9)
Albumin (g/L)	37 (20-48)	36 (21-47)
Creatinine (mmol/L)[1]	100 (42-516)	98 (39-489)
Bilirubin (mmol/L)[2]	20 (8-130)	23 (12-137)
Cholesterol (mmol/L)	3.8 (1.9-10.2)	3.9 (2-9.6)
Child-Pugh score[3]	7 (5-14)	7 (5-14)
Child-Pugh A, *n* (%)	67 (39)	30 (35)
Child-Pugh B, *n* (%)	75 (44)	34 (40)
Child-Pugh C, *n* (%)	30 (17)	21 (25)
MELD score	17 (9-42)	18 (10-43)
BioCliM score	7.7 (6.1-13.6)	7.9 (6-13.8)

HBV: Hepatitis B virus; HCV: Hepatitis C virus; NAFL: Non-alcoholic fatty liver; BEV: Bleeding esophageal varices; SBP: Spontaneous bacterial peritonitis; INR: International normalized ratio; MELD: Model for End-stage Liver Disease; BioCliM: Biochemical and Clinical Model. All quantitative variables are expressed as median (ranges). [1]To convert mmol/L into mg/dL, multiply by 0.01131. [2]To convert mmol/L into mg/dL, multiply by 0.0585. [3]The Child-Pugh, MELD and BioCliM scores are measures of the severity of liver disease.

an independent group of 85 patients who were evaluated at the "Calixto Garcia" Hospital of Havana from March 2005 to August 2007. The baseline characteristics of the patient population are summarized in Table 1.

Variables of interest, measurement, follow-up and ethics
Detailed medical history, complete physical examination, and a battery of laboratory tests were performed in all patients on the day of admission. Biochemical evaluations were carried out by the same laboratory. Prothrombin time expressed as PT-ratio (patient-to-normal coagulation time) was converted to prothrombin time international normalized ratio (INR) using an internal laboratory standard and was assessed by a single

operator. The main clinical complications of portal hypertension were initially evaluated and classified by an experienced hepatologist depending on the clinical response to medical treatment. Bleeding esophageal varices (BEV) were diagnosed by clinical signs of hematemesis and endoscopic signs of active bleeding or adherent clots on EV[10]; they were classified as absent, present with relapses or rebleeding (2 or more bleeding episodes in the last 3 mo) or without relapses (one bleeding episode in the last 3 mo). Variceal bleeding relapse or rebleeding was defined as the occurrence of hematemesis/melena, aspiration of more than 100 mL of fresh blood in patients with a nasogastric tube and decrease of 3 g in Hb if no transfusion was given. Portosystemic encephalopathy was defined according to the West Haven criteria for grading from 0 (subclinical) to 4 (coma)[11]; it was classified as absent (no episode of encephalopathy in the last year), medically controlled (episodic hepatic encephalopathy developing over hours to days, but does not persist with adequate medical treatment) or uncontrolled (persistent hepatic encephalopathy that develops upon discontinuation of medication), irrespective of disease severity. Ascites was classified as absent (no clinical and ultrasound evidence of ascites and without therapeutic intervention), medically controlled (no clinical and ultrasound evidence of ascites in patients undergoing full therapeutic intervention) or uncontrolled (ascites that requires repeated paracentesis for control or a sodium-restricted diet and intensive diuretic therapy). Diagnostic paracentesis and ascitic fluid culture were performed in all admitted cirrhotic patients. Spontaneous bacterial peritonitis (SBP) was diagnosed when the ascites polymorphonuclear leukocyte (PMN) count was $> 250/mm^3$, with or without positive ascites bacterial culture[12]; it was coded as absent, present without relapses (one SBP episode in the last year) and present with relapses (2 or more episodes in the last year). Patients were followed up from their date of initial evaluation until death (related or unrelated to liver disease), liver transplantation, or study closure. Patients with death unrelated to liver disease were excluded from the analysis. Patients lost to follow-up were censored at the last date known to be alive and patients undergoing liver transplantation were censored at the transplant date.

The study was conducted in compliance with the Declaration of Helsinki and approved by the ethics committee and the institutional review board of the National Institute of Gastroenterology. All patients provided written informed consent for participation.

Analysis for survival and derivation of the novel risk score

The probable prognostic predictors, including age, sex, serum biochemistry and clinical complications of portal hypertension, were analyzed to determine prognostic ability. To lessen the influence of extreme laboratory values, quantitative variables were transformed to their natural logarithms.

Univariate and multivariate forward stepwise Cox proportional hazards models were used to determine variables associated with survival. Variables that were significant ($P < 0.05$) in univariate analysis were included in multivariate analysis. Stepwise probabilities for entry or removal were set at 0.05 and 0.10, respectively.

With each Cox model, a risk score for each patient was calculated as: $R = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$, where X_1, $X_2, \ldots, X_k$ are the values of prognostic factors and $\beta_1, \beta_2, \ldots, \beta_k$ are the corresponding regression coefficients. A higher risk score corresponds to poorer prognosis.

The forward stepwise selection procedures were used for variable selection, assessment for interactions, and model development. The likelihood ratio statistic tested the significance of the addition of each variable separately to a predictive model that included ascites only. Furthermore, the c-statistic was computed as a criterion for the selection of a group of variables to be used in the new predictive model. The final criterion for inclusion in the model was minimization of the Bayes Information Criterion (BIC)[13]. The BIC is a likelihood-based measure in which lower values indicate better fit and in which a penalty is paid for increasing the number of variables. Thus, the variables selected for inclusion should provide not only the best fit but also a parsimonious prediction model.

Predictive models for survival and discrimination

The CP and MELD scores were calculated on parameters obtained at referral. The MELD score was calculated according to the original formula proposed by the Mayo Clinic group as follows: [$9.57 \times \log_e$ creatinine mg/dL + $3.78 \times \log_e$ bilirubin mg/dL + $11.20 \times \log_e$ INR + 6.43 (constant for liver disease etiology)]. To avoid negative scores, laboratory values less than 1 were rounded up to 1. The maximal value of creatinine was 4 mg/dL[3].

