Special Article

A Model to Predict Survival in Patients With End-Stage Liver Disease

PATRICK S. KAMATH,[1] RUSSELL H. WIESNER,[1] MICHAEL MALINCHOC,[2] WALTER KREMERS,[2] TERRY M. THERNEAU,[2] CATHERINE L. KOSBERG,[1] GENNARO D'AMICO,[3] E. ROLLAND DICKSON,[1] AND W. RAY KIM[1,2]

SEE EDITORIAL ON PAGE 473

A recent mandate emphasizes severity of liver disease to determine priorities in allocating organs for liver transplantation and necessitates a disease severity index based on generalizable, verifiable, and easily obtained variables. The aim of the study was to examine the generalizability of a model previously created to estimate survival of patients undergoing the transjugular intrahepatic portosystemic shunt (TIPS) procedure in patient groups with a broader range of disease severity and etiology. The Model for End-Stage Liver Disease (MELD) consists of serum bilirubin and creatinine levels, International Normalized Ratio (INR) for prothrombin time, and etiology of liver disease. The model's validity was tested in 4 independent data sets, including (1) patients hospitalized for hepatic decompensation (referred to as "hospitalized" patients), (2) ambulatory patients with noncholestatic cirrhosis, (3) patients with primary biliary cirrhosis (PBC), and (4) unselected patients from the 1980s with cirrhosis (referred to as "historical" patients). In these patients, the model's ability to classify patients according to their risk of death was examined using the concordance (c)-statistic. The MELD scale performed well in predicting death within 3 months with a c-statistic of (1) 0.87 for hospitalized patients, (2) 0.80 for noncholestatic ambulatory patients, (3) 0.87 for PBC patients, and (4) 0.78 for historical cirrhotic patients. Individual complications of portal hypertension had minimal impact on the model's prediction (range of improvement in c-statistic: <.01 for spontaneous bacterial peritonitis and variceal hemorrhage to ascites: 0.01-0.03). The MELD scale is a reliable measure of mortality risk in patients with end-stage liver disease and suitable for use as a disease severity index to determine organ allocation priorities. (HEPATOLOGY 2001;33: 464-470.)

In April of 1998, the Department of Health and Human Services issued the final rule on the Organ Procurement and Transplantation Network, in which the principles of organ allocation policies and procedures were defined.[1] These included allocating organs among transplant candidates in the order of medical urgency and minimizing the role of waiting time while avoiding futile transplantation and promoting efficient management of organ placement. A subsequent report from the Institute of Medicine (IOM)[2] and an independent study by Freeman and Edwards,[3] in which waiting times in transplant candidates were analyzed, concluded that waiting time was not an appropriate measure of the fairness of the organ allocation system. The report from the IOM recommended that the use of waiting time as an allocation criterion be discontinued and that an appropriate medical triage system be developed to ensure equitable allocation of organs based on medical characteristics and disease prognoses rather than waiting times.[2]

While the current allocation system utilizes the Child-Turcotte-Pugh (CTP) classification for determination of medical urgency,[4] implementation of the proposed new policies will require a more refined scale that accurately represents disease severity. A disease severity index for such a purpose should not only have a sound statistical and clinical validity, but should also rely on a few, readily available, objective parameters and be generalizable to a heterogeneous group of patients.[5] Thus, validation of the severity index should incorporate an assessment of its ability to assess the risk of death in independent groups of patients with varying etiology and severity of liver disease as well as geographical diversity.[6]

In this report, we examine the validity of the Model for End-Stage Liver Disease (MELD) as a disease severity index for patients with end-stage liver disease awaiting liver transplantation. This model was originally developed to assess the short-term prognosis of patients with cirrhosis undergoing the transjugular intrahepatic portosystemic shunt (TIPS) procedure.[7] While its development was based on a highly selected subgroup of patients with cirrhosis, we show in this study that the model is able to provide a reliable estimate of short-term survival over a wide range of liver disease severity and diverse etiology. Figure 1 summarizes our approach in the validation of the MELD scale. Starting with patients undergoing the TIPS procedure, the model's validity was tested with data obtained from increasingly heterogeneous patient populations ranging from patients hospitalized with advanced end-stage liver disease, to ambulatory patients with noncholestatic cirrhosis, to patients with primary biliary cirrhosis (PBC), and to an unselected group of heterogeneous cirrhotic patients from a time

Abbreviations: IOM, Institute of Medicine; CTP, Child-Turcotte-Pugh; MELD, Model for End-Stage Liver Disease; TIPS, transjugular intrahepatic portosystemic shunt; PBC, primary biliary cirrhosis; INR, International Normalized Ratio; CI, confidence interval; SBP, spontaneous bacterial peritonitis; ISI, International Sensitivity Index.

From the [1]Division of Gastroenterology and Hepatology, and the [2]Department of Health Science Research, Mayo Clinic and Foundation, Rochester, MN; and the [3]Divisione di Medicina, Ospedale V Cervello, Palermo, Italy.

Received November 13, 2000; accepted December 10, 2000.

