Prognostic Value of Metabolic Liver Function Tests: a Study on 711 Cirrhotic Patients

Fanny Lebossé1, Olivier Guillaud2, Julien Forestier2, Marie Ecochard2, Olivier Boillot2, Sabine Roman1, François Mion1, Jérôme Dumortier2

INTRODUCTION

Chronic liver diseases are a frequent cause of death in France, estimated between 2,000 and 3,300 cases per million inhabitants per year [1]. These diseases are related to different etiologies including mainly alcohol, hepatitis C virus, and non-alcoholic fatty liver disease. Patients can develop cirrhosis, which is characterized by portal hypertension, hepatocellular dysfunction, liver failure and a risk for hepatocellular carcinoma. Predicting evolution of chronic liver disease to improve therapeutic decision is a challenge, especially for patients who could obtain a liver transplant (LT). This option represents the ultimate therapy but it is resource-consuming. Therefore, prognostic factors to determine survival without LT are urgently required. Various scores combining clinical, biochemical and histological parameters have been proposed and are expected to optimize timing for LT and prioritize allocation of liver grafts. Assessment of liver function can be based on static tests such as Child-Pugh (CP) score [2] or Model for End Stage Liver Disease (MELD) score [3]. The CP score is simply obtained from clinical and current laboratory data. The MELD score was initially created to predict survival in patients with complications of portal hypertension undergoing elective placement of transjugular intrahepatic portosystemic shunt (TIPS) [3]. Nevertheless, despite the many advantages of the MELD score, there are approximately 15%-20% of patients whose survival cannot be accurately predicted by this score [4]. On the other hand, dynamic assessments such as clearance of indocyanine green (IGC), formation of metabolites such as

ABSTRACT

Background & Aims: The prognosis of cirrhotic patients is usually assessed by Child-Pugh and MELD scores. Metabolic liver function tests such as aminopyrine breath test (ABT) and indocyanine green clearance (IGC) have been shown to reveal hepatocellular dysfunction. The aim of this retrospective study was to compare the prognostic value of the MELD score, Child-Pugh score, ABT and IGC in a large cohort of cirrhotic patients.

Methods: Between January 1996 and June 2008, 711 cirrhotic patients were included and the primary endpoint was survival without LT. The ROC curves with c-statistics, correlation coefficient and survival were calculated.

Results: Metabolic function tests and scores were strongly correlated. At the time of evaluation, 111 patients had died and 520 had received a transplant. Prognostic ability (estimated by the AUROC curve) to predict survival without LT at 6 months was 0.662, 0.691, 0.738 and 0.715 for ABT, IGC, Child-Pugh score and MELD score, respectively. Similarly, at 1 year, AUROC was 0.738 for Child-Pugh score, 0.716 for MELD score, 0.693 for IGC clearance and 0.651 for ABT.

Conclusions: Our results strongly confirm that IGC and ABT have a high prognostic value in cirrhotic patients, similar to Child-Pugh and MELD scores. They could be developed to routinely evaluate the prognosis of patients in addition to clinical and biochemical data.

Key words: cirrhosis – prognosis – liver transplantation – MELD score – Child-Pugh score – aminopyrine breath test – indocyanine green clearance.

Abbreviations: ABT: Aminopyrine Breath Test; AUROC: area under the receiver operating characteristics curve; CP score: Child-Pugh score; DRI: deceased donor risk index; IGC: Indocyanine Green Clearance; LT: Liver transplantation; MELD: Model for End stage Liver Disease.
lidocaine to monoethylglycinxylidide (MEGX) or aminopyrine has been shown to estimate hepatocellular dysfunction and to predict mortality [5].

The aim of the present study was to compare the prognostic ability of two static tests, MELD and CP scores, and two dynamic tests, IGC and aminopyrine breath test (ABT).

PATIENTS AND METHODS

Study population
The present retrospective study is based on 711 patients with chronic liver disease, who were evaluated consecutively in our center (Hôpital Edouard Herriot, Lyon, France) between January 1996 and June 2008. All these patients had liver cirrhosis. The IGC and/or ABT were realized and the MELD and CP scores were calculated at the same time. Patients’ demographic characteristics (sex, age) and the cause of liver disease were available for all 711 patients.

The CP score was referenced for 707 patients (99.4%) and the MELD score for 708 patients (99.6%). Concerning liver function tests, IGC was available for 297 patients (41.8%) and ABT for 639 patients (89.9%).