Once a new risk model was determined, it was prospectively tested in the validation dataset of 85 patients from "Calixto Garcia" Hospital. The discrimination ability of the different models was measured by means of the concordance statistic (c-statistic), a measure of discrimination also known as a natural extension of the receiving operator characteristic (ROC) curve area in survival analysis. P values for the comparison of the c-statistic were computed using the bootstrap method. A c-statistic between 0.8 and 0.9 indicates excellent accuracy, and a value over 0.7 should be considered clinically useful. The concordance c-statistic was assessed for 12-, 52-, and 104-wk survival. A time-to-event with the censored data version for survival analysis was performed to compute the c-statistic.

The concordance probability estimates (CPE) were computed[14], because the c-statistic seems to overestimate the true concordance probability, especially if the censoring proportion is high. Since the CPE is a consistent estimate, it is a better measure in the context of using predictions from Cox regression models.

Calibration and external validation of the new risk score

To assess model calibration (or how closely the predicted

probabilities reflect actual risk), the Hosmer-Lemeshow calibration statistic, as modified by D'Agostino *et al*[15], comparing observed and predicted risk was implemented.

In addition, the new risk model was validated in a cohort of 85 independent patients from "Calixto Garcia" Hospital who were stratified into 2 risk (R) groups: R $\leq$ 8 and R > 8. Within each risk group the survival was calculated using the Kaplan-Meier procedure and the observed-predicted survivals were compared using the log-rank test.

Analyses were performed with the use of SAS software, version 9.1 (SAS Institute).

RESULTS

A total of 180 patients were examined for eligibility, and 172 were included in the study. The reasons for non-participation were: 3 patients with HCC, 3 with repeated alcohol use, and 2 with severe infection disease. The period of recruitment lasted from May 2003 to February 2004. One hundred and forty one patients completed the follow-up period. Thirty one patients died during the study, 29 liver-related and 2 unrelated to liver disease (myocardial infarction). One hundred and seventy patients were included in the outcome analyses.

The patients' clinical and serological features are summarized in Table 1.

In the derivation data set, the median follow-up period was 56 wk (range, 4-104 wk). The CP median score was 7 (range, 5-14) with 61% of the patients being CP class B and C. The MELD and BioCliM median scores were 17 (range, 8-42) and 7.7 (range, 5.7-13.6), respectively. During follow-up, 29 patients (17%) died. The 4-, 12-, 24-, 52- and 104-wk survival rates were 98%, 98%, 90%, 89% and 83%, respectively.

The patients of the validation group were followed for a median of 58 wk (range, 8–104 wk) during which 13 died. The 4-, 12-, 24-, 52- and 104-wk survival rates were 96%, 95%, 88%, 84% and 83%, respectively. The CP median score was 7 (range, 5-14) with 65% of the patients being CP class B and C. The MELD and BioCliM median scores were 18 (range, 9-43) and 7.9 (range, 6-13.8), respectively.

None of the patients in the derivation or validation groups underwent liver transplantation during the follow-up period.

Overall survival according to single prognostic factors
Univariate analysis for 104-wk overall survival:
Univariate analysis using Cox proportional hazards models showed that serum levels of creatinine, bilirubin, cholesterol, albumin, prothrombin time, partial thromboplastin time, ascites, spontaneous bacterial peritonitis, encephalopathy and bleeding esophageal varices were significantly associated with survival (Table 2).

Multivariate analysis for 104-wk overall survival:
Multivariate Cox regression analysis included those variables independently related to survival resulting from

Table 2 Association of baseline characteristics with mortality in 170 cirrhotic patients, results from univariate Cox proportional hazards models

Variables	*P*	Hazard ratio	95% CI for Hazard ratio
Age (yr)	0.68	0.56	0.36-1.06
Sex (male)	0.54	0.44	0.89-1.26
Etiology (viral)	0.66	0.58	0.40-1.11
ALT (IU/L) (log$_n$ value)	0.85	0.84	0.34-1.12
AST (IU/L) (log$_n$ value)	0.43	0.90	0.56-1.34
ALT/AST ratio	0.64	0.87	0.50-1.21
Platelet count ($\times 10^9$/L) (log$_n$ value)	0.54	0.89	0.52-1.30
Prothrombin time (s)[1] (log$_n$ value)	0.01	2.23	1.24-4.89
INR for prothrombin time (log$_n$ value)	0.03	1.99	1.13-3.96
Partial thromboplastin time (s)[2] (log$_n$ value)	0.04	1.78	1.10-3.23
Albumin (mg/dL) (log$_n$ value)	0.001	3.12	1.89-5.23
Bilirubin (mmol/L) (log$_n$ value)	< 0.001	3.89	2.12-6.14
Creatinine (mmol/L) (log$_n$ value)	< 0.001	3.95	2.18-6.56
Cholesterol (mmol/L) (log$_e$ value)	0.03	1.83	1.34-3,42
Ascites	< 0.001	4.05	2.27-6.33
Spontaneous bacterial peritonitis	0.001	3.05	2.10-5.07
Encephalopathy	< 0.001	4.50	2.90-6.50
Bleeding esophageal varices	< 0.001	4.78	3.11-7.11

CI: Confidence interval; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase. Hazard ratios (95% CI) for quantitative variables are expressed for 1 relevant unit increase of log. INR: International normalized ratio. [1]Prothrombin time (s): Value in seconds. [2]Partial thromboplastin time (s): Value in seconds.

univariate analysis. The selected variables were available in all patients that entered the forward stepwise model. Of the candidate variables, only ascites, encephalopathy, bleeding esophageal varices and serum creatinine were independently predictive of survival (Table 3).

The estimated hazard risk for ascites suggested that the risk of death for uncontrolled ascites was 10.2 times greater than for those with absent or controlled ascites. The risk of death in those patients with relapsing bleeding and uncontrolled encephalopathy increased 3.25 times compared to those without bleeding or with non-relapsing bleeding, and 2.5 times compared to those with absent or controlled encephalopathy. In terms of impact in prognosis, the ascites (hazard ratio (HR), 10.2) and serum creatinine (HR, 3.99) were the most important prognostic factors.