Supported by a grant from the National Institutes of Health (DK-34238).

The data presented in this manuscript have been presented and discussed at the OPTN/UNOS Liver and Intestinal Organ Transplantation Committee.

Address reprint requests to: W. Ray Kim, M.D., M.B.A., Gastroenterology and Hepatology (Ch10), Mayo Clinic and Foundation, 200 First Street, SW, Rochester, MN 55905. E-mail: kim.woong@mayo.edu; fax: 507-266-2810.

0270-9139/01/3302-0021$35.00/0

doi:10.1053/jhep.2001.22172

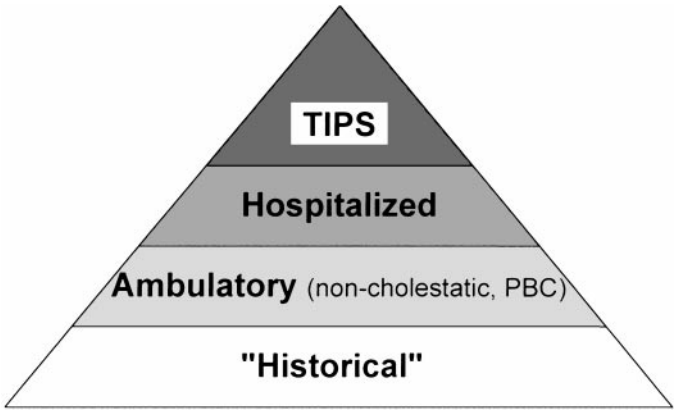

FIG. 1. Data sets used in the validation (total n = 2,278). While the model was developed in patients with end-stage liver disease undergoing the TIPS procedure, the model was validated in a large number of patients with a progressively wider spectrum of severity and etiology of liver disease.

when liver transplantation was not widely available. Based on these validity data, we propose that the MELD scale fulfills the requirement suggested by the IOM as an index by which to allocate organs for liver transplantation.

WHAT IS THE MELD SCALE?

We have previously reported a model that uses serum creatinine, total serum bilirubin, International Normalized Ratio (INR) for prothrombin time, and etiology of cirrhosis, which accurately predicted survival in patients with cirrhosis undergoing the TIPS procedure.[7] This model was derived from a heterogeneous group of patients at 4 medical centers in the United States and validated in an independent data set from the Netherlands. Because survival following portosystemic shunts is predominantly determined by the severity of the underlying liver disease,[8] we hypothesized that the same model could be used as a prognostic indicator in patients with advanced chronic liver disease in general and liver transplant candidates in particular.

The MELD score is a slight modification of the risk score used in the original TIPS model. For ease of use, the score was multiplied by 10 and then rounded to the nearest integer. Thus, the formula for the MELD score is $3.8*\log_e$(bilirubin [mg/dL]) + $11.2*\log_e$(INR) + $9.6*\log_e$(creatinine [mg/dL]) + 6.4*(etiology: 0 if cholestatic or alcoholic, 1 otherwise). An on-line worksheet is available over the Internet at www.mayo.edu/int-med/gi/model/mayomodl.htm.

WHAT CRITERIA DETERMINE THE VALIDITY OF THE MODEL?

Because the aim of this study was to validate the MELD scale as a liver disease severity index in determining short-term survival in liver transplant candidates, we chose 3-month survival as the primary outcome measure. In other words, the main study question was whether the model is able to rank patients according to their risk of death within 3 months. We also evaluated the model's validity in assessing very short-term (1 week) and longer-term (1 year) survival.

The mathematical measure to determine the validity of the model was the concordance (c)-statistic (equivalent to the area under receiver-operating-characteristic curve).[9] This statistic may range from 0 to 1, with 1 corresponding to perfect discrimination and 0.5 to what is expected by chance alone. For example, in determining the risk of death within 3 months in 2 individuals, if the one with a higher score dies before the one with lower score 100% of the time, the c-statistic will be 1.0. If, on the other hand, the prediction is correct only 50% of the time (*i.e.,* the same as a coin toss), the c-statistic is 0.5. A c-statistic of 0 would result if the prediction were wrong 100% of the time.

This statistic is used commonly in evaluating a diagnostic test. A c-statistic between 0.8 and 0.9 indicates excellent diagnostic accuracy and a c-statistic greater than 0.7 is generally considered as a useful test. For prognostic models, a c-statistics of 0.9 or greater is seldom seen.

DOES THE MODEL PREDICT MORTALITY IN HOSPITALIZED PATIENTS WITH CIRRHOSIS NOT UNDERGOING THE TIPS PROCEDURE?

Because the MELD scale was developed in patients undergoing the TIPS procedure, the first study we conducted was to examine the model's usefulness in patients with decompensated cirrhosis not undergoing the TIPS procedure ("hospitalized" data set in Fig. 1).