The proportion of patients with HCC was 60.1% in the CP class A group, 29.8% in the CP class B group, and 18.5% in the CP class C group. From the 711 patients’ entire cohort, 565 were listed for LT. Graft allocation policy was not based on the MELD score in France during the study period, and was mostly based on the waiting time on the LT list and on the decision of the transplantation team.

CP and MELD scores
The CP score was calculated as usually described [2], with laboratory data including bilirubin level, albumin concentration and prothrombin rate and two clinical data: presence of ascites and hepatic encephalopathy. The CP score was defined by values between 5 and 15. As usually, we divided patients in three classes: A, B and C. Patients with score of 5 or 6 were classified in class A, patients with score of 7, 8 or 9 in class B and patients with score above 9 in class C.

The MELD score was calculated with laboratory data: International Normalized Ratio (INR) and bilirubin level [6]. Patients were divided into two groups: MELD score below and above 20, which has been previously used to determine the need for LT [7].

Metabolic liver function tests
After a 12-hour fast, indocyanine green at a dosage of 0.5 mg per kg body-weight was intravenously administered (antecubital vein) [8]. We collected blood from the contralateral antecubital vein at pre-dose and at 5, 8, 11, 15 and 20 min after injection. Then, the plasma samples were read against the plasma blank at 805 nm by a spectrophotometer. We established concentrations using a standard calibration curve. The results were expressed as the percentage of indocyanine green retained 15 minutes after injection.

The ABT was performed using the oral method [9]. First, we collected a baseline breath sample. Then, fasting patients were given an oral dose of 140 mg of 13C-aminopyrine dissolved in water. One hour later, a breath sample was collected and the 13C/12C ratio in breath CO2 was measured in both samples with GC-IRMS. Results were expressed as percent of ingested dose of 13C excreted as breath 13CO2 per hour.

We stratified patients into groups according to their IGC results (<30%, 30 to 60%, >60%, tertile stratification was chosen) and their ABT results (>2%, 1-2%, <1%).

Patients’ outcome
Patients were followed from the date of their metabolic liver function tests and for at least one year. The primary endpoint was patients’ survival without LT.

Statistics
Quantitative variables were described using mean, range and standard deviation (SD). Qualitative values were tabulated and percentages were calculated.

p-Spearman correlation between scores and tests were calculated and their relationship were grafted on log scales.

The concordance statistic between the prediction of mortality by the scores and tests was calculated for the two periods of follow-up. The concordance statistic c is identical to the area under the receiver operating characteristics curve (AUROC). The c-statistic varies between 0.5 and 1.0 for sensible models; the higher the better. Only the results of c-statistics above 0.5 were considered as statistically significant.

Kaplan-Meier was used to graph survival according to the levels of scores and tests. A Cox regression model was used to test the significance of these scores and tests as predictors of survival.

Statistical tests were considered as significant if p<0.05.

RESULTS

Study population
The study population comprised 505 men (71%) and 206 women (29%). The median age was 53 years (20 to 77). The median follow-up until death, LT or the last visit was 296 days. At 6 months after inclusion, 277 (39%) patients survived without LT, and 434 (61%) patients did not survive without LT (71 deaths and 363 LT). At 1 year after inclusion, the rates were 21.7% (154) and 76.7% (545: 89 patients died without LT and 456 were transplanted), respectively (Table I).

The ABT was available for 639 patients (89.9%). The patients were classified into three groups: 112 patients (17.5%) in the ABT <1% group, 151 (23.6%) in the ABT > or = 1% and <2% and 376 (58.9%) in the group ABT > or = 2%.

The IGC was available for 297 patients (41.8%). The patients were classified into three groups according to the indocyanine green retention rate at 15 minutes: IGC< 30% (n=141 or 40.7%), IGC> or = 30% and <60% (n=102 or 34.4%) and IGC > 60% (n=74 or 24.9%).

The CP score was available for 707 patients (99.4%): 228 patients (32.2%) were classified as Child A, 252 patients (35.7%) as Child B and 227 patients (32.1%) as Child C.

The study population comprised 505 men (71%) and 206 women (29%). The median age was 53 years (20 to 77). The median follow-up until death, LT or the last visit was 296 days. At 6 months after inclusion, 277 (39%) patients survived without LT, and 434 (61%) patients did not survive without LT (71 deaths and 363 LT). At 1 year after inclusion, the rates were 21.7% (154) and 76.7% (545: 89 patients died without LT and 456 were transplanted), respectively (Table I).

