Published in final edited form as:
J Hepatol. 2012 May ; 56(5): 1112–1120. doi:10.1016/j.jhep.2011.11.020.

Alpha-fetoprotein Response Correlates with EASL Response and Survival in Solitary Hepatocellular Carcinoma Treated with Trans-arterial Therapies: A Subgroup Analysis

Khairuddin Memon, MD[1], **Laura Kulik, MD**[2], **Robert J Lewandowski, MD**[1], **Edward Wang, PhD**[3], **Robert K Ryu, MD**[1], **Ahsun Riaz, MD**[1], **Paul Nikolaidis, MD**[1], **Frank H Miller, MD**[1], **Vahid Yaghmai, MD**[1], **Talia Baker, MD**[3], **Michael Abecassis, MD MBA**[3], **Al B Benson 3rd, MD**[4], **Mary F Mulcahy, MD**[4], **Reed A Omary, MD**[1], and **Riad Salem, MD MBA**[1,3,4]

[1]Department of Radiology, Section of Interventional Radiology, Robert H. Lurie Comprehensive Cancer Center, Northwestern University, Chicago IL

[2]Department of Medicine, Division of Hepatology, Northwestern University, Chicago, IL

[3]Department of Surgery, Division of Transplantation, Comprehensive Transplant Center, Northwestern University, Chicago, IL

[4]Department of Medicine, Division of Hematology and Oncology, Robert H. Lurie Comprehensive Cancer Center, Northwestern University, Chicago, IL

Abstract

Background and Aims—Alpha-fetoprotein (AFP) is a universally recognized tumor marker in hepatocellular carcinoma (HCC). Its utility in assessing response to treatment remains controversial. We sought to study the: a) correlation between AFP response and imaging response, and b) ability of AFP, EASL and WHO response to predict survival outcomes in patients with solitary HCC.

Methods—629 HCC patients were treated with transarterial locoregional therapies over an 11-year period. To eliminate confounding factors, we included patients with single tumors, baseline AFP≥200 ng/mL, and no extrahepatic disease; this identified our study cohort of 51 patients. AFP response was defined as >50% decrease from baseline; this was correlated to EASL and WHO response criteria by Kappa agreement, Pearson correlation and receiver operating curves. Survival analyses were performed by Landmark, risk-of-death and Mantel-Byar methodologies. None of the patients received sorafenib.

Results—Three months post-treatment, AFP and EASL response correlated well (Kappa: 0.83; Pearson: 0.84); the sensitivity, specificity, positive and negative predictive values of AFP in predicting EASL response at 3 months were 96.6%, 85.7%, 92.3% and 93.3% respectively. Correlation with WHO response was low. From the 3-month landmark, WHO, EASL and AFP responders survived longer than nonresponders (P=0.006, 0.0001 and <0.0001 respectively). The

Corresponding Author: Riad Salem, MD, MBA Director, Interventional Oncology Department of Radiology 676 N. St. Clair, Suite 800 Chicago, Illinois 60611 USA Tel: 312-695-6371 Fax: 312-695-5654 r-salem@northwestern.edu.

Conflict of Interest: None of the authors report a conflict of interest.

risk of death was lower for EASL and AFP responders by both risk-of-death and Mantel-Byar methodologies (P<0.05).

Conclusion—Response by AFP and EASL are predictors of survival outcome in patients with solitary HCC. AFP correlates with imaging response assessment by EASL guidelines. Achieving AFP response should be one of the therapeutic intents of locoregional therapies.

Keywords

transarterial chemoembolization; radioembolization; hepatocellular carcinoma; imaging response, AFP response, correlation, survival

INTRODUCTION

The incidence of HCC is increasing;[1] it has tripled between 1975 and 2005.[2] Most patients present at an advanced stage beyond curative therapies, with sorafenib prolonging survival in advanced HCC.[3, 4] LRTs play a palliative role by inducing tumor necrosis, delaying progression and improving survival.[5-14] Following HCC treatment, it is the clinical standard of care to follow patients with CT/MR imaging. The utility of tumor markers to assess response, such as AFP, remains controversial.

AFP is the only universally recognized tumor marker for hepatocellular carcinoma. It has been investigated as a potential screening, diagnostic and a prognostic tool.[15-17] Several studies have reported the capability of AFP response in prognosticating response to therapy and survival outcomes. Riaz et al demonstrated that AFP response to LRTs can be used for assessing tumor response, time-to-progression and overall survival.[18] Such studies have also been reported with resection, chemotherapy and radiofrequency ablation.[19-21]

The observation of response to any treatment by imaging or AFP is time-dependent.[22] Since treatment algorithms for HCC using LRTs are based on staged sessions separated by weeks/months, it is of interest to correlate these variables in a time-dependent fashion. Does AFP response correlate with imaging response, or is it better able to predict survival than imaging response? [10, 23] Establishing a correlation between AFP and imaging response has the potential to help assess response in clinical scenarios where standard cross-sectional imaging findings are equivocal.

Recently, 3 novel statistical methods were used to demonstrate the importance of imaging response in HCC; the study concluded that tumor response was a potentially significant surrogate of survival.[22] Given the well-known difficulties in assessing treatment response in HCC (inter-observer subjectivity, scan thickness, variable enhancement, regenerative/dysplastic nodules, perfusional abnormalities), we hypothesized that AFP response (objective, no interobserver variability) may provide a simple, reproducible and potentially less subjective method of response assessment.[10, 23] We performed a comprehensive study addressing whether: a) AFP correlates with imaging response by WHO and EASL methodologies, and b) if AFP response can predict improved survival.

