Research Article

JOURNAL OF HEPATOLOGY

Validation of the AFP model as a predictor of HCC recurrence in patients with viral hepatitis-related cirrhosis who had received a liver transplant for HCC

Andrea Notarpaolo[1,9], Richard Layese[2], Paolo Magistri[3], Maria Gambato[4], Michele Colledan[5], Giulia Magini[5], Lucia Miglioresi[6], Alessandro Vitale[7], Giovanni Vennarecci[8], Cecilia D Ambrosio[6], Patrizia Burra[4], Fabrizio Di Benedetto[3], Stefano Fagiuoli[5], Marco Colasanti[8], Giuseppe Maria Ettorre[8], Arnoldo Andreoli[6], Umberto Cillo[7], Alexis Laurent[9], Sandrine Katsahian[2], Etienne Audureau[2], Françoise Roudot-Thoraval[2], Christophe Duvoux[9,*]

[1]*Arcispedale Santa Maria Nuova, Reggio Emilia, Italy;* [2]*Department of Public Health & Biostatistics, Henri Mondor Hospital, University of Paris-Est, Créteil, France;* [3]*Hepato-Pancreato-Biliary Surgery and Liver Transplantation Unit, Department of General Surgery, University of Modena and Reggio Emilia, Modena, Italy;* [4]*Multivisceral Transplant Unit, Department of Surgery, Oncology and Gastroenterology, Padova University Hospital, Padova, Italy;* [5]*Gastroenterology and Transplant Hepatology, Papa Giovanni XXIII Hospital, Bergamo, Italy;* [6]*UOC di epatologia, Ospedale San Camillo di Roma, Roma, Italy;* [7]*Hepatobiliary Surgery and Liver Transplant Unit, Padova University Hospital, Padova, Italy;* [8]*Multiorgan Transplantation Program-General Surgery and Transplantation Unit, Ospedale San Camillo di Roma, Roma, Italy;* [9]*Liver Transplant Unit- Department of Hepatology, Henri Mondor Hospital-APHP, University of Paris-Est, Creteil, France*

Background & Aims: The AFP model was shown to be superior to the Milan criteria for predicting hepatocellular carcinoma (HCC) recurrence after liver transplantation in a French population. Our aim was to test the AFP model in a non-French, post-hepatitic cirrhosis-based population of HCC candidates.

Methods: 574 patients transplanted for HCC in four Italian centers were studied. AFP score was assessed at the last evaluation before liver transplantation (LT). Probabilities of recurrence and survival were estimated by the log-rank test or competing risk analysis and compared according to the AFP model.

Results: 24.7% patients were beyond Milan criteria. HCC complicated hepatitis C virus (HCV) and hepatitis B virus (HBV) cirrhosis in 58.7% and 24% of the cases, respectively. Five-year probabilities of recurrence differed according to AFP score $\leqslant$2 *vs.* >2 in the whole population (13.2 ± 1.8% *vs.* 49.8 ± 8.7%, p <0.001, HR = 4.98), in patients within Milan criteria (12.8 ± 2.0% *vs.* 32.4 ± 12.1%, p = 0.009, HR = 3.51), beyond Milan criteria (14.9 ± 4.2% *vs.* 58.9 ± 11.5%, p <0.001, HR = 4.26), HCV patients (14.9 ± 2.5% *vs.* 67.6 ± 14.7%, p <0.001, HR = 6.56) and HBV patients (11.6 ± 3.4% *vs.* 34.3 ± 12.5%, p = 0.012, HR = 3.49). By net reclassification improvement analysis AFP score significantly improved prediction of non-recurrence compared to Milan criteria. Overall five-year survival rates according to AFP score $\leqslant$2 or >2 were 71.7 ± 2.2% *vs.* 42.2 ± 8.3% (p <0.001, HR = 2.14).

Conclusions: The AFP model identifies HCC candidates at low risk of recurrence, otherwise excluded by Milan criteria in a population with a predominance of post-hepatitic-related HCC. The AFP score can be proposed for selection of HCC candidates in programs with a high proportion of viral/HCV-related cirrhosis.

Lay summary: Selection criteria for liver transplantation of patients affected with hepatocellular carcinoma (HCC) are based on the Milan criteria, which have been shown to be too restrictive, precluding access to liver transplantation for some patients who might be cured by this operation. Recently, a French group of researchers developed a new selection model called the AFP model, or AFP score, allowing some patients with HCC not meeting Milan criteria to be transplanted with excellent results. In the present work, the AFP score was tested in a population of non-French patients transplanted for HCC occurring mainly on post-hepatitic (HCV or HBV) cirrhosis. The results confirm that in this specific population, as in the original French population of patients, the AFP model better selects patients with HCC eligible for transplantation, compared to Milan criteria. We conclude that the AFP score, which has been officially adopted by the French organization for Organ Sharing for HCC patients, can also be implemented in countries with an important burden of HCC occurring on post-hepatitic cirrhosis.

Keywords: Hepatocellular carcinoma; Liver transplantation; Alfa fetoprotein; AFP; AFP model; AFP score; Predictive model; Validation; Hepatitis; Liver cirrhosis.

Received 30 December 2015; received in revised form 16 September 2016; accepted 31 October 2016; available online 27 November 2016

* Corresponding author. Address: Department of Hepatology and Liver Transplant Unit, Henri Mondor Hospital-Paris-Est Créteil University (UPEC), 51 avenue du Maréchal de Lattre de Tassigny, 94000 Créteil, France. Tel.: +33 149812353; fax: +33 149812352.

E-mail address: christophe.duvoux@aphp.fr (C. Duvoux).

Introduction

Liver transplantation (LT) is considered the best treatment of hepatocellular carcinoma (HCC). However, its efficacy is limited by the risk of tumor recurrence, which results in rapid death and graft loss in patients who are not selected appropriately, making LT futile.