The ABT was available for 639 patients (89.9%). The patients were classified into three groups: 112 patients (17.5%) in the ABT <1% group, 151 (23.6%) in the ABT > or = 1% and <2% and 376 (58.9%) in the group ABT > or = 2%.

The IGC was available for 297 patients (41.8%). The patients were classified into three groups according to the indocyanine green retention rate at 15 minutes: IGC< 30% (n=141 or 40.7%), IGC> or = 30% and <60% (n=102 or 34.4%) and IGC > 60% (n=74 or 24.9%).

The CP score was available for 707 patients (99.4%): 228 patients (32.2%) were classified as Child A, 252 patients (35.7%) as Child B and 227 patients (32.1%) as Child C.

Table II summarizes the patients’ outcome according to the CP class.

The MELD score was calculated for 708 patients (99.6%): 540 patients (76.3%) presented a score equal or lower than 20 and 168 (23.7%) a score higher than 20.
Correlation between scores and metabolic liver function tests
The CP and MELD scores were strongly correlated ($p=0.801$, $p<0.0001$). The IGC and ABT were also strongly correlated one with another ($p=0.731$, $p<0.0001$). The CP score was better correlated to IGC and ABT ($p=0.781$, $p<0.0001$ and $p=0.691$, $p<0.0001$, respectively) than to the MELD score ($p=0.683$, $p<0.0001$ and $p=0.585$, $p<0.0001$, respectively).

Prognostic value of CP and MELD scores, and of the metabolic liver function tests
The prognostic ability (expressed as c-statistics) of the CP and MELD scores, and of the metabolic liver function tests for the two periods of follow-up (6 months and 1 year) are shown in Fig. 1. For the prediction of survival without LT at 6 months, the AUROC was 0.738 for CP score, 0.715 for MELD score,
The ROC curves of scores and tests for the prediction of patients’ survival without LT at 6 months (A) and 1 year (B).

0.691 for IGC and 0.662 for ABT. Similarly, for the prediction of survival without LT at 1 year, the AUROC was 0.738 for the CP score, 0.716 for the MELD score, 0.693 for the IGC and 0.651 for the ABT.

Actuarial survival (expressed as Kaplan-Meier curves) without LT according to CP class, MELD score, IGC and ABT are shown in Fig. 2. All scores and tests were of significant prognostic value (log-rank, p< 0.05). In our population, one-year probability of survival without LT decreased from 42% to 25% and to 7%, if patients were classified as CP class A, B or C, respectively. Similarly, one-year probability of survival without LT decreased from 28% to 6% if MELD score was ≤20 or >20, respectively. One-year probability of survival without LT decreased from 43% to 29% and to 12% if IGC rate was <30%, > or = 30% and <60% or > 60%, respectively. One-year probability of survival without LT decreased from 30% to 18% and to 11%, if ABT was > or = 2%, > or = 1% and <2% or <1%, respectively.

**DISCUSSION**

It is a challenging issue to elaborate reliable tools for predicting outcome of patients with severe chronic disease, such as cirrhosis. Thus, the main objective of the prognostic scores in cirrhotic patients is to estimate the probability of death (or death without LT). In this field, the oldest scores (especially the
CP score) included a limited number of variables, which were put together empirically, whereas more recent scores (including the MELD score) are based on a subset of variables that were shown to be significantly and independently correlated to survival by multivariate analysis. The CP score has been the gold standard for more than 30 years and the MELD score is its most serious alternative for overcoming its limitations. In addition, the metabolic liver function tests have also been shown to be efficient to assess the severity of liver failure and the prognosis for patients with cirrhosis [5, 10-12]. The results of the present study confirm that both IGC and ABT taken on their own are strong predictors of 6 and 12-month survival without LT, similarly to MELD and CP scores.