All cirrhotic patients 18 years or older, hospitalized at the Mayo Clinic between January 1994 and January 1999 were identified using a computerized diagnostic index. Individual hospital records were reviewed to verify the diagnosis of cirrhosis and extract information including patient demographic data, liver disease diagnosis, portal hypertensive complications, and laboratory data. The characteristics of these patients are summarized in Table 1. Patients with concurrent hepatocellular carcinoma, alcohol use within 1 month of hospitalization, advanced cardiopulmonary comorbidity, sepsis, or intrinsic renal disease and those who were hospitalized for liver transplantation were excluded. This yielded the final sample of 282 patients hospitalized for complications of liver disease.

Patient survival was assessed as the interval from the day of hospitalization until death or last follow-up. The median length of follow-up was 1.4 (range, 0-5.6) years. There were 129 deaths, 59 of which occurred during the first 3 months. The c-statistic for prediction of 3-month survival by the MELD score was 0.87 (95% confidence interval [CI], 0.82-0.92) (Table 2). Table 3 shows the relationship between the MELD score and 3-month mortality.

These data also allowed computation of the CTP score, which was also useful in classifying patients by their 3-month survival. The 3-month mortality in patients with CTP class A was 4%, for CTP class B it was 14%, and for class C it was 51%. The c-statistic associated with the CTP score in the prediction of 3-month survival was 0.84 (95% CI, 0.78-0.90).

DOES THE MODEL PREDICT MORTALITY IN AMBULATORY PATIENTS NOT REQUIRING HOSPITALIZATION?

Because the model was developed in patients with decompensated liver disease requiring the TIPS procedure and because, as described in the previous section, we were satisfactorily reassured that the model performed well among patients who were hospitalized for cirrhotic complications but did not require the TIPS procedure, we next examined the model's validity in ambulatory patients with cirrhosis. The issue that we wanted to address was whether the model is calibrated for evaluation of far-advanced patients, making it insensitive in patients not requiring hospitalization care.

TABLE 1. Patient Characteristics

	Hospitalized N = 282	Ambulatory Noncholestatic N = 491	Ambulatory PBC N = 326	Historical N = 1,179
Demographic				
Sex (% male)	55	53*	11	51
Age (yr)				
Median (range)	61 (50-69)	53 (44-59)*	49 (41-56)	60 (48-68)
Cause of cirrhosis (%)				
Alcoholic (%)	30	3	—	26
PBC (%)	13	0	100	18
PSC (%)	2	0	—	0
Viral hepatitis (%)	21	89	—	56†
Other (%)	33	8	—	
Laboratory parameters				
Serum bilirubin (mg/dL)	1.5 (0.8-3.2)	1 (1-1)	1.4 (0.8-3.5)	1.7 (0.9-3.5)
Serum creatinine (mg/dL)	1.0 (0.9-1.3)	1 (1-1)	0.9 (0.8-1.0)	0.9 (0.8-1.1)
INR for prothrombin time	1.3 (1.1-1.5)	1.4 (1.2-1.7)	1.1 (1.0-1.1)	1.2 (1.0-1.5)
MELD	9 (4-14)	10 (8-13)	1 (−2-5)	7 (3-13)

NOTE. Values reported are median (25th percentile-75th percentile).
* Age and gender were recorded in 74 patients.
† Postnecrotic cirrhosis.

Ambulatory Patients With Noncholestatic Liver Disease. Between June 1981 and June 1984, 491 consecutive patients at the Ospadale V Cervello in Palermo, Italy newly diagnosed with cirrhosis were included in a study to investigate the natural history of cirrhosis. This cohort consisted of ambulatory patients predominantly with viral hepatitis (noncholestatic ambulatory patients in Fig. 1). Detailed patient characteristics are shown in Table 1. Prothrombin expressed as percent activity was converted to prothrombin time INR using standard formula.[10]

After a median follow-up of 3.0 (range, 0.1-3.0) years, 117 patients had died, 34 of whom died in the first 3 months. The c-statistic in the MELD scale's prediction of 3-month mortality was 0.80 (95% CI, 0.69 to 0.90) (Table 2). Table 3 shows 3-month death rates by the MELD score.

Ambulatory Patients With PBC. We further evaluated the performance of the MELD score in PBC patients. This analysis is based on a data set that was used to create the Mayo model (n = 418).[11,12] These patients were accrued at the Mayo Clinic between 1973 and 1984. The diagnosis of PBC was based on established clinical, biochemical, serologic, and histologic criteria. Of these 418 patients, 92 lacked variables necessary for the MELD score, leaving 326 for analysis. The characteristics of the study subjects at baseline are described in Table 1 (ambulatory PBC data set in Fig. 1).

Subsequent to the initial enrollment in the study, patients were followed prospectively. Of the 326 patients, 127 or 39% died without liver transplantation, whereas 25 received liver transplantation. Because patients in this data set had early stage disease, there were only 5 deaths in the first 3 months. The c-statistic for the MELD score's prediction of 3-month mortality was 0.87 (95% CI, 0.71-1.00). There were a total of 23 deaths during the first year of follow-up. The c-statistic for 1-year mortality remained at 0.87 (95% CI, 0.80-0.93) (Table 2). This compares favorably to the Mayo PBC model, which has a c-statistic of 0.91 (95% CI, 0.83-0.99) in this group.