Model derivation and development
In the model derivation cohort, 11 potential variables selected from the univariate analysis (*P* < 0.05) were calculated for model inclusion. Of these only 5 were included in the model. The likelihood ratio statistic showed the significance of the addition of each variable separately to a predictive model that included ascites only (Table 3). The χ^2 statistic was progressively increased with the addition of creatinine, bleeding esophageal varices, hepatic encephalopathy, and bilirubin. The c-statistic in the model that included only ascites was 0.76, based on the c-statistic for censored data. When creatinine, BEV, HE and bilirubin were added to the model, the c-statistic was improved to 0.83, 0.85, 0.89, and 0.90, respectively. In the same context,

Table 3 Contributions of different variables to survival prediction at 104 wk, results from multivariate Cox regression models[1]

Variable	Variable χ^2	Regression coefficient	Hazard ratio	95% CI for Hazard ratio		P value	c-statistic	BIC
Ascites	53.90	2.310	10.2	3.78	28.1	< 0.0001	0.76	2014.15
+ Ln (creatinine)	63.43	1.370	3.99	1.57	10.9	0.006	0.83	1988.15
+ BEV	65.71	1.195	3.25	1.01	9.77	0.048	0.85	1970.65
+ HE	68.91	0.909	2.50	0.915	6.88	0.070	0.89	1961.89
+ Ln (bilirubin)[2]	70.11	0.349	1.46	0.66	3.33	0.427	0.90	1951.77

[1]Estimated from Cox proportional hazards models. [2]Biochemical (bilirubin and creatinine) and Clinical (ascites, encephalopathy and bleeding esophageal varices) Model; BEV: Bleeding esophageal varices; HE: Hepatic encephalopathy; BIC: Bayesian Information Criterion. Ln was used to normalize distributions and improve the fit for individual predictors. Hazard ratio for quantitative variables are expressed for 1 relevant unit increase of log. + indicates the addition of each variable separately to the model with ascites only. χ^2 is the likelihood ratio statistic for each group of variables when added to the model. The risk prediction was based on data from the model derivation cohort (n = 170) at 104 wk follow-up.
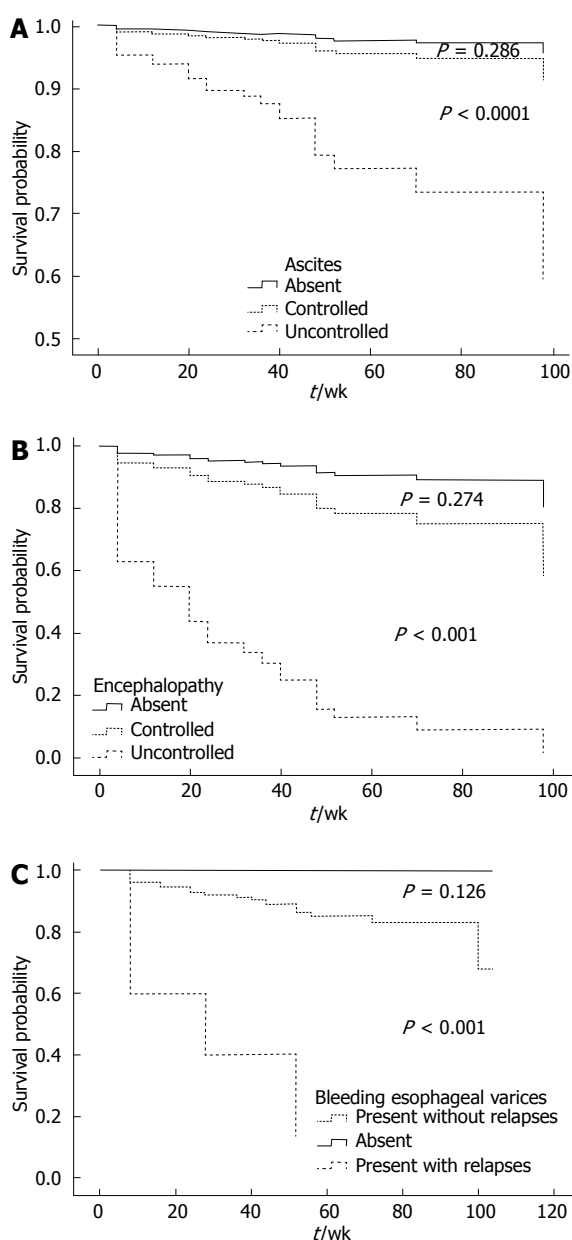

Figure 1 Kaplan-Meier estimated survival curves for clinical variables. A: Ascites; B: Encephalopathy; C: Bleeding esophageal varices.

the combination of ascites, creatinine, BEV, HE and bilirubin revealed the smallest BIC value (1951.77), thus,

in the derivation set, the model with the combination of clinical and biochemical variables appeared to improve the risk prediction.

Computational formula for 104-wk risk using best-fitting model

The regression coefficients of the formula for calculating the new risk score (biochemical and clinical model) were selected from a Cox regression model[16] and are reported in Table 3.

The risk scores for individual patients were calculated using the following equation: [1.370 × log$_e$ (creatinine mmol/L) + 0.349 × log$_e$ (bilirubin mmol/L) + 2.310 × (ascites: 0 if absent or medically controlled and 1 if uncontrolled) + 0.909 × (encephalopathy: 0 if absent or medically controlled and 1 if uncontrolled) + 1.195 × (bleeding esophageal varices: 0 if absent or present without relapses and 1 if present with relapses). The clinical variables were coded depending on the clinical response to medical treatment. The variables grouped together as "absent or medically controlled" (ascites and encephalopathy) and "absent or present without relapses" (bleeding esophageal varices) have been so grouped because their survival was similar in each one of them (Figure 1). The missing values were imputed for survival modeling.