The MELD and CP scores are undoubtedly useful but not totally efficient. The major advantages of the MELD score are the following: it is based on variables that have been selected by statistical analysis (and therefore each variable is weighted according to its proper influence on prognosis) rather than by clinical judgement, these variables are objective and unlikely to be influenced by external factors (such as albumin perfusion in CP score), and the score is continuous, which theoretically leads to classify patients more precisely. Nevertheless, the MELD score also has some limitations. The variables included in MELD (serum creatinine, prothrombin time, and bilirubin) are, compared to those included in the CP score, all objective, but they may be subjected to some variability depending on how they are measured (i.e. colorimetric method overestimates the serum creatinine if serum bilirubin >25mg/dl), on concomitant disease (i.e. dehydrratation, hemolysis, ...) or therapeutics (i.e. diuretics, ...). In addition, the MELD score has not been validated in some specific clinical situations and complications of liver cirrhosis (such as pulmonary complications, predominant portal hypertension, ...). As a result, it has been suggested that survival of 15 to 20% of cirrhotic patients cannot be accurately predicted [4]. Interestingly, in our study, the MELD score had a good prognostic ability but not greater than that of CP (AUROC for the prediction of survival without LT at 6 months was 0.738 for the CP score vs 0.715 for the MELD score). In a review on the performance of the MELD score in the setting of LT, Cholongitas et al. [13] found that only 4/11 studies showed MELD to be superior to CP in predicting short-term mortality. Similarly, a review of the studies comparing the accuracy of MELD vs CP in non-transplant settings suggested that MELD score does not perform better than CP score. In TIPS studies (including 1,360 patients) only one of five showed MELD to be superior to CP to predict 3-month mortality, but not 12-month mortality. In the studies in cirrhotic patients (including 2,569 patients) none of the four showed significant differences between the two scores for either short- or long-term prognosis [14]. More recently, a large systematic review and meta-analysis aimed at comparing the discriminative ability of CP vs. MELD score to assess the prognosis of cirrhotic patients [15]. Out of the 1,095 papers initially identified, 119 were eligible for the systematic review. They included 269 comparisons, of which 44 favored MELD score, 16 favored CP score, 99 did not find any significant difference between them, and 110 did not report the statistical significance. Forty-two papers were included in the meta-analysis. Interestingly, although CP and MELD scores had similar prognostic values in most of the cases, their benefits might be heterogeneous in some specific conditions. In patients with acute-on-chronic liver failure, the CP score had a higher sensitivity and a lower specificity than the MELD score. In patients admitted to the intensive care unit (ICU), the MELD score had a smaller negative likelihood ratio and a higher sensitivity than the CP score. In patients undergoing surgery, the CP score had a higher specificity than the MELD score. Finally, refinements in MELD score have been designed, including MELD-Na (containing Na), integrated MELD (iMELD) score (containing age, MELD and Na), and MESO index (MELD to Na ratio), to further improve survival prediction in cirrhotic patients.

Dynamic metabolic liver function tests are able to assess liver function [5]. Their interest in evaluating the prognosis for patients with chronic liver disease, especially for patients waiting for LT has not been extensively studied and many data are very old, from studies performed before the MELD era. Recently, we reported that ABT had been shown to improve the prognostic value of MELD score in patients waiting LT [16]. In that study, ABT provided additional information to that already given by the MELD score. In patients with a MELD score below 20 and even more in those with a MELD score above 20, ABT allowed the determination of subgroups with a statistically different survival rate. It can discriminate patients with a MELD quite low but with a poor prognosis: for example, patients with a MELD score below 20 and ABT <1% had a probability of a 1-year survival of 70% versus 87% if ABT >2%. Vice-versa, it was able to discriminate patients with a high MELD score but who still have a rather long survival (i.e., MELD>20 and ABT >2%, rate of 3-month survival 78% vs. 51% if ABT<1%). In the present study, in comparison to these two static tests (CP and MELD), IGC and ABT revealed a quite similar prognostic value based on the AUROCs.

Similarly to MELD and CP scores, the IGC and ABT have some advantages and limitations. ABT is a very simple and safe liver function test which needs neither blood injection nor blood sampling. It should be noticed that it is metabolized by cytochrome P450 enzymes and therefore can be induced or inhibited by many drugs, including alcohol and tobacco. It must also be taken into account that elevated serum bilirubin prevents an accurate measurement of IGC, which requires intravenous injection and sampling.

In a study on 604 patients with cirrhosis, the galactose-elimination-capacity, sorbitol clearance, lidocaine metabolism and indocyanin green half-life were compared to MELD and MELD-Na scores [17]. The ICG half-life was the most accurate in predicting survival. Interestingly, its incorporation into the MELD (leading to a new score “MELD-ICG”) rendered the more accurate in predicting survival in intermediate to advanced cirrhosis than the MELD and MELD-Na. Therefore, new scores including static and dynamic liver function tests could be the best valuable prognostic scores, and this undoubtedly needs a further and larger evaluation.