METHODS

This study was compliant with the Health Insurance Portability and Accountability Act and approved by the Northwestern University Institutional Review Board. Between 2000-2010, 629 HCC patients were treated with LRTs (^{90}Y: N=406; TACE: N=223); this constitutes the source population. Patients were eligible for LRTs if they exhibited unresectable HCC and bilirubin <3.0 mg/dL (discussed at weekly multidisciplinary HCC conference). To create the study population for this specific analysis, we selected patients who: a) had solitary tumors,

b) expressed baseline AFP >200 ng/mL and, c) did not have extrahepatic metastases (Figure 1: Flow chart). This minimized the number of patients with unknown confounding variables and maximized the number of patients reaching the 3 and 6-month landmarks. This methodology in HCC has recently been thoroughly described. [22, 24, 25] Furthermore, given that LRTs in multifocal HCC are performed as staged procedures, we eliminated patients with multifocal and extrahepatic disease in order to exclude the effect of AFP production by untreated disease and metastatic foci.[18] This resulted in the identification of our study population comprised of 51 patients with solitary HCC, no metastases and AFP >200 ng/mL. In such a cohort, survival becomes dependent on the HCC and background cirrhosis. Survival outcomes were studied with respect to AFP and imaging response using Landmark, risk-of-death and Mantel Byar methodologies.[22] 36 patients had died at the time of data closure. 13 patients received liver transplantation; their survival was censored at transplantation. In order to exclude the effect of transplantation on survival, a survival sub-analysis was performed in non-transplanted patients.

Evaluation/Staging

All patients underwent pretreatment assessment consisting of history, laboratory and imaging work-up. Diagnosis of HCC was made by following guidelines.[26] Baseline staging was performed using CP, UNOS and BCLC classification systems.[26] Patients were categorized as having portal hypertension if they exhibited varices, splenomegaly and/or thrombocytopenia (defined as <100,000/μL).

Locoregional Therapies

Chemoembolization was performed using 30 mg mitomycin, 30 mg adriamycin and 100 mg cisplatinum followed by embolization using 300-500 micron particles per previous reports.[8] Radioembolization was performed using glass microspheres per previous methodology.[27, 28]

Patient follow-up

Patients were followed using CT or MRI at scheduled 3 and 6 months as our standard of care. The median number of treatment sessions per patient was 1 (1, 2 and 3 treatments in 32, 15 and 4 patients, respectively). None of the patients who received ^{90}Y were retreated with TACE or vice versa; no patient received sorafenib. Subsequent to initial LRT, 4 patients received RFA at 1, 1, 4 and 7 years, respectively. One patient underwent surgical resection 1 year after treatment. These treatments did not affect response assessment at 3 and 6 months.

AFP and liver function testing coincided with the imaging scans (permitting time-dependant correlation). Patients were scanned by MRI (our institutional standard) or CT (in case of pacemakers, claustrophobia). For each patient, the imaging modality remained the same throughout the study period. Our protocols MR and CT have been described.[22]

Methodology of Response Assessment

AFP response—All patients exhibited baseline AFP ≥200 ng/mL. The rationale for selecting an AFP cut-off of 200 ng/mL includes: a) AFP>200 has been part of the AASLD guidelines to diagnose HCC,[29] b) this cutoff has been reported in another analysis relating AFP to response, TTP and survival,[18] c) it was deemed necessary to select a value of AFP sufficiently high that in its presence, HCC was likely present and active but not too low that the AFP level could be fluctuating as part of underlying cirrhosis, and d) a higher cut-off would have reduced the analyzable patient population. Hence, we chose a cut-off of 200 to achieve balance between study population and reasonable sensitivity/specificity of AFP. We

defined >50% AFP decrease as AFP responders; patients with <50% reduction or any increase were AFP nonresponders (Supplementary Table 1).[18] For assessment of response, AFP levels were obtained at the same time-points as cross-sectional imaging (3, 6 months). We also performed a quantitative sub-analysis at 3-months post treatment, where AFP response was defined as decreasing from >500 ng/mL to <500 ng/mL (Supplementary Table 1). To investigate if the magnitude of AFP reduction was dependent on baseline AFP level, the percentage reduction in AFP levels were analyzed at 3 and 6 months stratified by baseline AFP level of 200-1000, 1000-3000 and >3000 ng/mL. Median AFP at baseline, 3 and 6 months were also studied.

Imaging Response—Response status was assessed using WHO and EASL guidelines (Supplementary Table 1) at 3 and 6 months using the index (biomarker) lesion concept.[3, 10, 30-32] Patients with CR or PR were categorized as responders; those with SD or PD were categorized as nonresponders.[22] In order to report most conservatively, PVT did not affect the assessment of response in the index lesion, only progression (Supplementary Table 2). That is, a stable lesion by WHO/EASL with retracting/disappearing PVT following treatment was reported as SD. On the other hand, a stable lesion by WHO/EASL with progressing PVT following treatment was reported as PD.

Response assessment was performed by two board-certified radiologists expert in HCC (one specialist in cross-sectional imaging, one interventional radiologist) with blinding to AFP and survival outcomes. WHO/EASL response were correlated with survival. WHO (not RECIST) guidelines were used since we have previously demonstrated the high inter-method correlation between WHO/RECIST.[10] It is also the oncologic gold standard to report tumor size in two (not one) dimensions. EASL methodology is the bidimensional equivalent to the recently described mRECIST.[33]

Statistical Analyses

AFP-Imaging response correlation—For this study, imaging response was the gold standard, AFP was the test variable. We performed statistical correlation of AFP and imaging response using: 1) Kappa (κ) agreement, 2) Pearson coefficient (r) and 3) Receiver Operating Characteristics [to assess sensitivity, specificity, NPV, PPV for AFP response in predicting imaging response-only at 3 months].[34, 35]

Survival analysis—Sherman previously commented on AFP-Imaging correlation, stating that AFP response was only relevant if it correlated with survival outcomes better than the imaging.[36] To overcome guarantee-time bias, we assessed survival of AFP responders vs. nonresponders using 3 novel statistical tools.[24, 25, 37] These methods have been detailed elsewhere.[22, 24, 25, 38]

i. Landmark Method: Survival is calculated from the landmark, thereby eliminating patients with unfavorable biology. It decreases the probability of effects caused by unknown confounding variables on survival. We selected the 3 and 6-month landmarks; these were deemed clinically relevant.[24] Survival analysis was based on AFP and imaging (WHO/EASL) response status at each landmark (Supplementary Table 1). Survival curves were plotted by Kaplan-Meier and compared using the log-rank test.[39]

We performed two exploratory survival sub-analyses based on AFP response: 1) considering quantitative AFP response (i.e. from >500 ng/mL pre-treatment to <500 ng/ mL post-treatment); and 2) analyzing survival based on 50% AFP reduction after excluding patients with PVT.

ii. Risk-of-Death: This method compares the death rate by AFP and imaging response status in the 6 months following each landmark. The chi-square test was used to compare death-rates.

iii. Mantel-Byar Method: [38] Detailed methodology is described elsewhere.[22] It includes all patients from day 0 and treats response as a time-dependent covariate; all patients enter the study in the 'no-response' state. With time, AFP or imaging responders shift to 'responder' status; responders may ultimately progress and shift back to 'non-responder' state. At every endpoint (death), the number of patients in each response category (responder and nonresponder) is calculated, the risk-of-death is estimated for each response category, and a cumulative risk-of-death is generated.[38] Since no accepted definition of AFP progression exists, patients who AFP responders continued as 'responders' without ever shifting to a 'nonresponder' state (even if AFP increased significantly). The statistical significance is determined based on the difference in expected and actual deaths for responders and nonresponders, thereby minimizing biases and comparing patients dynamically by response status at multiple periods of time.