IS THE MODEL GENERALIZABLE TO PATIENTS WITH DIVERSE ETIOLOGY AND SEVERITY OF LIVER DISEASE?

To further validate the model in patients with diverse etiology and severity of liver disease while minimizing the impact of liver transplantation in altering the natural history and survival of patients with cirrhosis, we retrospectively generated a cohort of patients with cirrhosis diagnosed between 1984 and 1988 ("historical" data set in Fig. 1). This time period was chosen because it was the earliest time that laboratory data were electronically available at the Mayo Clinic.

A computerized institutional diagnostic index was used to identify patients with diagnostic codes that corresponded to cirrhosis in the following categories: postnecrotic cirrhosis, alcohol-induced cirrhosis, and primary biliary cirrhosis. Patients 17 years or younger or those who denied research authorization (<2%) were excluded. This process resulted in 5,125 potential cases.

TABLE 2. Validity of MELD in Predicting Mortality

	Hospitalized N = 282	Ambulatory Noncholestatic N = 491	Ambulatory PBC N = 326	Historical N = 1179
1-Week mortality	0.95 (0.88-1.00)	0.80 (0.67-0.94)	—	0.84 (0.78-0.89)
3-Month mortality	0.87 (0.82-0.92)	0.80 (0.69-0.90)	0.87 (0.71-1.00)	0.78 (0.74-0.81)
1-Year mortality	0.85 (0.80-0.90)*	0.78 (0.70-0.85)	0.87 (0.80-0.93)	0.73 (0.69-0.76)†

NOTE. Values reported are the concordance statistic (95% CI).
* N = 257; 25 patients lost to follow-up.
† N = 1,108; 71 patients lost to follow-up.

TABLE 3. Three-Month Death Rates

MELD	≤9	10-19	20-29	30-39	≥40
Hospitalized	4% (6/148)	27% (28/103)	76% (16/21)	83% (5/6)	100% (4/4)
Ambulatory noncholestatic	2% (5/213)	5.6% (14/248)	50% (15/30)	—	—
Ambulatory PBC	1% (3/308)	13% (2/16)	0% (0/2)	—	—
Historical	8% (55/711)	26% (90/344)	56% (47/84)	66% (23/35)	100% (5/5)

We screened these patients with the following laboratory criteria to select those who were potentially at a risk of death in the subsequent 3 months: total serum bilirubin concentration ≥2.0 mg/dL and serum albumin concentration ≤3.5 g/dL. While this process yielded 2,906 patients, an additional 1,469 were excluded because they did not have all the parameters necessary to compute the MELD score within a 7-day period. Finally an additional 258 were excluded because of lack of follow-up to determine the 3-month survival.

Hence, 1,179 patients were available for analysis. Of these, there were 220 deaths within 3 months after the laboratory values were measured. The c-statistic for prediction of 3-month survival by the MELD score was 0.78 (95% CI, 0.74-0.81) (Table 2). The mortality at 3 months is shown in Table 3.

DO INDIVIDUAL COMPLICATIONS OF PORTAL HYPERTENSION ADD TO THE MELD SCALE?

A large volume of literature indicates that portal hypertensive complications such as ascites,[13] encephalopathy,[14] variceal bleeding,[15] and spontaneous bacterial peritonitis (SBP),[16] adversely affect the survival of patients with cirrhosis. It is also recognized that in patients who experience these complications, the degree of hepatic dysfunction is the most important prognostic factor. Therefore, the question becomes whether considering these individual complications of portal hypertension provides further prognostic information in addition to the MELD score.

Not all the data sets that we have used for the validation of the MELD scale contained specific information with regard to the portal hypertensive complications. History of variceal bleeding, ascites, and encephalopathy were available in the "hospitalized" data set. The impact of SBP was analyzed by reconstructing the diagnosis of SBP using peritoneal fluid analysis data from the "historical" data set. The conventional criteria of absolute neutrophil count greater than 250 /mm^3 were used to identify patients with SBP.[17] History of variceal bleeding was also extracted from the medical index for the "historical" data set. Finally, the ambulatory noncholestatic data set contained information about the presence of ascites and encephalopathy. Table 4 shows the effect of adding the individual portal hypertensive complications to the MELD score on the overall c-statistic. There is minimal improvement in the prediction of the 3-month mortality, with the increase in c-statistics ranging from less than 0.01 for spontaneous bacterial peritonitis and variceal bleeding to 0.01 for hepatic encephalopathy and 0.01 to 0.03 for ascites.

HOW IMPORTANT IS THE LIVER DISEASE ETIOLOGY IN THE MELD MODEL?

In the current form of the MELD scale, patients with cholestatic and alcohol-induced liver disease are given lower risk scores, when other parameters of liver disease are comparable. Our original interpretation of the model was that the bilirubin level in cholestatic patients has a connotation different from that in patients with parenchymal liver disease. This is also accounted for in the CTP score in which modified criteria for bilirubin are applied to patients with cholestatic disease. With regard to patients with alcohol-induced liver disease, we had surmised that patients undergoing the TIPS procedure may be more likely to abstain from alcohol, leading to subsequent improvement in liver function and survival.