Finally, our present study has some limitations that must be taken into account. First of all, the metabolic liver function tests were not available for the entire study population: ABT was available for 639 patients (89.9%) and IGC was available for 297 patients (41.8%). Patients’ characteristics evaluated
to predict survival were registered only at a single time point and many chronic liver diseases run a course with phases of improvement and deterioration. It would probably be interesting, as suggested for the MELD score, to study the evolution of the tests at regular intervals of time. The choice of 20 as a cut-off for MELD score was made because it has been initially shown that significant transplant survival benefit is observed at a MELD score of 20 and above [7]. These US data were confirmed thereafter by data from the Scientific Registry of Transplant Recipients on 28,165 adult liver transplant candidates listed between 2001 and 2005. Survival benefit was estimated according to cross-classifications of the candidate MELD score and deceased donor risk index (DRI) using sequential stratification. All recipients with MELD ≥ 20 had a significant survival benefit from transplantation, regardless of the DRI [18]. Moreover, Lucey et al. evaluated if survival benefit can be related to the initial liver disease, and compared alcoholic liver disease and HCV-related cirrhosis [19], using again the data from the Scientific Registry of Transplant Recipients on 38,899 adult patients placed on the transplant waiting list between September 2001 and December 2006. The survival benefit of transplantation was significantly decreased among HCV+ compared with HCV- recipients with MELD scores 9-29, but was significantly increased at MELD ≥ 30. Alcoholic liver disease did not influence the survival benefit of transplantation at any MELD score. In addition to the US data, Ravaioli et al. reported in a European center (vs. US) that LT should be reserved for cirrhotic patients with MELD score ≥20 independently of other recipient and donor matches or for cases with lower MELD score but with HCC [20].

In our study, a significant proportion of patients (25.9%) presented HCC. We chose to include these patients in the analysis despite that the issue in this population is the prognosis of both cirrhosis and HCC. We believe that this did not affect our results since only a very small portion (2.7% of the deaths) of patients died from HCC (probably because they were in early stages). Our study took part close to a center of LT and a large proportion of patients were listed, so that we included a vast majority of patients with severe prognosis. This explains why death or LT occurred in almost 90% of our study population. The high rate of LT in the group of CP class A patients was probably due to a high rate of HCC in this group.

CONCLUSION

Our results in a large cohort of cirrhotic patients strongly demonstrate that the metabolic liver function tests, ICG and ABT, are useful tools to improve the prognostic evaluation of survival, with a high sensitivity and specificity, similar to the CP and MELD scores. In our experience, they are easy to use, without side-effects and could be developed to routinely evaluate the prognosis of patients in addition to clinical and biochemical data.

Conflicts of interest: None to declare.


REFERENCES


Valoarea prognostică a testelor metabolice funcționale hepatice: un studiu pe 711 pacienți cu ciroză

ABSTRACT / REZUMAT

Premize și scop: Prognosticul pacienților cu ciroză hepatică este de obicei evaluat prin scorurile Child-Pugh și MELD. Testele metabolice funcționale hepatice, ca testul respirator cu aminopirină (ABT) și clearance-ul verdelui de indocianină (IGC) pot evidenția disfuncția hepatică. Scopul acestui studiu retrospectiv este de a compara valoarea prognostică a scorului MELD, Child-Pugh, a testului ABT și a IGC la o serie mare de pacienți cu ciroză hepatică.


Rezultate: Testele metabolice funcționale și scorurile studiate au fost strâns correlate. În momentul evaluării, 111 pacienți erau decesați și 520 fuseseră transplantați. Capacitatea prognostică (apreciată prin curbele AUROC) de a prevedea supraviețuirea la 6 luni fără transplant hepatic a fost de 0,662, 0,691, 0,738 și 0,715 pentru ABT, IGC, scorul Child-Pugh și, respectiv, scorul MELD. În mod similar, evaluarea supraviețuirii la 1 an a constatat curbe AUROC de 0,738 pentru scorul Child-Pugh, 0,716 pentru scorul MELD, 0,693 pentru IGC, respectiv 0,651 pentru ABT.

Concluzii: Rezultatele noastre confirmă faptul că IGC și ABT au o valoare prognostică ridicată pentru pacienții cirotici, similată scorurilor Child-Pugh și MELD. Ele pot fi utilizate de rutină pentru a evalua prognosticul pacienților, împreună cu datele clinice și biochimice.