Because of this intrinsic limitation, considerable efforts have been attempted to select HCC candidates with the lowest risk of tumor recurrence. For this purpose, the Milan criteria were proposed 18 years ago [1], and have been adopted by a number of LT programs and centers around the year 2000, notably in the USA. Over the last decade, some groups have reported on expanded HCC criteria, which were associated with an acceptable risk of recurrence of around 10–15% on average [2–9], and with 5-year survival rates similar to those observed after LT for benign liver diseases. These findings indicate that some patients can be transplanted beyond Milan criteria with excellent results and point out that Milan criteria are probably too restrictive. However, no consensus has been achieved on such expanded criteria, which were mostly derived from retrospective analysis of explant pathology with no prospective validation on external cohorts, nor direct comparison to Milan criteria. Therefore, the 2010 international consensus conference on HCC and LT [10] stated that Milan criteria remained the benchmark for selection of HCC patients for LT, and the basis for comparison with any other suggested criteria. Yet, recommendation 10 [10] opened a door to an expanded criteria, provided such criteria would not significantly affect LT for other benign indications.

Recently, the French study group for LT reported on a new predictive model for HCC recurrence, namely the AFP model [11], which was based on tumor staging and AFP values at listing and follow-up time points. Adding AFP to tumor size and number increased the accuracy for predicting recurrence as AFP is a surrogate marker of both tumor differentiation and vascular invasion [11–14], two features which cannot be assessed by conventional imaging-based tumor staging. Accordingly, high AFP levels have been reported to be associated with high recurrence rates [2–3,11,15–16]. The AFP model was shown to be superior to Milan criteria in predicting recurrence [11] in a training set of HCC patients, and was subsequently validated in a cohort of 460 French patients followed prospectively under the control of the French organization for organ sharing (ABM). On these grounds, the AFP model was officially adopted in January 2013 in France by ABM for selecting HCC candidates. However, whether the AFP model may appropriately select non-French HCC candidates, with different distribution of underlying liver diseases, remains unknown. As recently stated by a European expert panel [17], incorporating a biomarker-based predictive model on a large scale deserves confirmation of results using the same technology in external cohorts reported by independent investigators.

The aim of this study, therefore, was to test the predictive value of the AFP model for recurrence and survival in an Italian population of HCC patients which differed from the French cohort by the predominance of HCC complicating post-hepatitic cirrhosis.

Patients and methods

The study population consisted of adult patients who had been listed and had undergone LT for HCC in the centers of Bergamo, Modena, Padova and Roma San Camillo between 2002 and 2010.

Inclusion criteria were: (i) patients listed for HCC diagnosed either on preoperative imaging according to the European Association for the Study of the Liver (EASL)/American Association for the Study of Liver Diseases (AASLD) criteria [17] or on preoperative tumor biopsy; (ii) absence of tumor venous involvement on preoperative ultrasound or CT scan examination of the liver; and (iii) histopathological proof of HCC on the explanted liver.

Exclusion criteria were: (i) incidental HCC; (ii) diagnosis of tumor vascular invasion on preoperative imaging, at listing or during follow-up; (iii) diagnosis of HCC after listing; (iv) patients younger than 18 years of age.

A total of 684 patients were screened to participate in this retrospective study. After exclusion of patients with missing data, the final study population consisted of 574 patients (Modena n = 210, Padova n = 168, Bergamo n = 135 Roma n = 61), the characteristics of which are listed in Table 1.

Data collection

Pretransplant data at listing and post-transplant events

Demographics, cause of cirrhosis, MELD scores, imaging tumor features, type of pre-LT bridging therapies, liver function tests and AFP values were retrospectively collected at listing and during the waiting phase by local investigators. Imaging features of HCC had been collected from imaging reports. Response to treatment after loco-regional therapy was assessed according to mRECIST criteria, taking into account the size and number of the residual viable tumor tissue as assessed in the arterial phase of contrast-enhanced CT or MRI. Pathological features of HCCs, including vascular invasion, tumor differentiation, tumor size and number were collected after LT from pathological reports of the explants. Monitoring and modalities of diagnosis of HCC recurrence have been reported elsewhere [1,18–20]. Post LT follow-up data included death, cause of death, HCC recurrence and dates of last follow-up visit, death or recurrence.

Collection of data

Data were collected by independent local investigators blinded to the final data base and blinded to statistical analysis. Data collection was supervised by AN.

Statistical analysis

AFP model

The AFP score was calculated for each patient enrolled in the study at listing and at last evaluation before LT, using a simplified version of the AFP model (11, Table 1). However, due to a median waiting time of 8.6 months, data and AFP values closest to LT were eventually used to test the ability of the AFP model to predict both recurrence and death.

Probabilities of HCC recurrence and death were estimated and compared according to Milan criteria or the AFP score at a cut-off of 2 by the means of the log-rank test. Hazard ratio between low and high risk groups as defined by either the AFP model or Milan criteria were determined from univariate Cox models. Competing risks analysis [21] was used to compare the probabilities of HCC-related and unrelated deaths. The ability of Milan and AFP models to predict

Table 1. Simplified, user-friendly version of the AFP model.

Variables	β coefficient	Hazard ratio	Points
Largest diameter			
⩽3 cm	0	1	0
3–6 cm	0.272	1.31	1
>6 cm	1.347	3.84	4
Number of nodules			
1–3	0	1	0
4 and more	0.696	2.01	2
AFP level (ng/ml)			
⩽100	0	1	0
[100–1000]	0.668	1.95	2
>1000	0.945	2.57	3

The score is calculated by adding the individual points for each obtained variable. A cut-off of 2 separates between patients at high and low risk of recurrence. In this simplified version, a cut-off of 2 selected exactly the same patients as the original Cox score 0.7 cut-off.
Taken from Duvoux *et al.* Gastroenterology 2012 [11].

recurrence was also tested by the means of net reclassification improvement analysis applied to patients with a minimal follow-up of 2 years [22]. Analysis of histo-pathological features of HCC associated with recurrence post-LT was performed using uni- and multivariate Cox models. Histo-pathological features of HCC were also compared according to AFP scores >or ⩽2 and Milan criteria.

SPSS (V.18) and Stata (V11) software were used for statistical analysis.

Statistical analysis was performed by a team of statisticians (RL, SK, EA), independently of the investigators involved in data collection, after ruling out files with relevant missing data. In addition, the team of statisticians had not been involved in the design and validation of the original score and had no *a priori* on the expected end-points.