In the application of the model, however, particularly as a disease severity index for organ allocation, several issues may be relevant with regard to the inclusion of the etiology in the model. First, for patients with alcohol-induced liver disease, the vast majority of liver transplantation programs require a certain length of time, most commonly 6 months, of abstinence before registration onto the waiting list. Thus, unlike the selected patients with alcohol-induced liver disease included in the development and validation of the model, liver transplantation candidates may have little room for improvement from subsequent abstinence and their survival may be no different from other parenchymal diseases.

Second, patients with cholestatic disease may experience different types of complications than parenchymal liver disease patients, such as pruritus, osteoporosis, fat malabsorption, and fatigue.[18] In addition, patients with primary sclerosing cholangitis may have further complications such as ascending cholangitis, cholangiocarcinoma, or problems re-

TABLE 4. Consideration of Portal Hypertensive Complications in the MELD Score

	Hospitalized N = 282			Ambulatory Noncholestatic N = 491			Historical N = 1179		
Complication	N*	MELD	MELD + Complication	N*	MELD	MELD + Complication	N*	MELD	MELD + Complication
SBP	—	—	—	—	—	—	18	0.78	0.78
Variceal bleed	30	0.87	0.88	—	—	—	107	0.78	0.78
Ascites	116†	0.87	0.88	94	0.80	0.83	—	—	—
Encephalopathy	52†	0.87	0.88	21	0.80	0.81	—	—	—

* Number of patients with the given complication.

† Due to missing values, 271 patients were included in the analysis for ascites or encephalopathy among hospitalized patients.

TABLE 5. Impact of Exclusion of the Etiology of Liver Disease

	Hospitalized N = 282	Ambulatory Noncholestatic N = 491	Ambulatory PBC N = 326*	Historical N = 1,179
With etiology in the model	0.87 (0.82-0.92)	0.80 (0.69-0.90)	0.87 (0.71-1.00)	0.78 (0.74-0.81)
Without etiology in the model	0.86 (0.81-0.92)	0.82 (0.73-0.91)	0.87 (0.71-1.00)	0.78 (0.74-0.81)

NOTE. Values reported are the concordance statistic (95% CI).

* Because all patients in these data had PBC, there is no effect from excluding the etiology.

lated to inflammatory bowel disease.[19,20] As these factors are not incorporated in the current MELD scale, one may consider it inequitable to assign lower scores for patients with cholestatic liver disease. Furthermore, patients with primary sclerosing cholangitis are often put on ursodeoxycholic acid, which may decrease the bilirubin level but does not provide survival benefit.[21]

Third, difficulties in model application may arise when a patient has a combination of 2 or more causes for the liver disease, such as alcohol abuse and hepatitis C or PBC and autoimmune hepatitis. Among patients with hepatitis C, which is currently the leading indication for liver transplantation in the United States, many have coexistent alcohol-induced liver disease, making it difficult, if not impossible to determine which is the primary cause of liver disease.

These factors led us to consider examining the validity of a modified model that excludes the liver disease diagnosis. Table 5 shows in the 4 validation sets, the exclusion of the diagnosis has minimal influence on the c-statistics of the model.

IS THE MELD SCALE USEFUL IN ASSESSING VERY SHORT-TERM SURVIVAL?

If the MELD scale is going to be used as the disease severity index for the organ allocation system in which liver disease severity will be the most weighed criterion, it is important for the model to be valid in patients with the most advanced liver disease and short-term survival likelihood. As shown in Table 2, the c-statistics when the MELD scale was used for prediction of 1-week mortality strongly suggest validity of its use to determine short-term outcome (c-statistic range, 0.80-0.95).

DOES BODY SIZE AFFECT THE VALIDITY OF THE MELD SCALE?

To the extent that the serum creatinine is determined in part by the total body muscle mass as well as the renal function, the MELD score may underestimate the severity of liver disease in patients with small body size.[22] In examining the impact of body size on the MELD score, patient body weight and height were only available in the "hospitalized" patient data set. Table 6 summarizes the serum creatinine concentration by age, sex, and body mass index. In both genders, there was no significant correlation between body mass index and MELD score ($r = -0.05$ for men [$P = .6$], $r = 0.19$ for women [$P = .1$]).

TABLE 6. Relationship Between Body Size and Creatinine by Age and Sex

	Men		Women	
	Age < 50	Age ≥ 50	Age < 50	Age ≥ 50
Creatinine				
BMI < 25	0.8 [0.6-1.3]	1.3 [1.0-2.1]	0.9 [0.8-1.1]	0.9 [0.7-1.0]
BMI ≥ 25	1.1 [0.8-1.2]	1.1 [0.9-1.3]	0.8 [0.7-0.9]	0.9 [0.8-1.1]

NOTE. Data are presented as median [25th percentile-75th percentile].