Uni/Multivariate analysis—Uni/multivariate analyses were performed using Cox proportional hazards model at each landmark, investigating whether survival was affected by imaging or AFP response status, liver function and disease stage (at the landmark). Hazard ratio estimates were based on simultaneous analysis of all variables. P-values <0.05 on univariate analyses were corrected for multiple comparisons using Bonferroni methodology.[40, 41] Variables included in univariate analysis were AFP, WHO and EASL response status and bilirubin/albumin. AFP 200-1000 or >1000 ng/mL was included to assess if higher baseline AFP influenced survival outcomes irrespective of imaging or AFP response. Baseline tumor size and presence/absence of PVT were added to investigate the effect of tumor characteristics on outcomes. Variables with P<0.25 (with Bonferroni correction) on univariate analysis were included in the multivariate analysis. All analyses were conducted using SAS 9.2 (SAS, Cary, NC). P<.05 was considered significant.

RESULTS

Baseline Characteristics

Table 1 describes the baseline characteristics. 28 (55%) were ≥65, 30 (59%) were male, and 48 (94%) were treatment naive. Eighteen patients were diagnosed by biopsy. Tumor grade information was only available for 9 patients (well-differentiated: N=3, moderately differentiated: N=2, poorly-differentiated: N=4). Baseline imaging, laboratory characteristics and cancer stages are also summarized.

AFP-Imaging response correlation

Table 2 summarizes the correlation between AFP and imaging response by WHO and EASL guidelines at 3 and 6 months. The highest Kappa agreement (κ=0.83) and Pearson correlation (r=0.84) were seen between AFP and EASL response at 3 months. The sensitivity, specificity, positive and negative predictive values of AFP in predicting EASL response at 3 months were 96.6%, 85.7%, 92.3% and 93.3% (Figure 2). Correlation between AFP and EASL at 6 months was moderate (κ=0.59, r=0.59). Correlation between AFP and WHO response was low. The sensitivity of AFP at detecting EASL and WHO response was moderate-high; however, specificity fluctuated between time points (high for EASL at 3 months and moderate at 6 months, low for WHO).

Pattern of AFP reduction

The pattern of AFP reduction is summarized in Supplementary Tables 2 and 3. Median AFP level at baseline, 3 and 6 months was 1360, 145 and 52 ng/mL, respectively. At 3 and 6 months, 47% and 74% of patients exhibited >90% reduction in AFP levels. The finding of >90% AFP reduction was relatively uniform when stratified by baseline AFP of 200-1000, 1000-3000 and >3000 ng/mL. Thus, there was no significant heterogeneity in AFP response when stratified by baseline AFP. (Supplementary Table 2).

Survival Analyses (Table 3)

i. Landmark Method—From the 3-month landmark, median survival for responders and nonresponders was: WHO [NC and 9.3 months (P=0.006)]; EASL [27.6 and 4.0 months (P=0.0001)] and AFP [27.6 and 2.6 months (P<0.0001)]. From the 6-month landmark, median survival for responders and nonresponders was: WHO [NC and 12.7 months (P=0.34)]; EASL [24.6 and 10.8 months (P=0.353)] and AFP [24.6 and 5.6 months (P=0.044)].

Exploratory survival analysis (Supplementary Table 4) at 3-month landmark revealed the following: 1) patients achieving quantitative AFP response (i.e. AFP<500 ng/mL) survived longer than nonresponders (29.4 and 6.8 months, P=0.003); 2) survival sub-analysis excluding patients with PVT also suggested that AFP responders (>50% reduction) survived longer than nonresponders (27.6 and 4.0 months, P=0.0003)

ii. Risk-of-Death—From the 3-month landmark, the death rate in responders and nonresponders was: WHO [0% and 32% (P=0.06)]; EASL [7% and 57% (P=0.001)] and AFP [7% and 61% (P=0.0004)]. From the 6-month landmark, the death rate in responders and nonresponders was: WHO [20% and 23% (P=0.73)], EASL [17% and 40% (P=0.612)] and AFP [16% and 51% (P=0.4)].

iii. Mantel-Byar Method—By Mantel-Byar, survival outcomes favored EASL (P=0.002) and AFP response (P=0.0009) over WHO (P=0.177), confirming the association of EASL and AFP response with survival.

Analyses i, ii, and iii when repeated excluding transplanted patients demonstrated consistent results (Table 3).

Uni/Multivariate analyses (Table 4)

At the 3-month landmark, univariate analysis confirmed the following as independent prognosticators of survival: WHO response (P=0.006, HR:0.10, CI:0.04-0.27), EASL response (P=0.0001, HR:0.18, CI:0.06-0.50), and AFP response (P<0.0001, HR:0.16, CI: 0.05-0.52); multivariate analysis confirmed only AFP response (P=0.03, HR:0.14, CI: 0.02-0.83) and bilirubin ≤1.2 mg/dL (P=0.01, HR: 0.2, CI:0.06-0.72) as independent prognosticators of survival. Since no variable attained significance on univariate analysis, no multivariate analysis was performed at the 6-month landmark.

Effect of tumor size/liver function on response/survival

Supplementary Table 5 illustrates that responders and nonresponders were comparable by their baseline tumor size and liver function at each landmark. Baseline tumor size did not affect survival on univariate analysis (Table 4).