Results

Characteristics of the study population

Baseline characteristics of the study population are presented in Table 2. The median MELD score at listing was 12. The majority of HCC complicated post-hepatitic cirrhosis, i.e., HCV cirrhosis in 58.7% and HBV cirrhosis in 24% of the cases.

Of note, causes of liver diseases differed significantly in this cohort from those reported in the original French cohort [11] with a significantly higher number of HCC occurring on post-hepatitic liver diseases and a lower number of HCC complicating alcoholic liver disease in the present cohort, compared to the French one (Supplementary Table 1A, 1B).

Assessment of HCC was performed by contrast-enhanced CT scan, MRI or contrast-enhanced ultrasound in 77.2%, 19.3% and 3.5% of the cases, respectively. Median time [IQ] from last imaging to LT was 2.2 [1.1–4.1] months. AFP value used for calculation of the AFP score was determined a median [IQ] of 5.4 [1.16–11.6] months after listing and 1.2 [0.4–2.8] months before LT. Twenty-five percent of HCCs were beyond Milan criteria at listing and AFP score was >2 in almost 11% of the cases. Median waiting time was 8.6 months and 84.7% of the patients had received loco-regional bridging therapies during the wait phase, including thermo-ablation in 254 cases (associated with transarterial chemoembolization (TACE) in 111 cases, ethanol ablation in 65, and surgical resection in 22 cases), TACE in 160 cases, ethanol ablation in 23 cases and surgical resection in 67 cases (surgery only: 12, combined with loco-regional therapies: 55). Overall, post-operative mortality was 7.7%, and crude incidence of recurrence was 13.5%.

Probabilities of recurrence according to pre-LT AFP values

Five year probability of recurrence significantly differed according to pre-LT AFP thresholds as defined by the AFP model [11], ranging from 13.0 ± 1.8% to 34.9 ± 6.8% and 75.0 ± 15.3%, in patients with pre-LT AFP values ⩽100 ng/ml, [100–1000 ng/ml] and >1000 ng/ml, respectively (Fig. 1), p <0.001.

Overall probabilities of recurrence and survival according to the AFP score cut-off of 2 or Milan criteria

Five-year probability of recurrence, as assessed by Kaplan-Meier estimates were 13.2 ± 1.8% in 512 patients with AFP score ⩽2 *vs.* 49.8 ± 8.7%, in 62 patients with AFP score >2 (p <0.001, HR = 4.98 [3.06–8.10]) (Fig. 2A). Accordingly, 5-year survival rates were 71.7 ± 2.2% *vs.* 42.2 ± 8.3% (p <0.001, HR = 2.14 [1.43–3.20]), among patients with AFP score ⩽2 or >2 (Fig. 2B), indicating that in this cohort the AFP model discriminated appropriately between low and high risk patients for both recurrence and survival. These figures compared favorably with the risk of recurrence as assessed by Milan criteria (Fig. 2C, D). Risks of recurrence in patients within and beyond Milan criteria were 13.6 ± 2.0% and 27.4 ± 4.6%, respectively (p <0.001), with corresponding 5-year survival rates of 73.5 ± 2.3% and 54.3 ± 5.0%, respectively (p = 0.01). Of note, risks of recurrence as assessed by competing risk analysis, taking into account the competing risk of non-HCC-related death (Supplementary material, Fig. 1A, B) were estimated as 11.1 ± 1.0% and 43.0 ± 7.7% (p <0.001) in patients with AFP score ⩽ and >2, and 11.6 ± 1.9% and 22.2 ± 3.8% (p <0.001,) in patients within or beyond Milan criteria, indicating that Kaplan-Meier estimates only slightly overestimated the risk of recurrence. Again, these figures indicated that the AFP model better discriminated between high and low risk patients than Milan criteria. Finally, based on competing risk analysis, probabilities of death not related to HCC recurrence were similar, 20.6 ± 1.9% and 20.1 ± 5.8%, (p = 0.76) in patients with AFP score ⩽2 or >2 (Fig. 3) indicating that differences in survival rates according to the AFP model were actually due to HCC recurrence but not to other causes of deaths.

Probabilities of recurrence according the AFP score cut-off of 2, in patients fulfilling or not Milan criteria (Fig. 4A, B)

Among 432 patients fulfilling Milan criteria, 5-year risk of recurrence was 12.8 ± 2.0% in patients with AFP score <2 and 32.4 ± 12.1% in patients with AFP score >2 (p = 0.009, HR = 3.51 [1.39–8.88]) (Fig. 4A).

Among 142 patients beyond Milan criteria, the risk of recurrence was 14.9 ± 4.2% among patients with an AFP score <2 and 58.9 ± 11.5% in patients with an AFP score >2, (p <0.001, HR = 4.26 [2.10–8.67]) (Fig. 4B). These results show that the AFP score could identify patients at low and high risk of recurrence both in patients fulfilling or not fulfilling Milan criteria.

Table 2. Baseline characteristics of the study population.

Males (n, %)	497 (86.6)
Age at listing/at LT (yr)	55.8 ± 7.5/56.9 ± 7.6
MELD score (median, [IQR])	12 [10–16]
Child-Pugh (A/B/C) (n, %)	196 (34.1%)/268 (46.7%)/110 (19.2%)
Causes of liver disease (HCV/HBV/alcohol/others)	387 (58.7%)/138 (24%)/67 (11.7%)/32 (5.6%)
Number of nodules (median, [IQR]), (range)	2 [1–2], (1–8)
Max diameter (cm) (median, [IQR], (range)	2.5 [2–3.5], (1–21)
AFP (ng/ml) at listing (median, [IQR], (range)	9 [3.9–30.1], (0.4–17,500)
AFP (ng/ml) at last evaluation (median, [IQR], (range)	10.4 [4.3–33.3], (0.5–22,455)
Milan criteria [in/out, (%)]	432/142 (75.3% *vs.* 24.7%)
AFP score: ⩽2 *vs.* >2	512/62 (89.2% *vs.* 10.8%)
Median waiting time (months) [IQR]	8.6 [3.6–16.0]
Bridging therapies (n, %)	486 (84.7)
Post-operative deaths (n, %)	44 (7.7)
Overall recurrence rate (n, %)	81 (13.5)
Follow-up (months) (median, [IQR])	40.9 [18.4–73.6]