Abbreviation: BMI, body mass index.

We tested whether adding body mass index to the MELD score improves the model. In a logistic model predicting death within 3 months in hospitalized patients, once the MELD score is considered, the addition of the body mass index did not improve the model ($P = .14$).

DISCUSSION

This study corroborates that the model previously developed to predict survival following the TIPS procedure may be used as a reliable tool to assess survival in patients with chronic liver disease not undergoing TIPS. The model, referred to as Model for End-Stage Liver Disease (MELD), is based on serum creatinine, serum bilirubin, and INR for prothrombin time with or without liver disease etiology and is generalizable to patient populations of diverse etiologies and wide ranges of severity. Several aspects of the parameters used in our model, in contrast to the Child-Turcotte-Pugh classification, need to be addressed further.

The CTP classification is the model most widely used to determine prognosis in patients with liver disease.[23,24] Although the original purpose of this system was to assess the operative risk in patients undergoing surgical portosystemic shunt, the classification has been used to stratify patients on the waiting list for liver transplantation.[4] The CTP classification is based on serum albumin, serum bilirubin, prothrombin time, ascites, and encephalopathy. Although the classification system has not been formally evaluated for its statistical accuracy, it has been shown to be useful in the assessment of prognosis in patients with cirrhosis.[25] However, when used as a disease severity index to determine priority in organ allocation, the CTP system has a number of limitations. These include (1) limited discriminatory ability, (2) subjective interpretation of parameters, and (3) variability in the measurement of the laboratory parameters.

First, the CTP classification has a limited discriminatory ability. Even when the CTP score, as opposed to CTP class, is used, there are only 8 levels of difference between the least sick transplant candidates (CTP score = 7) and the most advanced (CTP score = 15). This creates many ties (*i.e.,* transplant candidates with the same score) and necessitates an emphasis on waiting time as a tie breaker. In addition, it is not able to evaluate patients with markedly abnormal laboratory parameters (so-called ceiling effect). For example, a patient with a bilirubin of 3 mg/dL and a patient with a bilirubin of 30 mg/dL are given the same score and thus are determined to have the same severity of liver disease. Likewise, patients with an albumin of 2.8 g/dL or 1.5 g/dL are again determined to have the same severity of liver disease. Finally, all the parameters in the system are given the same weight. For instance, a

patient with an albumin of 3.5 g/dL and a patient with spontaneous hepatic encephalopathy are both given a score of 2, even though the prognostic implication of the 2 abnormalities is likely to be dissimilar. In contrast, the MELD scale is a continuous system with no ceiling or floor in the score and the coefficients in the scale are derived statistically so that appropriate weights are given to variables according to their relative importance.

Second, the CTP score has parameters that require subjective assessment, namely ascites and encephalopathy. When originally described, determination of ascites was based on physical examination. Currently, however, ultrasound examination is frequently used to detect the presence of ascites. Detection of encephalopathy is dependent on the examiner, and there is no accepted definition of refractory encephalopathy. Thus, the scores given for the degree of ascites and encephalopathy are difficult to standardize. Furthermore, these parameters may change with treatment. In contrast, the MELD scale is based on objective parameters. The only parameter in the model that required interpretation was the etiology of liver disease. As described previously, when the etiology of the liver disease was excluded, the accuracy of the model did not suffer appreciably. Thus, the modified model using only the variables of bilirubin, prothrombin time, and creatinine completely exclude interpretative subjectivity while preserving the accuracy of the model.

Third, even the more objective elements in the CTP system, namely albumin and prothrombin time, may vary from one laboratory to another. For instance, in an informal survey we conducted, the normal range of albumin across the United States varies from 2.9-4.5 g/dL in some laboratories to 3.8-5.1 g/dL at other laboratories. Because of this variability, a patient with a 5% decrease in albumin synthesis will be given 1 point in the CTP system in a laboratory where the lower limit of normal for albumin is 3.8 g/dL (5% decrease = albumin 3.6 g/dL), but 3 points if the albumin is measured in a laboratory where the lower limit of normal is 2.9 g/dL (5% decrease = albumin 2.8 g/dL). The prothrombin time is currently standardized and reported in most places worldwide as the INR for prothrombin time. The prothrombin time in seconds, which is used in the CTP classification, however, depends on the sensitivity of the thromboplastin reagent used (International Sensitivity Index, ISI) and, based on this sensitivity, the prothrombin time in seconds can vary greatly from laboratory to laboratory. For example, for a prothrombin time INR of 2.0, which is standardized across the country, a patient will have a prolongation of the prothrombin time of 10 seconds if the ISI for thromboplastin is 1, 4 seconds if the ISI for thromboplastin is 2, and 2.6 seconds if the ISI is 3. Thus, for an identical INR for prothrombin time, a patient can get a score of 1, 2, or 3 in the CTP classification depending on which laboratory the prothrombin time is measured. Although there is some suggestion that prothrombin activity may more accurately reflect liver function than the INR does,[26] prothrombin activity is not routinely measured in the United States.