DISCUSSION

HCC patients usually present beyond potentially curative options.[42] In this scenario, systemic agents and LRTs have an established palliative role.[3, 4, 7, 8, 12, 26] Consequently, response assessment following LRTs has also been extensively studied in order to develop appropriate guidelines for accurate response monitoring.[10, 26, 33]. AFP may play a potential role in this scenario, where, combined with imaging, it may improve the ability to assess treatment response and consequently, directly impact clinical care and future therapy.[10]

Radiological response has been established to correlate with pathological response (gold standard).[23, 43, 44] The purpose of this study was to analyze if AFP response correlates with imaging response. If a correlation can be demonstrated, the potential advantages of AFP assessment in patient follow-up become apparent: 1) AFP response is a test that may reduce the cost burden of repeat imaging scans; 2) high-quality imaging scans are not readily available in developing countries (where HCC is potentially the most relevant health crisis), limiting its universal role; 3) difficulties persist in assessing response and progression in cirrhotic livers (infiltrative tumors, dyplastic nodules vs early HCC); 4) patients responding by one guideline (e.g. EASL) may not respond by another (e.g. WHO) and, 5) controversies persist in the optimal response assessment tool. Therefore, it is critical to establish response algorithms that incorporate multiple variables and parallel clinical practice, including AFP.

AFP-imaging response correlation was performed at multiple time-points by 3 separate statistical methods. Our findings suggest that AFP has a strong correlation with EASL response at 3 months and maintains a moderate correlation at 6 months. Translated clinically, AFP responders have a high likelihood of exhibiting EASL response; AFP nonresponders don't exhibit EASL response. The sensitivity of AFP in detecting imaging response by both EASL and WHO is moderate to high; when there is imaging response, AFP is likely to detect it. The specificity for detecting the absence of EASL response at 3-months (85.7) and moderate at 6-months (60); however, specificity was low for detecting the absence of WHO response, where size (rather than necrosis) is taken into consideration. This leads to the question of whether achieving imaging response is necessary, or if AFP response is sufficient from an overall survival standpoint. Studies have shown that patients may experience symptomatic improvement and pathological remission despite the absence of imaging response.[45] This question was investigated by our AFP response-survival analysis.

AFP response was shown to be a strong predictor of survival outcomes, with better consistency than WHO or EASL response. As demonstrated by 3 statistical tools, AFP responders seem to survive much longer than nonresponders; this was consistent after excluding transplanted and PVT patients. Supplementary Table 4 also demonstrates that not only the percentage reduction, a quantitative reduction in AFP (i.e. from >500 ng/mL to <500 ng/mL) is also a prognosticator of better survival outcomes. Although a 500 ng/mL cut-off was chosen arbitrarily, it has been reported that pre-transplant AFP level <500 ng/mL leads to lower transplant dropout compared to >500 ng/mL.[46]

Multivariate analysis reinforced this concept where AFP response was noted to predict longer survival independent of EASL and WHO response. Longer survival for AFP responders was also found to be independent of baseline tumor size and liver function (Supplementary Table 5). Potentially, AFP responders survive longer than nonresponders since AFP is more reflective of overall subclinical disease than imaging. At 6 months, AFP-imaging correlation was lower than 3 months. Moreover, although survival for responders at

6 months landmark is longer than nonresponders, it was not found to be significant. This is likely explained by extrahepatic progression/multifocality that developed at the 6 month landmark confounding any correlation, suggesting that AFP-imaging-survival correlations are best studied at the 3-month landmark.

The median times-to-response were: WHO: 5.9 months; EASL; 1.2 months and AFP: 1.2 months. This shows that patients exhibit AFP/EASL response earlier than WHO, supporting their role in early response assessment.

This analysis studied AFP-imaging response in a time-dependent fashion, with 2 time-points and 3 robust methodologies. The 2 landmarks predicted survival outcomes after exclusion of patients with aggressive tumor biology and underlying liver disease. On the other hand, the Mantel-Byar method included all patients from day 0 and calculated the risk-of-death at multiple time-points based on respective response status, confirming longer survival for EASL and AFP responders. WHO response did not demonstrate survival benefit by Mantel-Byar, potentially explained by the lower number of patients reaching WHO response endpoints (compared with EASL). Alternatively, these results may imply that necrosis and AFP reduction is a superior indicator of tumor response than size decrease, and indirect measure of the regenerative capacity of cirrhotic livers. Finally, we observed that the magnitude of AFP reduction was relatively uniform when stratified by baseline AFP.

Strengths/Limitations

There are strengths to this analysis. First, this is a novel study where AFP and imaging response were correlated at multiple time-points in a time-dependent manner. Second, Landmark/Mantel-Byer methods corrected for responder versus nonresponder guarantee-time bias, enabling a biologic “test-of-time” minimizing unknown confounders. Multiple statistical methods converging to the same conclusion (as in this study) lend strength to the conclusions presented.[47] Third, the statistical methods included adjusted P-values; conclusions were drawn following adjustment, permitting cautious interpretation.[40, 41] Finally, imaging was recognized as the gold standard; AFP correlative analyses were based on imaging as the reference standard.[36] There are limitations. First, the limited size of the study is recognized. Highly conservative selection criteria were deemed essential to minimize variables that would confound AFP levels (multiple lesions, extrahepatic metastases) and permit AFP-producing single lesion imaging-AFP-survival analyses. Second, given the absence of accepted definitions of AFP progression, this could not be incorporated in our analysis. Third, the small sample size prevented TACE/^{90}Y subset analyses; these findings may be equally applicable to TACE or ^{90}Y treated solitary HCCs. Fourth, this study only establishes individual (not trial) level association; since trial level association would require RCTs to establish AFP response as a surrogate of the true endpoint (survival), these findings should be considered hypothesis-generating. Fifth, only 30% of patients are AFP producers, and these levels may fluctuate because of underlying liver disease; this may impact AFP response assessment. Sixth, we do recognize that as time goes on, solitary HCCs progress and AFP becomes confounded by multifocality/extrahepatic disease. However, our study does suggest that AFP and imaging do correlate strongly, most evident at the 3-month landmark. Finally, we cannot conclude that AFP response directly causes longer survival. Rather, AFP response as a biomarker likely identifies patients with unknown characteristics that favor longer survival.[24, 25]

CONCLUSION

This study investigates AFP response in a time-dependent fashion. AFP response assessment is simple, reproducible, operator independent and is highly sensitive for detecting radiologic response. Response by AFP and EASL predicts improved survival. Consideration should be

made to develop HCC treatments that not only prolong TTP, but also elicit AFP and tumor response.[22] Future research should focus on incorporating AFP in response assessment methodologies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Role of Funding: There was no funding provided for this study. RS and RAO are supported in part by NIH grant CA126809.