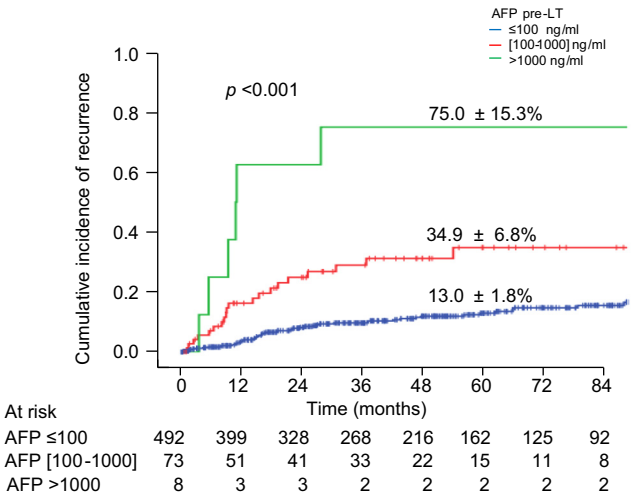

Fig. 1. Risk of recurrence according to pre-LT AFP thresholds as defined in the AFP model. (This figure appears in colour on the web.)

Comparison of AFP model and Milan criteria according to net reclassification improvement

Net reclassification improvement table for recurrence at 2 years is presented in Supplementary Table 2.

The AFP score significantly improved classification of patients without recurrence compared to Milan (net reclassification improvement [NRI] for non-event = 0.137, z = 6.81, p <0.001) confirming that AFP score performed better than Milan criteria to select patients at low risk of recurrence, even among patients exceeding Milan criteria. However, global NRI was not significantly different between the two scores (NRI = 0.0303, z = 0.434, n.s.) because NRI for event was similar for Milan criteria and AFP score.

Probabilities of recurrence and survival in HCV and HBV patients (Fig. 5)

In the subgroup of 337 patients transplanted for HCV-related HCC, the 5-year risk of recurrence was 14.9 ± 2.5% in patients with AFP score ⩽2 and 67.6 ± 14.7% in patients with AFP score >2 (p <0.001, HR = 6.56 [3.61–11.92]) (Fig. 5A). Corresponding 5-year survival rates in HCV patients were 67.8 ± 3.0% and 25.6 ± 11.0% (p <0.001) (Fig. 5B). Similar results were found in the subgroup of 138 patients transplanted for HBV-related HCC in terms of recurrence, (p = 0.012, HR = 3.49 [1.23–9.93]) (Fig. 5C) although 5-year survival rates according to AFP score did not achieve statistical significance (Fig. 5D).

Probability of recurrence and survival in 46 patients transplanted after down staging from AFP score >2 to AFP score <2.

The features of 46 patients with successful down staging from AFP score >2 to <2 after loco-regional therapy are shown in Supplementary Tables 3A and B.

Median AFP scores before and after loco-regional therapy were 3 (3.00; 3.00) and 0 (0.00; 1.00), respectively; 30.4% of the

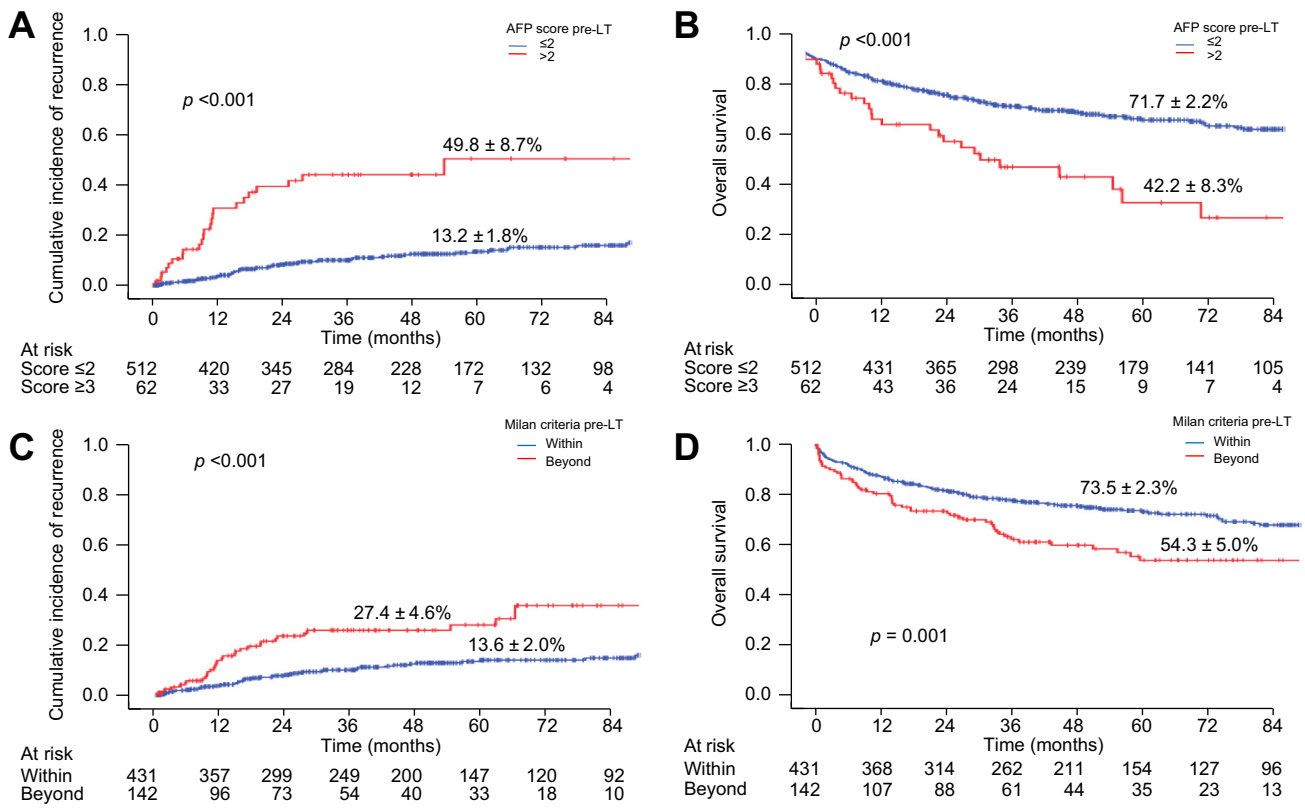

Fig. 2. Overall probabilities of recurrence and survival. Probabilities according to the AFP score cut-off of 2 (A and B) or Milan criteria (C and D). (This figure appears in colour on the web.)