Survival probabilities were derived from the Cox proportional hazards model: $S(t) = S_0(t)^{\exp(R-R_o)}$. $S(t)$ is the survival probability in wk, $S_0(t)$ the baseline survival function, R the individual risk score and R_0 the risk score of the average patient in the series. For example, the 12-wk survival probability is calculated as: $S_{(12\ wk)} = 0.981^{\exp(BioCliM\ score-7)}$, where 0.981 is the 12-wk baseline survival and 7 is the reference BioCliM score. To ease its use, the score was multiplied by 100.

Predictive models for 12-, 52- and 104-wk survival

Comparison of the c-statistic values among the CP, MELD and BioCliM scores was performed. All scoring systems were found to have diagnostic accuracy in predicting survival. The BioCliM score, however, showed to have better discriminative power in predicting short- (12 wk), intermediate- (52 wk) and long-term survival (104 wk) than the rest of the scores (Figure 2).
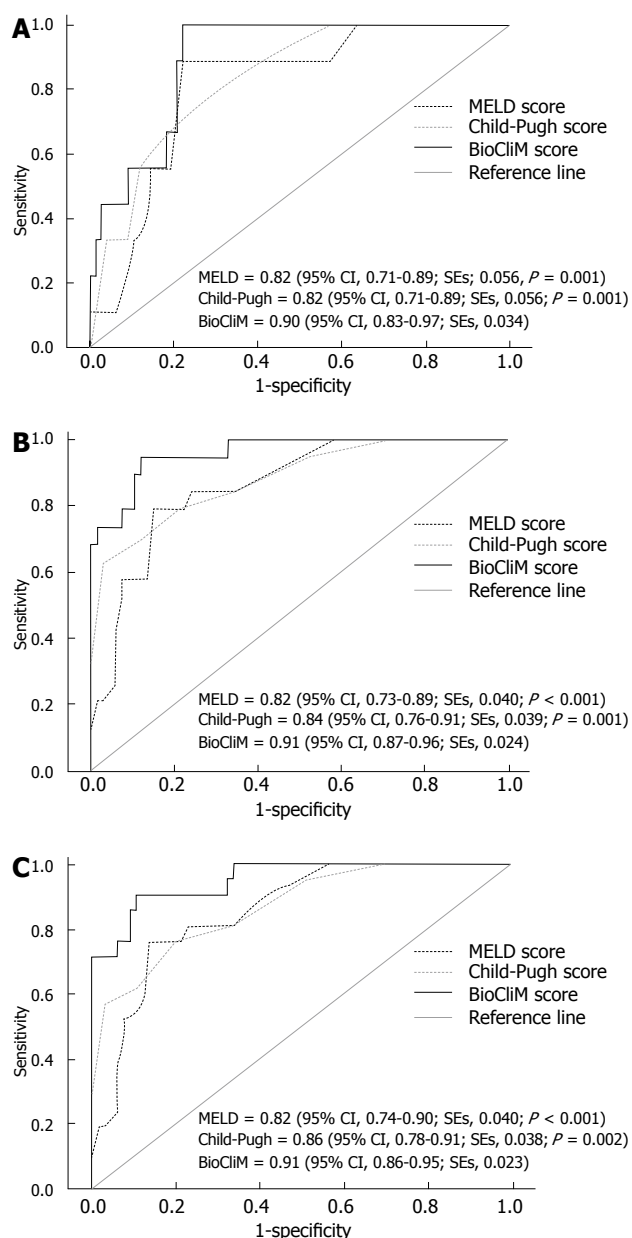
Figure 2 Comparison of the c-index values of the MELD, Child-Pugh and BioCliM scores for 12- (A), 52- (B) and 104-wk (C) survival. SE indicates standard errors. The different values were compared with BioCliM score using the bootstrap method.

The c-statistic for the CP and MELD scores were almost identical for 12-wk survival (0.82 and 0.82), and slightly higher for CP as compared with MELD for 52-wk (0.84 and 0.82) and 104-wk (0.86 and 0.82) survival.

We used an alternative way of computing the concordance probability for a censored outcome to estimate the true concordance probability in samples with a high censored proportion. The concordance probability estimates for the CP (CPE, 0.71; SE, 0.042), MELD (CPE, 0.74; SE, 0.043) and BioCliM (CPE, 0.78; SE, 0.050) models were lower at 12 wk in comparison with those obtained using the standard c-statistic value. Finally, the CPE at 12 wk was consistently higher for BioCliM as compared with CP and MELD scores.

Discrimination and model validation

The Hosmer-Lemeshow statistic (H-L) is a measure of the discrepancy between the observed and predicted risk. A better calibrated model would have a smaller discrepancy between the observed and predicted and thus a smaller H-L statistic.

A significant *P* value for the H-L statistic indicates a significant deviation between predicted and observed outcomes. Figure 3 compares the calibration of the BioCliM, MELD and CP scores in predicting the probability of death at 104 wk. The H-L statistic was 4.69 for the BioCliM score, 17.06 for the MELD score and 14.23 for the CP score, indicating a good calibration for all models; however, this analysis clearly shows that BioCliM is better calibrated.

Figure 4 illustrates the observed and expected Kaplan-Meier survival curves for each score in 2 patient subgroups divided according to risk score as low risk (R ≤ 8) and high risk (R > 8), selected from the "Calixto Garcia" Hospital. Using a cutoff value of 8 (risk score) to predict probability of survival within 104 wk, the sensitivity and specificity of the BioCliM score was 90% and 87%, respectively. Median survival was 104 wk and 47 wk for low- and high-risk groups, respectively. There were no significant differences between the observed and expected survival curves in the stratified risk groups (low risk, *P* = 0.61; high risk, *P* = 0.77). Thus, the BioCliM score allowed accurate prediction of survival in the cirrhotic patient validation group.

Survival according to the BioCliM score

The differences in the short-, intermediate- and long-term survival between patients with low risk (≤ 8), and high risk (> 8) scores were compared (Figure 5).