Abbreviations

HCC	Hepatocellular Carcinoma
LRT	Locoregional therapy
CT	triphasic contrast-enhanced computerized tomography
MRI	magnetic resonance imaging
AFP	Alpha-fetoprotein
WHO	World Health Organization
EASL	European Association for the Study of the Liver
TACE	Transarterial Chemoembolization
^{90}Y	Yttrium-90 radioembolization
CP	Child-Pugh
UNOS	United Network for Organ Sharing
BCLC	Barcelona Clinic Liver Cancer
RFA	Radiofrequency ablation
AASLD	American Association for the Study of Liver Diseases
CR	Complete Response
PR	Partial Response
SD	Stable Disease
PD	Progressive Disease
RECIST	Response Evaluation Criteria in Solid Tumors
PVT	Portal venous thrombosis
mRECIST	modified Response Evaluation Criteria in Solid Tumors
NPV	negative predictive value
PPV	positive predictive value
NC	Not calculable
HR	Hazard Ratio
CI	95% Confidence Interval
RCTs	randomized control trials

TTP time-to-progression

REFERENCES

1. El-Serag HB. Hepatocellular carcinoma and hepatitis C in the United States. Hepatology. 2002; 36:S74–83. [PubMed: 12407579]
2. Alterkruse SF MK, Reichman ME. Hepatocellular carcinoma incidence, mortality, and survival benefits in the United States From 1975 to 2005. J Clin Oncol. 2009; 27:1485–1491. [PubMed: 19224838]
3. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, et al. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med. 2008; 359:378–390. [PubMed: 18650514]
4. Cheng AL, Kang YK, Chen Z, Tsao CJ, Qin S, Kim JS, et al. Efficacy and safety of sorafenib in patients in the Asia-Pacific region with advanced hepatocellular carcinoma: a phase III randomised, double-blind, placebo-controlled trial. Lancet Oncol. 2009; 10:25–34. [PubMed: 19095497]
5. Maddala YK, Stadheim L, Andrews JC, Burgart LJ, Rosen CB, Kremers WK, et al. Drop-out rates of patients with hepatocellular cancer listed for liver transplantation: outcome with chemoembolization. Liver Transpl. 2004; 10:449–455. [PubMed: 15004776]
6. Kulik LM, Carr BI, Mulcahy MF, Lewandowski RJ, Atassi B, Ryu RK, et al. Safety and efficacy of (90)Y radiotherapy for hepatocellular carcinoma with and without portal vein thrombosis. Hepatology. 2007; 47:71–81. [PubMed: 18027884]
7. Salem R, Lewandowski RJ, Mulcahy MF, Riaz A, Ryu RK, Ibrahim S, et al. Radioembolization for hepatocellular carcinoma using Yttrium-90 microspheres: a comprehensive report of long-term outcomes. Gastroenterology. 2010; 138:52–64. [PubMed: 19766639]
8. Lewandowski RJ, Mulcahy MF, Kulik LM, Riaz A, Ryu RK, Baker TB, et al. Chemoembolization for hepatocellular carcinoma: comprehensive imaging and survival analysis in a 172-patient cohort. Radiology. 2010; 255:955–965. [PubMed: 20501733]
9. Lewandowski RJ, Kulik LM, Riaz A, Senthilnathan S, Mulcahy MF, Ryu RK, et al. A comparative analysis of transarterial downstaging for hepatocellular carcinoma: chemoembolization versus radioembolization. Am J Transplant. 2009; 9:1920–1928. [PubMed: 19552767]
10. Riaz A, Miller FH, Kulik LM, Nikolaidis P, Yaghmai V, Lewandowski RJ, et al. Imaging response in the primary index lesion and clinical outcomes following transarterial locoregional therapy for hepatocellular carcinoma. JAMA. 2010; 303:1062–1069. [PubMed: 20233824]
11. Llovet JM, Real MI, Montana X, Planas R, Coll S, Aponte J, et al. Arterial embolisation or chemoembolisation versus symptomatic treatment in patients with unresectable hepatocellular carcinoma: a randomised controlled trial. Lancet. 2002; 359:1734–1739. [PubMed: 12049862]
12. Salem R, Lewandowski RJ, Kulik L, Wang E, Riaz A, Ryu RK, et al. Radioembolization Results in Longer Time-to-Progression and Reduced Toxicity Compared With Chemoembolization in Patients With Hepatocellular Carcinoma. Gastroenterology. 2011; 140:497–507. e492. [PubMed: 21044630]
13. Salem R, Lewandowski RJ, Atassi B, Gordon SC, Gates VL, Barakat O, et al. Treatment of unresectable hepatocellular carcinoma with use of 90Y microspheres (TheraSphere): safety, tumor response, and survival. J Vasc Interv Radiol. 2005; 16:1627–1639. [PubMed: 16371529]
14. Sangro B, Carpanese L, Cianni R, Golfieri R, Gasparini D, Ezziddin S, et al. Survival after yttrium-90 resin microsphere radioembolization of hepatocellular carcinoma across Barcelona clinic liver cancer stages: A European evaluation. Hepatology. 2011; 54:868–878. [PubMed: 21618574]
15. Trevisani F, De Notariis S, Rossi C, Bernardi M. Randomized control trials on chemoembolization for hepatocellular carcinoma: is there room for new studies? J Clin Gastroenterol. 2001; 32:383–389. [PubMed: 11319307]
16. Nomura F, Ohnishi K, Tanabe Y. Clinical features and prognosis of hepatocellular carcinoma with reference to serum alpha-fetoprotein levels. Analysis of 606 patients. Cancer. 1989; 64:1700–1707. [PubMed: 2477133]