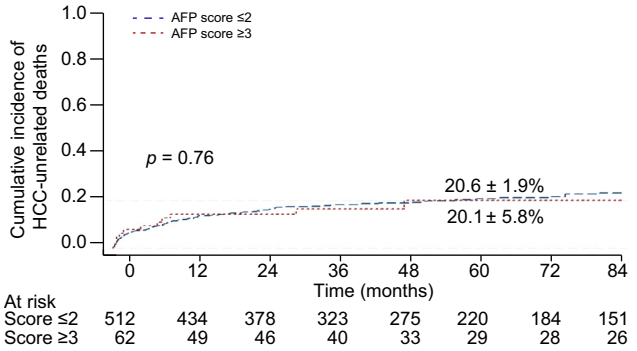

Fig. 3. Probabilities of death not related to HCC recurrence as assessed by competing risk analysis. Data according to the AFP model. (This figure appears in colour on the web.)

patients remained out of Milan criteria but with AFP score ⩽2 after down staging. The majority of patients had been down staged by means of percutaneous ablation techniques, in combination with TACE in nearly half of them.

The median time from down staging procedure to transplantation was 81.00 (22.00; 148.00) days, i.e., nearly 3 months.

According to competing risks analysis, the 5-year risk of recurrence was 16.4 ± 5.7% (Supplementary Fig. 2), with a corresponding overall 5-year survival rate of 71.8 ± 7.0% (Supplementary Fig. 3)

The pathological features of tumors after down staging are summarized in Supplementary Table 3B. Features of those tumors were quite similar to those of the whole group of patients with AFP ⩽2 in terms of number and size of nodules, and also for prevalence of micro-vascular invasion and poor differentiation on the explant (see below).

HCC pathological features according to AFP score and comparison of AFP model and Milan criteria according to explant findings

Explant-based tumor features according to AFP score are summarized in Table 3. Multivariate analysis of histo-pathological predictors of recurrence show that micro-vascular invasion (OR 4.02 [2.51–6.44], p <0.001) and poor tumor differentiation (OR 1.98 [1.24–3.15], p = 0.004) were significantly associated with the risk of recurrence. Risks of micro-vascular invasion and poor differentiation were higher in patients with AFP score >2 than in patients with AFP score ⩽2. In addition, tumor size was larger and tumor number was higher in patients with AFP score >2 than in patients with AFP score ⩽2.

Comparisons of histo-pathological features of HCC according to Milan criteria and AFP scores are shown in Supplementary Table 4. In patients with AFP score >2 (high risk of recurrence), prevalence of both micro- and macro-vascular invasion as well as poor differentiation were high and did not differ whether HCCs were in or out Milan criteria on the explant. In particular, prevalence of micro-vascular invasion and poor differentiation for HCC within Milan criteria but with AFP score >2 were 46.7% and 60%, respectively. This reflected a better association of the AFP score with high risk pathological predictors of poor prognosis, compared to Milan criteria. In patients with AFP score ⩽2, prevalence of macro-vascular invasion and poor differentiation did not differ whether HCC were in or out Milan, this again indicates a better association of AFP score with these two pathological predictors

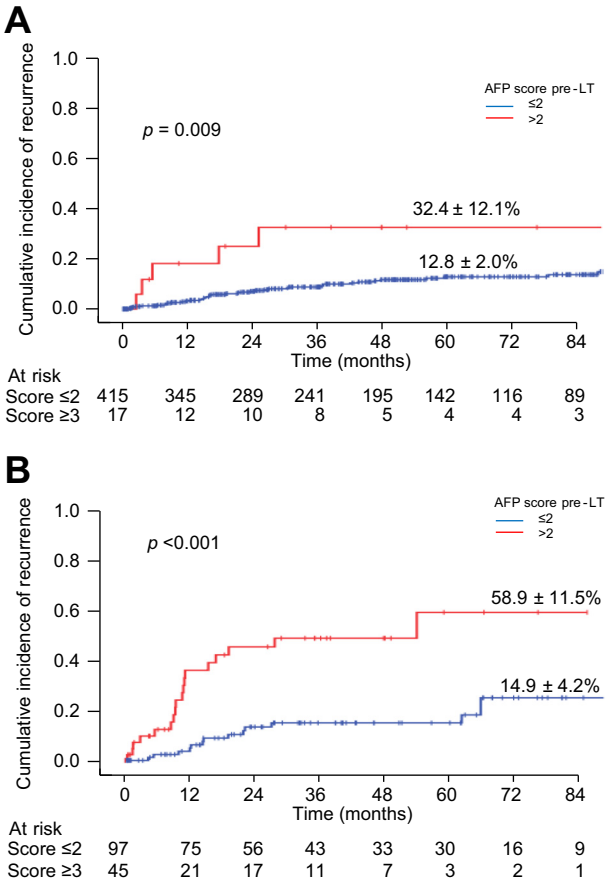

Fig. 4. Probabilities of recurrence according to the AFP score cut-off of 2, in patients fulfilling or not Milan criteria. (A) Patients within Milan criteria. (B) Patients beyond Milan criteria. (This figure appears in colour on the web.)

of recurrence, compared to Milan criteria. Yet, in patients with AFP score ⩽2, the prevalence of micro-vascular invasion was higher in patients beyond than within Milan criteria.