Overall survival rates were significantly different between low-risk and high-risk patients (*P* < 0.0001). The 12-wk survival rates were 98% and 64% for low and high risk, respectively. For low and high risk, 1-year survival rates were 97% and 3%, and 2-year survival rates were 95% and 0%, respectively. Patients with a high risk score had the highest risk of mortality compared to patients with low values. Patients with a BioCliM score of ≥ 8 had a median survival of < 47 wk in comparison to patients with a median survival of 104 wk for patients with a BioCliM score of < 8.

DISCUSSION

The most widely used prognostic model to predict survival in cirrhotic patients has been the CP score. It is an important tool for the prognostic evaluation of cirrhotic patients and the current organ allocation policy. It has, however, several drawbacks such as the subjectivity of clinical parameters, limited discriminative capability and variability in the measurements of laboratory parameters[17,18]. Current CP score modifications by adding new variables or utilizing sophisticated measures did not improve its accuracy to predict survival[19-25]. A relatively new score, the MELD,
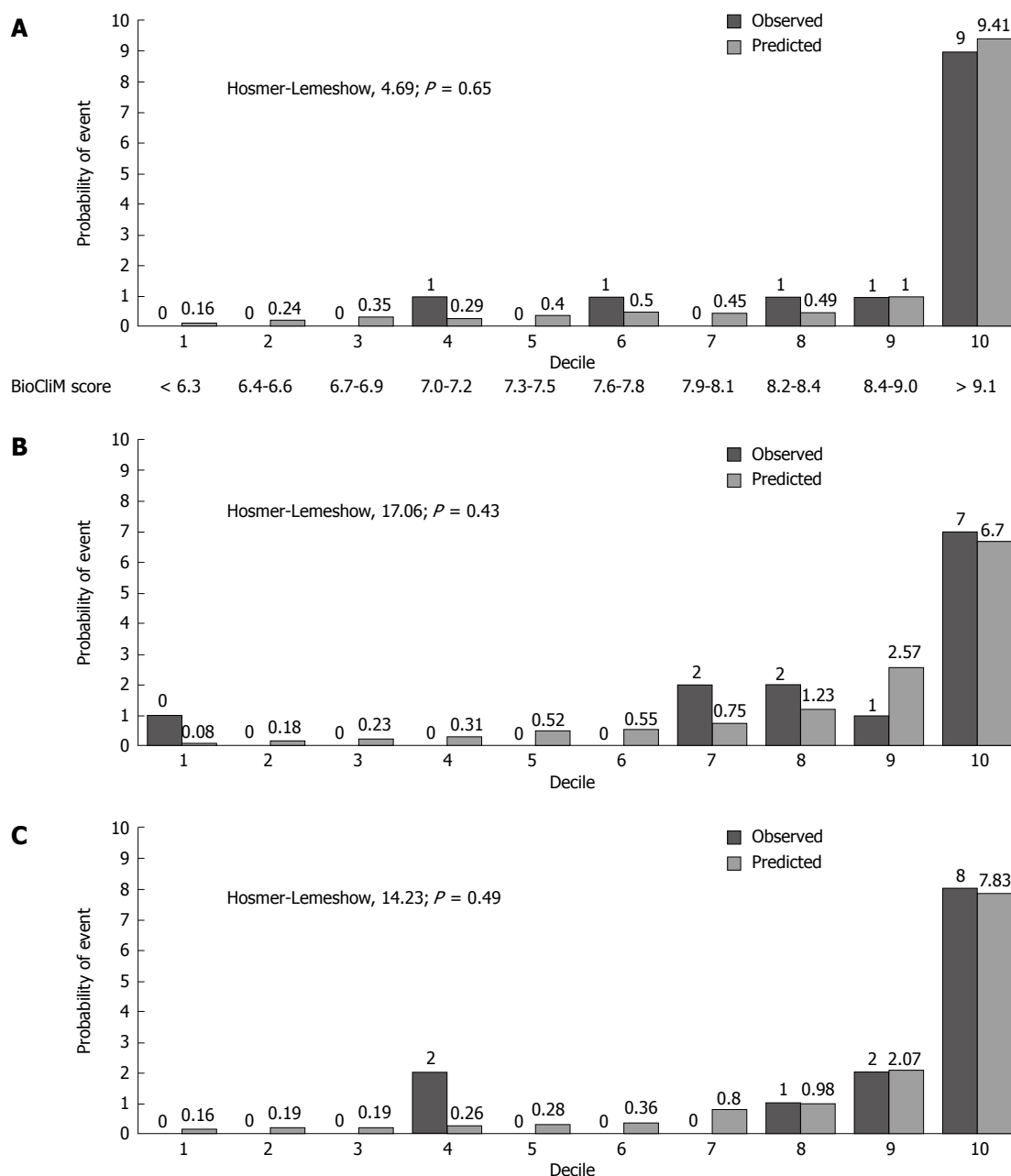
Figure 3 Observed and predicted probability of events at 104 wk. A, B and C shows the observed and predicted probability of death according to BioCliM, MELD and Child-Pugh scores in 10 groups (deciles) of patients, respectively. A significant *P*-value for the Hosmer-Lemeshow statistic indicates a significant deviation between predicted and observed outcomes.

has been instituted in patients with end-stage liver disease awaiting liver transplantation. MELD has shown an advantage over CP by using continuous objective variables that are not open to observer interpretation and are appropriately weighted according to their impact on prognosis[3,4,26]. Its ability to predict mortality, however, has been found to be similar or slightly superior to the traditional CP score[27-30]. These controversies suggest that a better predictive model is necessary to predict survival in cirrhotic patients.

In our study, the baseline characteristics were comparable with similar studies evaluating survival in cirrhotic patients[31-36]. Furthermore, all clinical and biochemical variables included in the CP and MELD scores were associated with survival in univariate

analysis. Multivariate Cox proportional hazards analysis identified serum creatinine, ascites, encephalopathy and bleeding esophageal varices as independent prognostic factors for overall survival. The strongest predictors of mortality were ascites and serum creatinine. In our proposed model, ascites, encephalopathy and variceal bleeding were evaluated depending on medical treatment response, and the diagnosis and treatment of each of these was based on the most recent published guidelines[10-12,37]. The used nomenclature appeared to be more uniform and less subjective than the commonly applied classification into CP or MELD scores[2-4].