All in all, in the comparison between the MELD scale and the CTP score, the MELD scale is at least as good as the CTP score in predicting short-term mortality, while it is devoid of many limitations of the CTP score. In addition, it incorporates serum creatinine level, a measure of renal function in patients with liver disease and a well-recognized predictor of survival in patients with liver disease,[13,27] and outcome post liver transplantation.[28-31]

Of interest, the addition of spontaneous bacterial peritonitis, encephalopathy, variceal bleeding, or ascites to the MELD scale did not improve the accuracy of the model. This indicates that patients with severe liver disease are the ones likely to get complications of liver disease and, thus, the complications are not independent predictors of survival. We have previously reported that hospital mortality in patients with SBP is determined by serum bilirubin, creatinine, and prothrombin time,[32] identical factors included in the MELD scale. Finally, the comparison between the MELD scale and the Mayo PBC model further reassures the validity of the MELD scale. Given that the data set in which the two models were compared was the one with which the PBC model was developed and that the validity of the latter model has been repeatedly shown, the small difference in the accuracy of prediction (0.91 for the PBC model versus 0.87 for the MELD scale) bolsters our confidence in the MELD scale.

There are limitations to the MELD scale. First, although we tried to maximize the range of disease severity that the scale may be applied to, most of the data available to us were mostly based on patients with advanced liver disease. Overall, the 3-month mortality among the data sets used in this report range between 2% and 21%. This is in contrast to recent reports on the natural history of compensated cirrhosis from hepatitis C, in which the 5-year mortality is described between 9% to 24%.[33-36] Although this may warrant further validation of the MELD scale in patients with early stage cirrhosis, we believe this report provides sufficient evidence for its use in liver transplant candidates with end-stage liver disease. Second, although we presented data to support that the influence of the age, sex, and body mass on the MELD score is unlikely to be clinically significant, it is possible that a more direct measurement of renal function, such as iothalamate clearance, may improve the accuracy of the model. Further investigation on the optimal measurement of renal function in this context may be needed. Another caveat regarding the creatinine level is that excessive use of diuretics may result in dehydration and worsening in the renal function. We recommend that for accurate application of the model, the value of creatinine should be used to determine survival when the patient is hemodynamically stable and adequately hydrated.

In summary, MELD is a reliable measure of short-term mortality risk in patients with end-stage liver disease of diverse etiology and severity. As a comprehensive indicator of physiologic reserve of patients with decompensated cirrhosis, the validity of the MELD scale is shown in patients with advanced liver disease independent of complications of portal hypertension. In addition to the validity of the score in diverse clinical circumstances, many of its characteristics such as the use of easily available, objective parameters as well as its advantage over the CTP score such as continuous, interval scale, lack of ceiling or floor effect, and stability of laboratory measures make it compatible with the criteria proposed by the IOM for application in the allocation decisions for liver transplantation.

Acknowledgment: The authors thank Drs. Todd Howard and Richard Freeman and members of the Committee for their thoughtful input. The authors also thank Drs. Luigi Pagliaro and Linda Pasta for providing data on Italian patients.