Discussion

Over the last decade, an increasing perception has emerged among the community of LT teams that Milan criteria, which were adopted almost two decades ago as a selection tool for HCC candidates, have become too restrictive [10,23]. However, although several new selection criteria have been proposed for expanding HCC indications [2–9], no consensus has been achieved so far for their use in clinical practice [23]. In the present study, we tested the predictive value for recurrence of the AFP model in an Italian population of HCC candidates [11], a recently proposed prognostic tool which was designed in a French training cohort of HCC candidates, and tested further in an external, prospectively followed, validation set. The AFP model has been shown to be more accurate than Milan criteria for selecting HCC candidates in this French population and as a results, has been adopted as an official selection tool by the French organization for organ sharing (ABM) by 2013. Of note, the Italian cohort of HCC candidates differed from the French population because of a much higher proportion of HCC resulting

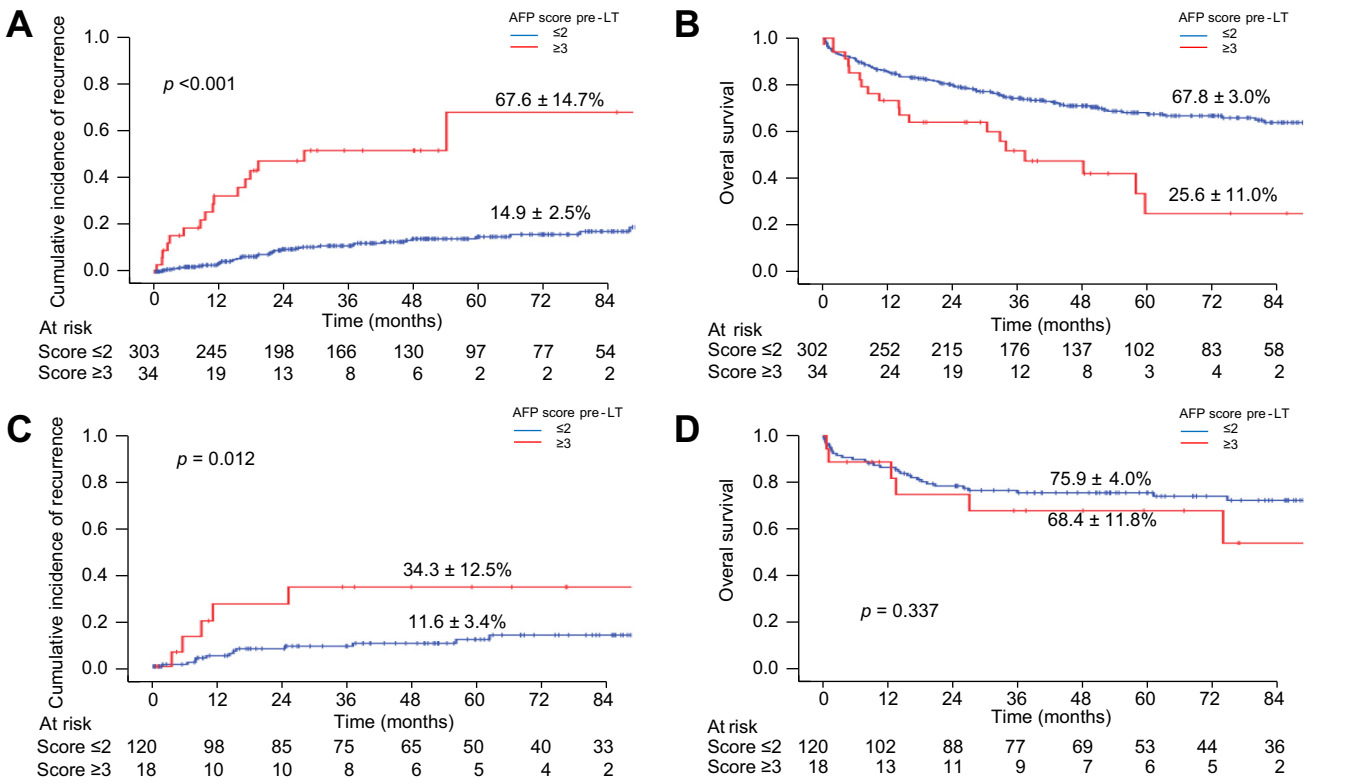

Fig. 5. HCC recurrence in hepatitis populations. Probabilities of recurrence (A and C) and overall survival (B and D), according to the AFP score cut-off of 2 in the HCV population (A and B) and in the HBV population (C and D). (This figure appears in colour on the web.)

Table 3. Explant-based comparison of pathological features of HCC according to the AFP model.

	AFP score ⩽2 n = 512	AFP score >2 n = 62	*p* value
Macro-vascular invasion [n, (%)]	16 (3.2)	5 (8.3)	0.051
Micro-vascular invasion [n, (%)]	96 (19.4)	27 (45.0)	<0.001
Poorly differentiated tumour [n, (%)]	116 (32.2)	28 (51.9)	0.009
Number of nodules (median, [IQR])	2 [1–3]	3 [1–5]	0.001
Diameter of nodules (median, [IQR])	2.5 [1.8–3.5]	4.5 [2.5–6]	<0.001

from post-hepatitic cirrhosis, with a 58% prevalence of HCV-related cirrhosis and 24% prevalence of HBV cirrhosis. The aim of this study was therefore to test the AFP model in an additional external cohort of HCC candidates which differed from the original one in order to ensure consistency. Subject to the retrospective design of the study, the results presented herein show that, as in the French cohort, the AFP model could discriminate correctly and more accurately than Milan criteria between patients at low and high risk of recurrence in the Italian, HCV/HBV-based population. The 5-year incidence of recurrence and probability of survival were significantly better among patients with AFP score ⩽2 than in patients with AFP score >2: 13.2 ± 1.8% and 71.7 ± 2.2% *vs.* 49.8 ± 8.7% and 42.2 ± 8.3%, respectively (p <0.001). In addition, competing analysis censoring HCC-related deaths have shown that the 5-year incidence of HCC unrelated deaths were similar in patients with low and high AFP scores (19.0% *vs.* 21.9%, n.s.). This finding demonstrated that the better survival observed in patients with AFP score ⩽2 was not due to a lower incidence of HCC unrelated deaths but actually to a lower incidence of recurrence.