The major finding of this prospective study is that the BioCliM score, which is based on a combination of 3 clinical indices (ascites, encephalopathy and bleeding
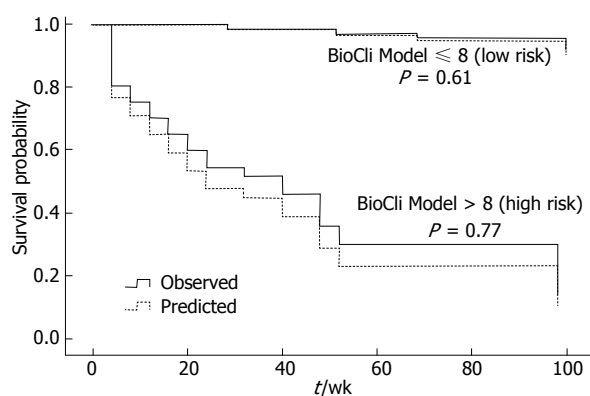
Figure 4 Observed and expected 104-wk survival curves for the BioCliM score. Survival of 85 independent patients from the "Calixto Garcia" Hospital who were stratified according to their risk score into two risk groups (low-risk ≤ 8 and high-risk > 8). The observed and predicted BioCli Model survival curves were compared using log-rank test. The observed and expected survival was similar for the low- and high-risk groups.
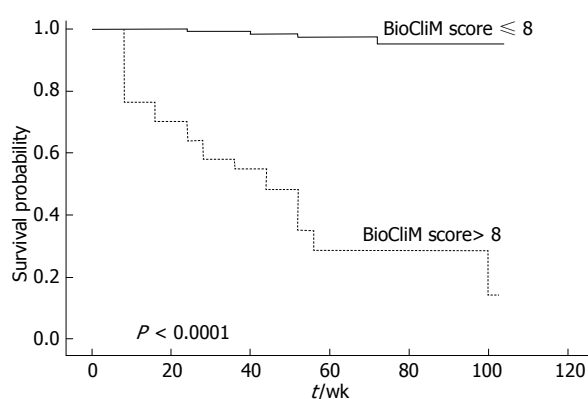

Figure 5 Kaplan-Meier estimated survival curves for the BioCli Model score.

esophageal varices) and 2 biochemical parameters (creatinine and bilirubin), is able to accurately predict short-term (12 wk), intermediate-term (52 wk) and long-term (104 wk) mortality in cirrhotic patients. Our results showed that the BioCliM score is superior to the CP and MELD scores in ranking patients according to their risk of death. In addition, the BioCliM score showed a sustained discriminative power to predict survival through the different evaluated periods (12-104 wk). Our data further support, as well as previous findings, that the MELD score is not significantly superior to the CP score in predicting survival in patients with hepatic cirrhosis[25,27-30]. Theoretically, the MELD score is undoubtedly more objective and robust than the CP score for the previously mentioned reasons; however, a major limitation of the MELD score is the poor discriminative power to predict survival among patients whose clinical course is often affected by other factors which are excluded by the model[38]. Recent studies have demonstrated that ascites, encephalopathy and hyponatremia are important independent predictors of early pretransplant mortality, especially for patients with low MELD scores[8,9,39,40], thus affecting the consideration for an expedited liver transplant under

the "sickest first" model. In consequence, as the MELD score does not reflect the presence of ascites and encephalopathy, these patients need to be allocated separately for liver transplantation if MELD is used to prioritize organ allocation. By contrast, the BioCliM scale is able to accurately predict survival in patients with clinical complications of portal hypertension, thus the BioCliM score could be recommended in the individual management of these patients. Further studies are needed to validate its prognostic accuracy in patients undergoing liver transplantation.

Possibly the most important study limitations were the relatively small sample size, the poor geographic diversity of the patients included (single and tertiary center) and the major drawbacks of the MELD score related to wide variability of laboratory parameters such as serum creatinine and bilirubin[28,41,42].

In conclusion, both the CP and MELD scores can accurately predict short-term survival in cirrhotic patients, while the BioCliM score appears to have great discriminative power for short- (4 and 12 wk), intermediate- (24 and 52 wk) and long-term (104 wk) survival. In contrast to the MELD score, the use of the BioCliM score in patients with ascites, encephalopathy and variceal bleeding could significantly increase survival predictive values in patients with end-stage liver disease.

ACKNOWLEDGMENTS

COMMENTS

Background
The Child-Pugh and Model for End-stage Liver Disease (MELD) scores are important tools for the prognostic evaluation of cirrhotic patients and the current organ allocation policy. These have, however, several drawbacks such as the subjectivity of clinical parameters, limited discriminative capability and variability in the measurements of laboratory parameters. The current evidence suggests that a better predictive model is necessary to predict survival in cirrhotic patients.

Research frontiers
The clinical complications of portal hypertension such as ascites, encephalopathy, spontaneous bacterial peritonitis (SBP) and gastrointestinal bleeding are not considered in the MELD score, probably underestimating that they may have a direct association with the severity of liver disease. The classification applied to the clinical complications of portal hypertension (ascites, encephalopathy, variceal bleeding and SBP) in the MELD score does not clearly reveal the different grades of severity of liver disease and its clinical response to medical treatment. Therefore, its utility in the prognostic model could be limited. In this study, the authors have evaluated a new paradigm for clinical variables, depending on the severity and medical treatment response and how they have an influence, as prognostic factors, in the survival of cirrhotic patients.

Innovations and breakthroughs
Recent reports have demonstrated that ascites, encephalopathy and hyponatremia are important independent predictors of early pretransplant mortality, especially for patients with low MELD scores, thus affecting the consideration for an expedited liver transplantation under the "sickest first" model. In consequence, as the MELD score does not reflect the presence of

ascites and encephalopathy, these patients need to be allocated separately for liver transplantation if MELD is used to prioritize organ allocation. By contrast, the new biochemical and clinical model is able to accurately predict survival in patients with clinical complications of portal hypertension; thus the BioCliM score could be recommended in the individual management of these patients.