REFERENCES

1. Anonymous. Organ Procurement and Transplantation Network—HRSA. Final rule with comment period. Federal Register 1998;63:16296-16338.
2. Institute of Medicine. Analysis of waiting times. In: Committee on Organ Procurement and Transplantation Policy, ed. Organ Procurement and Transplantation: Assessing Current Policies and the Potential Impact of the DHHS Final Rule. Washington, DC: National Academy Press, 1999: 57-78.
3. Freeman R, Edwards E. Liver transplant waiting time does not correlate with waiting list mortality: implications for liver allocation policy. Liver Transplantation 2000;6:543-552.
4. Lucey MR, Brown KA, Everson GT, Fung JJ, Gish R, Keeffe EB, Kneteman NM, et al. Minimal criteria for placement of adults on the liver transplant waiting list: a report of a national conference organized by the American Society of Transplant Physicians and the American Association for the Study of Liver Diseases. Liver Transplant Surg 1997;3:628-637.
5. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000;19:453-473.
6. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Int Med 1999;130:515-524.
7. 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.
8. Langer B, Taylor BR, Mackenzie DR, Gilas T, Stone RM, Blendis L. Further report of a prospective randomized trial comparing distal splenorenal shunt with end-to-side portacaval shunt. An analysis of encephalopathy, survival, and quality of life. Gastroenterology 1985;88:424-429.
9. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36.
10. Van der Besselarr A. Precision and accuracy of the international normalized ratio for prothrombin time in oral anticoagulation control. Hemostasis 1996;26:248-265.
11. Dickson ER, Grambsch PM, Fleming TR, Fisher LD, Langworthy A. Prognosis in primary biliary cirrhosis: model for decision making. HEPATOLOGY 1989;10:1-7.
12. Grambsch PM, Dickson ER, Kaplan M, LeSage G, Fleming TR, Langworthy AL. Extramural cross-validation of the Mayo primary biliary cirrhosis survival model establishes its generalizability. HEPATOLOGY 1989;10:846-850.
13. Arroyo V, Gines P, Gerbes A, Dudley F, Gentilini P, Laffi G, Reynolds T. Definition and diagnostic criteria of refractory ascites and hepatorenal syndrome in cirrhosis. HEPATOLOGY 1996;23:164-176.
14. Mullen K, Dasarathy S. Hepatic encephalopathy. In: Schiff ER, Maddrey WC, ed. Shiff's Disease of the Liver. Philadelphia: Lippincott-Raven, 1999:545-581.
15. Graham D, Smith J. The course of patients after variceal hemorrhage. Gastroenterology 1981;80:800-806.
16. Altman C, Grange JD, Amiot X, Pelletier G, Lacaine F, Bodin F, Etienne JP. Survival after a first episode of spontaneous bacterial peritonitis. Prognosis of potential candidates for orthotopic liver transplantation. J Gastroenterol Hepatol 1995;10:47-50.
17. Rimola A, Garcia-Tsao G, Navasa M, Piddock L, Planas R, Bernard B, Inadomi J. Diagnosis, treatment and prophylaxis of spontaneous bacterial peritonitis: a consensus document. HEPATOLOGY 2000;32:142-153.
18. Kaplan MM. Primary biliary cirrhosis. N Engl J Med 1996;335:1570-1580.
19. Angulo-Hernandez P, Lindor K. Primary sclerosing cholangitis. HEPATOLOGY 1999;30:325-332.
20. deGroen P, Gores GJ, LaRusso N, Gunderson L, Nagorney D. Medical progress-biliary tract cancers. New England Journal of Medicine 1999; 28:1368-1378.
21. Lindor K. Ursodiol for primary sclerosing cholangitis. N Engl J Med 1997;336:691-695.
22. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron 1976;16:31-41.
23. 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.
24. Child C, Turcotte J. Surgery and portal hypertension. In: Child CG, ed. The Liver and Portal Hypertension. Philadelphia: Saunders, 1964:50.
25. Christensen E, Schlichting P, Fauerholdt L, Gluud C, Anderson P, Juhl E, Poulson H, et al. Prognostic value of Child-Turcotte criteria in medically treated cirrhosis. HEPATOLOGY 1984;4:430-435.
26. Robert A, Chazoulleres O. Prothrombin time in liver failure: time, ratio, activity percentage, or international normalized ratio? HEPATOLOGY 1996; 24:1392-1394.
27. Gines A, Escorsell A, Gines P, Salo P, Jimenez W, Inglada L. Incidence, predictive factors, and prognosis of the hepatorenal syndrome in cirrhosis with ascites. Gastroenterology 1993;105:229-236.
28. Gonzalez E, Rimola A, Navasa M, Adnreu H, Grande L, Garcia-Valdecasas J, Cirera I, et al. Liver transplantation in patients with non-biliary cirrhosis: prognostic value of preoperative factors. J Hepatol 1998;28: 320-328.
29. Gonwa TA, Klintmalm GB, Levy M, Jennings LS, Goldstein RM, Husberg BS. Impact of pretransplant renal function on survival after liver transplantation. Transplantation 1995;59:361-365.
30. Cuervas-Mons V, Millan I, Gavaler JS, Starzl TE, Van Thiel DH. Prognostic value of preoperatively obtained clinical and laboratory data in predicting survival following orthotopic liver transplantation. HEPATOLOGY 1986;6:922-927.
31. Ricci P, Therneau TM, Malinchoc M, Benson JT, Petz JL, Klintmalm GB, Crippen JS, et al. A prognostic model for the outcome of liver transplantation in patients with cholestatic liver disease. HEPATOLOGY 1997;25: 672-677.
32. Angulo-Hernandez P, Pasha T, Lindor K, Kamath P. A model for prediction of in hospital mortality in cirrhotic patients with spontaneous bacterial peritonitis [Abstract]. HEPATOLOGY 1997;26:355A.
33. Fattovich G, Giustina G, Degos F, Tremolada F, Diodati G, Almasio P, Nevens F, et al. Morbidity and mortality in compensated cirrhosis type C: a retrospective follow-up study of 384 patients. Gastroenterology 1997; 112:463-472.
34. Serfaty L, Aumaitre H, Chazouilleres O, Farmachidi JP, Degott C, Guettier C, Trinchet JC, et al. Determinants of outcome of compensated hepatitis C virus-related cirrhosis. HEPATOLOGY 1998;27:1435-1440.
35. Hu KQ, Tong MJ. The long-term outcomes of patients with compensated hepatitis C virus-related cirrhosis and history of parenteral exposure in the United States. HEPATOLOGY 1999;29:1311-1316.
36. Degos F, Christidis C, Ganne-Carrie N, Bonnand AM, Rosmorduc O, Poupon RE, Poupon R, et al. Hepatitis C virus related cirrhosis: time to occurrence of hepatocellular carcinoma and death. Gut 2000;47:131-136.