The lower incidence of recurrence and higher survival rates in patients with AFP score ⩽2 were observed whether patients met Milan criteria or not. In particular, an AFP score ⩽2 identified a subgroup of patients with a low 5-year 14.9 ± 4.2% risk of recurrence, among patients beyond Milan criteria. On the other hand, AFP score >2 identified a subgroup of patients with quite a high 5-year risk of recurrence of 32.4 ± 12.1%, among patients within Milan criteria. This latter finding indicates that special attention should be paid to patients within Milan criteria and high AFP levels at listing before considering them fully eligible for transplantation. Indeed, among such patients, those with AFP levels >1000 ng/ml should be considered at high risk for recurrence, a finding already observed in the French cohort. A careful down staging strategy to AFP score ⩽2 in this subgroup of patients can reasonably be advised before considering LT. Indeed, the results shown in the subgroup of patients who underwent a successful down staging procedure before LT

indicate that a reasonable risk of recurrence (i.e.,16.4%) and excellent 5-year survival rate (i.e., 71.8%) may be achieved after down staging to AFP score <2. Our results also confirm that AFP brings up additional information about tumor behavior, compared to imaging staging, making possible the identification of aggressive tumors despite reasonable tumor size and number. Analysis of the relationship between AFP scores and histopathological features of HCC was in agreement with this finding: HCCs with AFP scores >2 had significantly more aggressive pathological features than HCC with scores <2. This was observed not only in the whole population but also in patients within Milan criteria: AFP score >2 was associated with 46.7% and 60% prevalence of micro-vascular invasion and poor differentiation respectively in this subgroup of patients.

Interestingly, the AFP model performed in a population, which differed notably from the French population in whom it has been developed and tested originally. A stated above, prevalence of HCC complicating post-hepatitic cirrhosis was >80% in the Italian cohort *vs.* 44% in the French validation set. However in the HCV population, accounting for almost 60% of the Italian cohort, 5-year probabilities of recurrence were 14.9 ± 2.5% *vs.* 67.6 ± 14.7% in patients with AFP score ⩽2 or >2 (p <0.001) with corresponding highly different survival rates in this group (67.8 ± 3.0% *vs.* 25.6 ± 11.0% in patients with AFP score ⩽or >2 (p <0.001)), indicating that the AFP model prediction was independent of liver disease etiology and may therefore be applied in programs with a majority of HCV-related HCCs. The reason why the incidence of recurrence in the HCV population with AFP score >2 was particularly high is unclear and further comparisons of pathological features of HCC in the HCV *vs.* non-HCV populations according to the AFP model are ongoing. A similar although less pronounced trend was observed in HBV-related HCCs. In this subgroup, 5-year HCC recurrence rate was significantly higher in patients with AFP score >2 compared to AFP score <2. However, 5-year survival rate although lower in patients with AFP score >2 did not achieve statistical significance. This might be due to the small number of patients in this subgroup with only 18 HBV patients with AFP score >2.

An important issue is also to determine whether adopting the AFP model may significantly impact the burden of HCC candidates and may further increase the competition with non-HCC patients. The results presented herein show that in programs strictly adopting Milan criteria, expanding selection criteria to AFP model may result in a 14% increase in the number of patients eligible for LT (in this present series, 80/574 (14%) patients were beyond Milan criteria but had AFP score ⩽2). However, denying LT to such candidates appears no longer ethical given the excellent, 72%, 5-year survival rate observed in the AFP score ⩽2 patients. To balance the limited expansion of indications of LT for HCC resulting from adoption of the AFP model, additional allocation rules for HCC patients should be encouraged, based on baseline staging of HCC and responses to bridging therapies as recently implemented in the French program. On the other hand, in programs not strictly based on Milan criteria, the AFP model gives the opportunity of a better selection of high risk patients and therefore reduces the probability of futile transplantation for HCC. As so the AFP model has recently been strongly discussed by the UK LT program for selection of HCC candidates [24]. Recent data from Latin America also give additional background to support the use of the AFP model [25].

Although more accurate than Milan criteria for prediction of recurrence, the AFP score deserves further improvement. Some patients with AFP score >2 do not recur and it is of utmost importance to identify them more specifically. Future research aiming at improving prediction of recurrence of HCC before LT is therefore mandatory. Extensive analysis of larger data sets, new predictive models integrating functional imaging [26–28] or/and molecular tools may overcome this issue in the future.

In conclusion, the AFP model which was designed in a French population also performs appropriately in an Italian cohort, characterized by a large predominance of HCV-related HCCs. As in the French population, the AFP model discriminates between patients with low and high risk of recurrence, both in patients within and beyond Milan criteria, indicating a better accuracy than Milan criteria for selection of HCC candidates. This study therefore shows that the performance of AFP model is reproducible and fulfills recommendations of the European expert panel [17] for incorporating AFP and the AFP model in the clinical management of HCC candidates. This important finding strongly supports the adoption of the AFP model as a selection tool for HCC patients in programs with a high proportion of HCC complicating post-viral/HCV cirrhosis.

Financial support

This investigators-driven study was supported by the Association Henri Mondor de Recherche en Pathologie Hépatique et Digestive.

Conflict of interest

The authors who have taken part in this study declared that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

Authors' contributions

AN, CD and FRT designed the study; AN, AB, PM, MG, GM, LM, AV, GV, CDA, MC and GM carried out data collection; RL, SK and EA analyzed the results and conducted the statistical analysis; AN and CD drafted the publication. All authors significantly contributed to the manuscript.

Acknowledgements

The authors thank all Italian participating centers and investigators for their contribution and G Bizouard for statistical advice. They also thank the Association Henri Mondor de Recherche en Pathologie Hépatique et Digestive for supporting this study.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhep.2016.10.038.

References

Author names in bold designate shared co-first authorship

[1] Mazzaferro VV, Regalia EE, Doci RR, Andreola SS, Pulvirenti AA, Bozzetti FF, et al. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. N Engl J Med 1996;334:693–699.

[2] Yao FY, Ferrell L, Bass NM, Watson JJ, Bacchetti P, Venook A, et al. Liver transplantation for hepatocellular carcinoma: expansion of the tumor size limits does not adversely impact survival. Hepatology 2001;33:1394–1403.

[3] Toso C, Trotter J, Wei A, Bigam DL, Shah S, Lancaster J, et al. Total tumor volume predicts risk of recurrence following liver transplantation in patients with hepatocellular carcinoma. Liver Transpl 2008;14:1107–1115.