Applications

In contrast to the MELD score, BioCliM is able to accurately predict survival in patients with clinical complications of portal hypertension, thus the BioCliM score could be recommended in the individual management of these patients. Further studies are needed to validate its prognostic accuracy in patients undergoing liver transplantation.

Terminology

BioCliM is a new biochemical and clinical model that is able to accurately predict survival in patients with end-stage liver disease.

Peer review

The authors examined the prognostic value and predictive capability of a new prognostic model in patients with end-stage liver disease. A less subjective nomenclature to assess the clinical complications of portal hypertension was evaluated in combination with biochemical variables to determine their influence as prognostic factors of survival in cirrhotic patients. The biochemical and clinical model was shown to accurately predict survival in patients with clinical complications of portal hypertension and it appeared to have great discriminative power for short- (4 and 12 wk), intermediate- (24 and 52 wk) and long-term (104 wk) survival.

REFERENCES

1   **Child CG**, Turcotte JG. Surgery and portal hypertension. In: Child CG, editor. The Liver and Portal Hypertension. Philadelphia: W.B. Saunders Co, 1964: 1-85

2   **Pugh RN**, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. *Br J Surg* 1973; **60**: 646-649

3   **Kamath PS**, Wiesner RH, Malinchoc M, Kremers W, Therneau TM, Kosberg CL, D'Amico G, Dickson ER, Kim WR. A model to predict survival in patients with end-stage liver disease. *Hepatology* 2001; **33**: 464-470

4   **Wiesner R**, Edwards E, Freeman R, Harper A, Kim R, Kamath P, Kremers W, Lake J, Howard T, Merion RM, Wolfe RA, Krom R. Model for end-stage liver disease (MELD) and allocation of donor livers. *Gastroenterology* 2003; **124**: 91-96

5   **Freeman RB Jr**, Wiesner RH, Harper A, McDiarmid SV, Lake J, Edwards E, Merion R, Wolfe R, Turcotte J, Teperman L. The new liver allocation system: moving toward evidence-based transplantation policy. *Liver Transpl* 2002; **8**: 851-858

6   **Everson GT**. MELD: the answer or just more questions? *Gastroenterology* 2003; **124**: 251-254

7   **Reuben A**. Child comes of age. *Hepatology* 2002; **35**: 244-245

8   **Yoo HY**, Edwin D, Thuluvath PJ. Relationship of the model for end-stage liver disease (MELD) scale to hepatic encephalopathy, as defined by electroencephalography and neuropsychometric testing, and ascites. *Am J Gastroenterol* 2003; **98**: 1395-1399

9   **Heuman DM**, Abou-Assi SG, Habib A, Williams LM, Stravitz RT, Sanyal AJ, Fisher RA, Mihas AA. Persistent ascites and low serum sodium identify patients with cirrhosis and low MELD scores who are at high risk for early death. *Hepatology* 2004; **40**: 802-810

10 **D'Amico G**, De Franchis R. Upper digestive bleeding in cirrhosis. Post-therapeutic outcome and prognostic indicators. *Hepatology* 2003; **38**: 599-612

11 **Ferenci P**, Lockwood A, Mullen K, Tarter R, Weissenborn K, Blei AT. Hepatic encephalopathy--definition, nomenclature, diagnosis, and quantification: final report of the working party at the 11th World Congresses of Gastroenterology, Vienna, 1998. *Hepatology* 2002; **35**: 716-721

12 **Rimola A**, García-Tsao G, Navasa M, Piddock LJ, Planas R, Bernard B, Inadomi JM. Diagnosis, treatment and prophylaxis of spontaneous bacterial peritonitis: a consensus document. International Ascites Club. *J Hepatol* 2000; **32**: 142-153

13 **Harrell FE Jr**. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer Verlag, 2001

14 **Gönen M**, Heller G. Concordance probability and discriminatory power in proportional hazards regression. *Biometrika* 2005; **92**: 965-970

15 **D'Agostino RB**, Nam BH. Evaluation of the performance of survival analysis models: discrimination and calibration measures. In: Balakrishnan N, Rao CR, editors. Advances in survival analysis. Amsterdam: Elsevier, 2004: 1-25

16 **Christensen E**. Multivariate survival analysis using Cox's regression model. *Hepatology* 1987; **7**: 1346-1358

17 **Forman LM**, Lucey MR. Predicting the prognosis of chronic liver disease: an evolution from child to MELD. Mayo End-stage Liver Disease. *Hepatology* 2001; **33**: 473-475

18 **Pagliaro L**. MELD: the end of Child-Pugh classification? *J Hepatol* 2002; **36**: 141-142

19 **Degré D**, Bourgeois N, Boon N, Le Moine O, Louis H, Donckier V, El Nakadi I, Closset J, Lingier P, Vereerstraeten P, Gelin M, Adler M. Aminopyrine breath test compared to the MELD and Child-Pugh scores for predicting mortality among cirrhotic patients awaiting liver transplantation. *Transpl Int* 2004; **17**: 31-38

20 **Adler M**, Verset D, Bouhdid H, Bourgeois N, Gulbis B, Le Moine O, Van de Stadt J, Gelin M, Thiry P. Prognostic evaluation of patients with parenchymal cirrhosis. Proposal of a new simple score. *J Hepatol* 1997; **26**: 642-649

21 **Merkel C**, Bolognesi M, Bellon S, Bianco S, Honisch B, Lampe H, Angeli P, Gatta A. Aminopyrine breath test in the prognostic evaluation of patients with cirrhosis. *Gut* 1992; **33**: 836-842

22 **Testa R**, Valente U, Risso D, Caglieris S, Giannini E, Fasoli A, Botta F, Dardano G, Lantieri PB, Celle G. Can the MEGX test and serum bile acids improve the prognostic ability of Child-Pugh's score in liver cirrhosis? *Eur J Gastroenterol Hepatol* 1999; **11**: 559-563