17. Lok AS, Lai CL. alpha-Fetoprotein monitoring in Chinese patients with chronic hepatitis B virus infection: role in the early detection of hepatocellular carcinoma. Hepatology. 1989; 9:110–115. [PubMed: 2461890]
18. Riaz A, Ryu RK, Kulik LM, Mulcahy MF, Lewandowski RJ, Minocha J, et al. Alpha-fetoprotein response after locoregional therapy for hepatocellular carcinoma: oncologic marker of radiologic response, progression, and survival. J Clin Oncol. 2009; 27:5734–5742. [PubMed: 19805671]
19. Chan SL, Mo FK, Johnson PJ, Hui EP, Ma BB, Ho WM, et al. New utility of an old marker: serial alpha-fetoprotein measurement in predicting radiologic response and survival of patients with hepatocellular carcinoma undergoing systemic chemotherapy. J Clin Oncol. 2009; 27:446–452. [PubMed: 19064965]
20. Han K, Tzimas GN, Barkun JS, Metrakos P, Tchervenkov JL, Hilzenrat N, et al. Preoperative alpha-fetoprotein slope is predictive of hepatocellular carcinoma recurrence after liver transplantation. Can J Gastroenterol. 2007; 21:39–45. [PubMed: 17225881]
21. Tsai MC, Wang JH, Hung CH, Kee KM, Yen YH, Lee CM, et al. Favorable alpha-fetoprotein decrease as a prognostic surrogate in patients with hepatocellular carcinoma after radiofrequency ablation. J Gastroenterol Hepatol. 2010; 25:605–612. [PubMed: 20074164]
22. Memon K, Kulik L, Lewandowski RJ, Wang E, Riaz A, Ryu RK, et al. Radiographic Response to Locoregional Therapy in Hepatocellular Carcinoma Predicts Patient Survival Times. Gastroenterology. 2011; 141:526–535. e522. [PubMed: 21664356]
23. Riaz A, Memon K, Miller FH, Nikolaidis P, Kulik LM, Lewandowski RJ, et al. Role of the EASL, RECIST, and WHO response guidelines alone or in combination for hepatocellular carcinoma: Radiologic–pathologic correlation. Journal of Hepatology. 2011; 54:695–704. [PubMed: 21147504]
24. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. J Clin Oncol. 1983; 1:710–719. [PubMed: 6668489]
25. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response and other comparisons of time-to-event by outcome variables. J Clin Oncol. 2008; 26:3913–3915. [PubMed: 18711176]
26. Llovet JM, Di Bisceglie AM, Bruix J, Kramer BS, Lencioni R, Zhu AX, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. J Natl Cancer Inst. 2008; 100:698–711. [PubMed: 18477802]
27. Salem R, Thurston KG. Radioembolization with 90Yttrium Microspheres: A State-of-the-Art Brachytherapy Treatment for Primary and Secondary Liver Malignancies: Part 1: Technical and Methodologic Considerations. J Vasc Interv Radiol. 2006; 17:1251–1278. [PubMed: 16923973]
28. Salem R, Lewandowski RJ, Gates VL, Nutting CW, Murthy R, Rose SC, et al. Research reporting standards for radioembolization of hepatic malignancies. J Vasc Interv Radiol. 2011; 22:265–278. [PubMed: 21353979]
29. Bruix J, Sherman M. Management of hepatocellular carcinoma. Hepatology. 2005; 42:1208–1236. [PubMed: 16250051]
30. Bruix J, Sherman M, Llovet JM, Beaugrand M, Lencioni R, Burroughs AK, et al. Clinical management of hepatocellular carcinoma. Conclusions of the Barcelona-2000 EASL conference. European Association for the Study of the Liver. J Hepatol. 2001; 35:421–430. [PubMed: 11592607]
31. Forner A, Ayuso C, Varela M, Rimola J, Hessheimer AJ, de Lope CR, et al. Evaluation of tumor response after locoregional therapies in hepatocellular carcinoma: are response evaluation criteria in solid tumors reliable? Cancer. 2009; 115:616–623. [PubMed: 19117042]
32. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst. 2000; 92:205–216. [PubMed: 10655437]
33. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. Semin Liver Dis. 2010; 30:52–60. [PubMed: 20175033]
34. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–174. [PubMed: 843571]

35. Svanholm H, Starklint H, Gundersen HJ, Fabricius J, Barlebo H, Olsen S. Reproducibility of histomorphologic diagnoses with special reference to the kappa statistic. Apmis. 1989; 97:689–698. [PubMed: 2669853]
36. Sherman M. The resurrection of alphafetoprotein. J Hepatol. 2010; 52:939–940. [PubMed: 20395007]
37. Blade J, Lopez-Guillermo A, Bosch F, Cervantes F, Reverter JC, Montserrat E, et al. Impact of response to treatment on survival in multiple myeloma: results in a series of 243 patients. Br J Haematol. 1994; 88:117–121. [PubMed: 7803233]
38. Mantel N, Byar D. Evaluation of Response-Time Data Invloving Transient States: An Illustration Using Heart-Transplant Data. Journal of the American Statistical Association. 1974; 69:81–86.
39. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958; 53:457–481.
40. Gyorffy B, Gyorffy A, Tulassay Z. [The problem of multiple testing and solutions for genome-wide studies]. Orv Hetil. 2005; 146:559–563. [PubMed: 15853065]
41. Cleophas TJ, Zwinderman AH. Clinical trials are often false positive: a review of simple methods to control this problem. Curr Clin Pharmacol. 2006; 1:1–4. [PubMed: 18666373]
42. Mazzaferro V, Regalia E, Doci R, Andreola S, Pulvirenti A, Bozzetti F, et al. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. N Engl J Med. 1996; 334:693–699. [PubMed: 8594428]
43. Riaz A, Lewandowski RJ, Kulik L, Ryu RK, Mulcahy MF, Baker T, et al. Radiologic-pathologic correlation of hepatocellular carcinoma treated with chemoembolization. Cardiovasc Intervent Radiol. 2010; 33:1143–1152. [PubMed: 19967371]
44. Riaz A, Kulik L, Lewandowski RJ, Ryu RK, Giakoumis Spear G, Mulcahy MF, et al. Radiologic-pathologic correlation of hepatocellular carcinoma treated with internal radiation using yttrium-90 microspheres. Hepatology. 2009; 49:1185–1193. [PubMed: 19133645]
45. Leung TW, Patt YZ, Lau WY, Ho SK, Yu SC, Chan AT, et al. Complete pathological remission is possible with systemic combination chemotherapy for inoperable hepatocellular carcinoma. Clin Cancer Res. 1999; 5:1676–1681. [PubMed: 10430068]
46. Washburn K, Edwards E, Harper A, Freeman R. Hepatocellular carcinoma patients are advantaged in the current liver transplant allocation system. Am J Transplant. 2010; 10:1643–1648. [PubMed: 20486906]
47. Green E, Yothers G, Sargent DJ. Surrogate endpoint validation: statistical elegance versus clinical relevance. Stat Methods Med Res. 2008; 17:477–486. [PubMed: 18285438]