[4] Yao FY, Xiao L, Bass NM, Kerlan R, Ascher NL, Roberts JP. Liver transplantation for hepatocellular carcinoma: validation of the UCSF-expanded criteria based on preoperative imaging. Am J Transplant 2007;7:2587–2596.

[5] Ito T, Takada Y, Ueda M, et al. Expansion of selection criteria for patients with hepatocellular carcinoma in living donor liver transplantation. Liver Transpl 2007;13:1637–1644.

[6] Herrero JI, Sangro B, Pardo F, et al. Liver transplantation in patients with hepatocellular carcinoma across Milan criteria. Liver Transpl 2008;14:272–278.

[7] Onaca N, Davis GL, Goldstein RM, Jennings LW, Klintmalm GB. Expanded criteria for liver transplantation in patients with hepatocellular carcinoma: a report from the International Registry of Hepatic Tumors in Liver Transplantation. Liver Transpl 2007;13:391–399.

[8] Kneteman NM, Oberholzer J, Al Saghier M, et al. Sirolimus-based immunosuppression for liver transplantation in the presence of extended criteria for hepatocellular carcinoma. Liver Transpl 2004;10:1301–1311.

[9] Mazzaferro V, Llovet JM, Miceli R, et al. Predicting survival after liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: a retrospective, exploratory analysis. Lancet Oncol 2009;10:35–43.

[10] Clavien PA, Lesurtel M, Bossuyt PM, et al. Recommendations for liver transplantation for hepatocellular carcinoma: an international consensus conference report. Lancet Oncol 2012;13:e11–e22.

[11] **Duvoux C**, **Thoraval FR**, **Decaens T**, Pessione F, Badran H, Piardi T, et al. Liver transplantation for hepatocellular carcinoma: a model including α-fetoprotein improves the performance of Milan criteria. Gastroenterology 2012;143:e3.

[12] Fujioka M, Nakashima Y, Nakashima O, Kojiro M. Immunohistologic study on the expressions of alpha-fetoprotein and protein induced by vitamin K absence or antagonist II in surgically resected small hepatocellular carcinoma. Hepatology 2001;34:1128–1134.

[13] Fujiki M, Takada Y, Ogura Y, Oike F, Kaido T, Teramukai S, et al. Significance of des-gamma-carboxy prothrombin in selection criteria for living donor liver transplantation for hepatocellular carcinoma. Am J Transplant 2009;9:2362–2371.

[14] Liu C, Xiao GQ, Yan LN, Li B, Jiang L, Wen TF, et al. Value of α-fetoprotein in association with clinicopathological features of hepatocellular carcinoma. World J Gastroenterol 2013;19:1811–1819.

[15] Ravaioli M, Grazi GL, Piscaglia F, Trevisani F, Cescon M, Ercolani G, et al. Liver transplantation for hepatocellular carcinoma: results of down-staging in patients initially outside the milan selection criteria. Am J Transplant 2008;8:2547–2557.

[16] Hameed B, Mehta N, Sapisochin G, Roberts JP, Yao FY. Alpha-fetoprotein level >1000 ng/mL as an exclusion criterion for liver transplantation in patients with hepatocellular carcinoma meeting the Milan criteria. Liver Transpl 2014;20:945–951.

[17] European Association for the Study of the LiverEuropean Organization for Research and Treatment of Cancer. EASL-EORTC Clinical Practice Guidelines: Management of hepatocellular carcinoma. J Hepatol 2012;56:908–943.

[18] Decaens T, Roudot-Thoraval F, Bresson-Hadni S, et al. Impact of pretransplantation transarterial chemoembolization on survival and recurrence after liver transplantation for hepatocellular carcinoma. Liver Transpl 2005;11:767–775.

[19] Decaens T, Roudot-Thoraval F, Bresson-Hadni S, et al. Role of immunosuppression and tumor differentiation in predicting recurrence after liver transplantation for hepatocellular carcinoma: A multicenter study of 412 patients. World J Gastroenterol 2006;12:7319–7325.

[20] Decaens T, Roudot-Thoraval F, Hadni-Bresson S, et al. Impact of UCSF criteria according to pre- and post-OLT tumor features: Analysis of 479 patients listed for HCC with a short waiting time. Liver Transpl 2006;12:1761–1769.

[21] Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. JASA 1999;94:496–509.

[22] Pencina MJ, D'Agostino Sr RB, D'Agostino Jr RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008;27:157–172, [Discussion 207–212].

[23] Dutkowski P, Linecker M, DeOliveira ML, Müllhaupt B, Clavien PA. Challenges to liver transplantation and strategies to improve outcomes. Gastroenterology 2015;148:307–323, [NHS blood and transplant liver advisory group].

[24] Liver transplantation for HCC in the UK – Report from a National Consensus Meeting (Birmingham, January 2014): LT for HCC-Consensus meeting summary. LAG(14)9b. Online access: http://www.odt.nhs.uk/pdf/advisory_group_papers/LAG/HCC_recommendations_IR_TS_b_NAS_Work_in_Progress.pdf.

[25] Piñero F, Tisi Baña M, de Ataide EC, Hoyos Duque S, Marciano S, Varón A, et al. Liver transplantation for hepatocellular carcinoma: evaluation of the alpha-fetoprotein model in a multicenter cohort from Latin America. Liver Int 2016. http://dx.doi.org/10.1111/liv.13159.

[26] Lee JW, Paeng JC, Kang KW, Kwon HW, Suh KS, Chung JK, et al. Prediction of tumor recurrence by 18F-FDG PET in liver transplantation for hepatocellular carcinoma. J Nucl Med 2009;50:682–687.

[27] Kornberg A, Küpper B, Tannapfel A, Büchler P, Krause B, Witt U, et al. Patients with non-[18 F]fludeoxyglucose-avid advanced hepatocellular carcinoma on clinical staging may achieve long-term recurrence-free survival after liver transplantation. Liver Transpl 2012;18:53–61.

[28] Hong G, Suh KS, Suh SW, Yoo T, Kim H, Park MS, et al. Alpha-fetoprotein and (18)F-FDG positron emission tomography predict tumor recurrence better than Milan criteria in living donor liver transplantation. J Hepatol 2016;64:852–859.