23 **Giannini E**, Botta F, Fumagalli A, Malfatti F, Testa E, Chiarbonello B, Polegato S, Bellotti M, Milazzo S, Borgonovo G, Testa R. Can inclusion of serum creatinine values improve the Child-Turcotte-Pugh score and challenge the prognostic yield of the model for end-stage liver disease score in the short-term prognostic assessment of cirrhotic patients? *Liver Int* 2004; **24**: 465-470

24 **Angermayr B**, Koening F, Cejna M, Karnel F, Gschwantler M, Ferenci P. Creatinine-modified Child-Pugh score (CPSP) compared with MELD-score to predict survival in patients undergoing TIPS. *Hepatology* 2002; **36**: 860A

25 **Papatheodoridis GV**, Cholongitas E, Dimitriadou E, Touloumi G, Sevastianos V, Archimandritis AJ. MELD vs Child-Pugh and creatinine-modified Child-Pugh score for predicting survival in patients with decompensated cirrhosis. *World J Gastroenterol* 2005; **11**: 3099-3104

26 **Malinchoc M**, Kamath PS, Gordon FD, Peine CJ, Rank J, ter Borg PC. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. *Hepatology* 2000; **31**: 864-871

27 **Schepke M**, Roth F, Fimmers R, Brensing KA, Sudhop T, Schild HH, Sauerbruch T. Comparison of MELD, Child-Pugh, and Emory model for the prediction of survival in patients undergoing transjugular intrahepatic portosystemic shunting. *Am J Gastroenterol* 2003; **98**: 1167-1174

28 **Angermayr B**, Cejna M, Karnel F, Gschwantler M, Koenig F, Pidlich J, Mendel H, Pichler L, Wichlas M, Kreil A, Schmid M, Ferlitsch A, Lipinski E, Brunner H, Lammer J, Ferenci P, Gangl A, Peck-Radosavljevic M. Child-Pugh versus MELD score in predicting survival in patients undergoing transjugular intrahepatic portosystemic shunt. *Gut* 2003; **52**: 879-885

29 **Ferral H**, Gamboa P, Postoak DW, Albernaz VS, Young CR, Speeg KV, McMahan CA. Survival after elective transjugular intrahepatic portosystemic shunt creation: prediction with

model for end-stage liver disease score. *Radiology* 2004; **231**: 231-236

30   **Cholongitas E**, Papatheodoridis GV, Vangeli M, Terreni N, Patch D, Burroughs AK. Systematic review: The model for end-stage liver disease--should it replace Child-Pugh's classification for assessing prognosis in cirrhosis? *Aliment Pharmacol Ther* 2005; **22**: 1079-1089

31   **Ginés P**, Quintero E, Arroyo V, Terés J, Bruguera M, Rimola A, Caballería J, Rodés J, Rozman C. Compensated cirrhosis: natural history and prognostic factors. *Hepatology* 1987; **7**: 122-128

32   **Cooper GS**, Bellamy P, Dawson NV, Desbiens N, Fulkerson WJ Jr, Goldman L, Quinn LM, Speroff T, Landefeld CS. A prognostic model for patients with end-stage liver disease. *Gastroenterology* 1997; **113**: 1278-1288

33   **Christensen E**, Schlichting P, Fauerholdt L, Gluud C, Andersen PK, Juhl E, Poulsen H, Tygstrup N. Prognostic value of Child-Turcotte criteria in medically treated cirrhosis. *Hepatology* 1984; **4**: 430-435

34   **Infante-Rivard C**, Esnaola S, Villeneuve JP. Clinical and statistical validity of conventional prognostic factors in predicting short-term survival among cirrhotics. *Hepatology* 1987; **7**: 660-664

35   **Zoli M**, Cordiani MR, Marchesini G, Iervese T, Labate AM, Bonazzi C, Bianchi G, Pisi E. Prognostic indicators in compensated cirrhosis. *Am J Gastroenterol* 1991; **86**: 1508-1513

36   **Salerno F**, Borroni G, Moser P, Badalamenti S, Cassarà L, Maggi A, Fusini M, Cesana B. Survival and prognostic factors of cirrhotic patients with ascites: a study of 134 outpatients. *Am J Gastroenterol* 1993; **88**: 514-519

37   **Runyon BA**. Management of adult patients with ascites due to cirrhosis. *Hepatology* 2004; **39**: 841-856

38   **Cholongitas E**, Marelli L, Shusang V, Senzolo M, Rolles K, Patch D, Burroughs AK. A systematic review of the performance of the model for end-stage liver disease (MELD) in the setting of liver transplantation. *Liver Transpl* 2006; **12**: 1049-1061

39   **Said A**, Williams J, Holden J, Remington P, Gangnon R, Musat A, Lucey MR. Model for end stage liver disease score predicts mortality across a broad spectrum of liver disease. *J Hepatol* 2004; **40**: 897-903

40   **Huo TI**, Wu JC, Lin HC, Lee FY, Hou MC, Lee PC, Chang FY, Lee SD. Evaluation of the increase in model for end-stage liver disease (DeltaMELD) score over time as a prognostic predictor in patients with advanced cirrhosis: risk factor analysis and comparison with initial MELD and Child-Turcotte-Pugh score. *J Hepatol* 2005; **42**: 826-832

41   **Cholongitas E**, Marelli L, Kerry A, Senzolo M, Goodier DW, Nair D, Thomas M, Patch D, Burroughs AK. Different methods of creatinine measurement significantly affect MELD scores. *Liver Transpl* 2007; **13**: 523-529

42   **Trotter JF**, Brimhall B, Arjal R, Phillips C. Specific laboratory methodologies achieve higher model for endstage liver disease (MELD) scores for patients listed for liver transplantation. *Liver Transpl* 2004; **10**: 995-1000

S- Editor Li LF   **L- Editor** Cant MR   **E- Editor** Yin DH