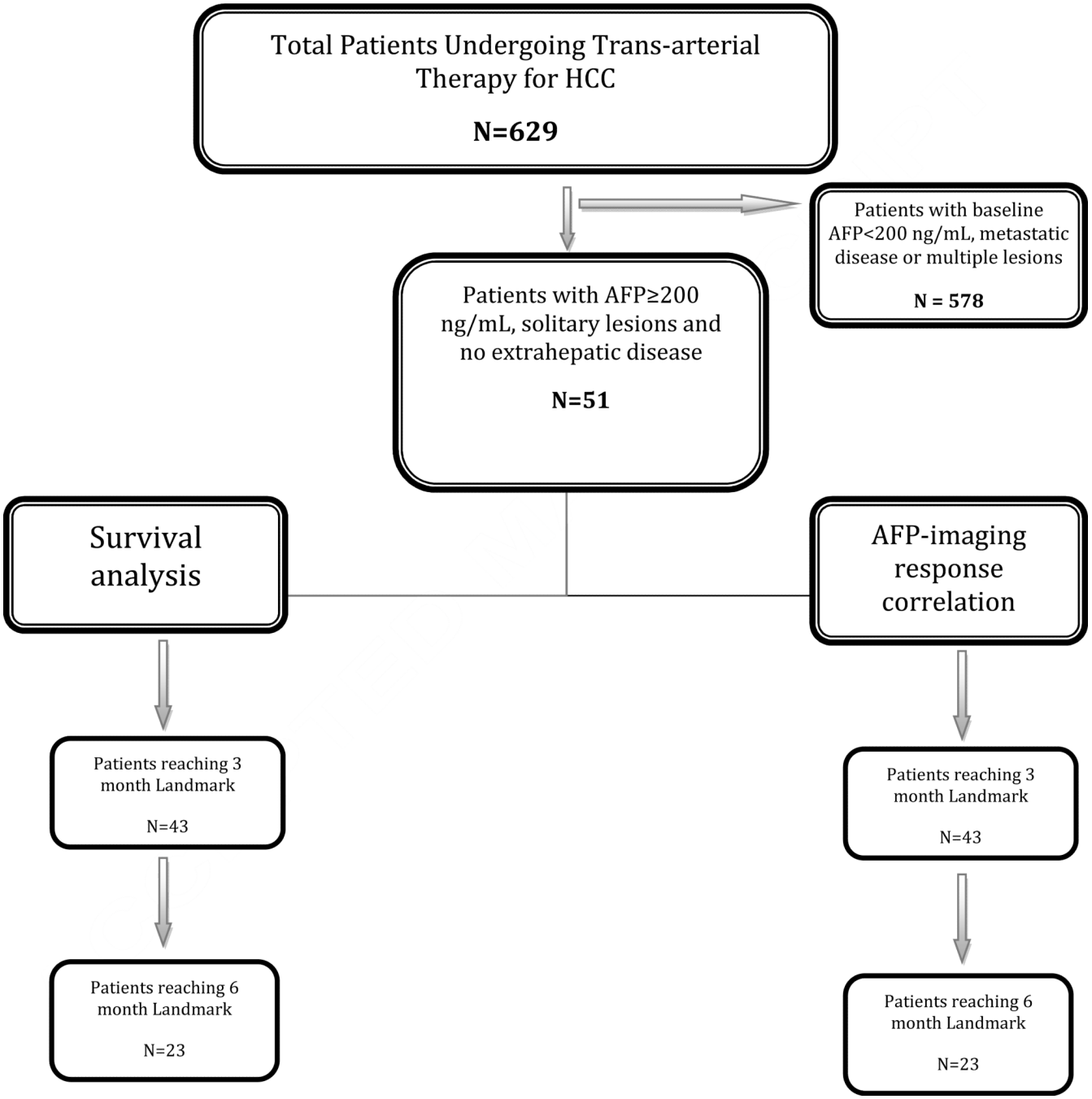

Figure 1.
Study flow chart

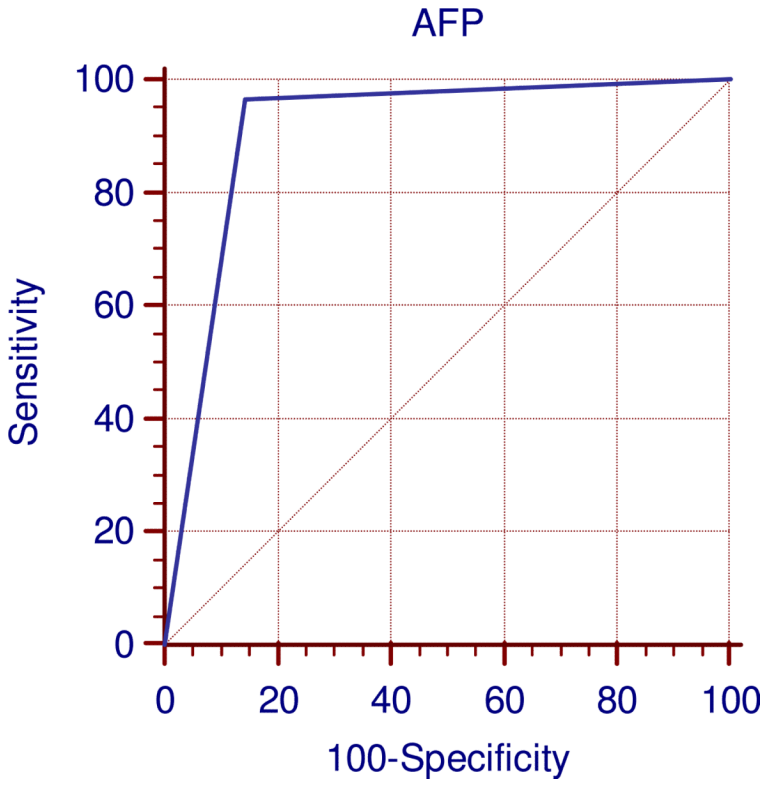

Figure 2.
ROC curve at 3 months between AFP and EASL response
Abbreviations: AFP: Alpha-fetoprotein; EASL: European Association for the Study of the Liver; ROC: Receiver Operating Characteristics

Table 1

Baseline patient characteristics

Characteristic	Category	Patients N=51 (%)	Characteristics	Category	Patients N=51 (%)
Treatment	^{90}Y Radioembolization	33 (65)	**Platelets**	Mean	163,000/μL
	Chemoembolization	18 (35)		Median	132,000/μL
				Range	39000 – 666000/μL
Age (years)	<65	23 (45)	**PVT**	Present	7 (14)
	≥65	28 (55)		Absent	44 (86)
Ethnic Group	Caucasian	29 (56)	**Lobar Distribution**	Unilobar	48 (94)
	Asian	7 (14)		Bilobar	3 (6)
	Hispanic	5 (10)	**Tumor Size**	≤5 cm	29 (57)
	African-American	10 (20)		5.1-10 cm	16 (31)
Gender	Male	30 (59)		>10 cm	6 (12)
	Female	21 (41)	**AFP (ng/mL)**	200-1000	21 (41)
Etiology	Alcohol	7 (14)		>1000	30 (59)
	Hepatitis C Virus (HCV)	20 (39)	**ECOG**	0	25 (49)
	HCV + Alcohol	2 (4)		1	24 (47)
	Hepatitis B Virus	4 (8)		2	2 (4)
	Nonalcoholic Steatohepatitis	1 (2)	**UNOS**	T1	1 (2)
	Cryptogenic	8 (15)		T2	25 (49)
	Autoimmune Hepatitis	1 (2)		T3	18 (35)
	Primary Biliary Cirrhosis	3 (6)		T4b	7 (14)
	Hemochromatosis	2 (4)	**BCLC**	A	23 (45)
	Unknown	3 (6)		B	10 (20)
Method of Diagnosis	Imaging	33 (65)		C	18 (35)
	Biopsy	18 (35)	**Child-Pugh**	A	29 (57)
Portal Hypertension	Present	37 (73)		B	22 (43)
	Absent	14 (27)	**Previous Therapy**	RFA	2 (4)
Ascites	Present	5 (10)		Resection	1 (2)
	Absent	46 (90)		Systemic Agents	0 (0)

Characteristic	Category	Patients N=51 (%)	Characteristics	Category	Patients N=51 (%)
Cirrhosis	Present	48 (94)		None	48 (94)
	Absent	3 (6)			

Abbreviations: AFP: Alpha-fetoprotein; BCLC: Barcelona Clinic Liver Cancer; ECOG: Eastern Cooperative Oncology Group; PVT: Portal Vein Thrombosis; RFA: radiofrequency ablation; UNOS: United Network for Organ Sharing

Table 2

Correlation and receiver operating characteristics between AFP and radiological tumor response by WHO and EASL guidelines

	3 months		6 months	
	WHO	EASL	WHO	EASL
N	43	43	23	23
Kappa agreement	0.20	0.83	0.11	0.59
Pearson Correlation	0.29	0.84	0.17	0.59
Sensitivity	91.7	96.6	90	94.4
Specificity	38.7	85.7	23.1	60
Negative Predictive Value	92.3	92.3	75	75
Positive Predictive Value	36.7	93.3	47.4	89.5

Abbreviations: AFP: Alpha-fetoprotein; EASL: European Association for the Study of the Liver; WHO: World Health Organization

Table 3

Survival analysis by response status

		WHO				EASL				AFP			
All Patients													
Landmark Method		**3 months**		**6 months**		**3 months**		**6 months**		**3 months**		**6 months**	
		R	**NR**	**R**	**NR**	**R**	**NR**	**R**	**NR**	**R**	**NR**	**R**	**NR**
	N	12	31	10	13	29	14	18	5	30	13	19	4
	***Survival**	NC	9.3	NC	12.7	27.6	4.0	24.6	10.9	27.6	2.6	24.6	5.6
	P value	**0.006**		0.34		**0.0001**		0.353		**<0.0001**		**0.044**	
	HR (CI)	0.10 (0.04 – 0.27)		0.52 (0.15 – 1.83)		0.18 (0.06 – 0.50)		0.56 (0.13 – 2.25)		0.16 (0.05 to 0.52)		0.28 (0.03 to 2.07)	
Risk Of death In following 6 Months from Landmark		0/12 (0)	10/31 (32)	2/10 (20)	3/13 (23)	2/29 (7)	8/14 (57)	3/18 (17)	2/5 (40)	2/30 (7)	8/13 (61)	3/19 (16)	2/4 (50)
	Chi-Square	3.39		0.11		10.68		0.25		12.28		0.70	
	P value	0.06		0.73		**0.001**		0.612		**0.0004**		0.40	
****Mantel Byar Method**	**N**	51				51				51			
	D	24				24				24			
	A+	3				7				13			
	C++	21				17				11			
	Exp A	6.33				16.09				19.61			
	Chi-Square	1.81				14.04				10.98			
	Relative Odds	0.363				0.209				0.24			
	P value	0.177				**0.002**				**0.0009**			

Excluding Transplanted Patients													
Landmark Method		**3 months**		**6 months**		**3 months**		**6 months**		**3 months**		**6 months**	
		R	**NR**	**R**	**NR**	**R**	**NR**	**R**	**NR**	**R**	**NR**	**R**	**NR**
	N	7	24	8	9	18	13	13	4	20	11	14	3
	***Survival**	NC	9.2	NC	12.7	27.6	4.0	24.6	10.9	27.6	2.6	24.6	5.6
	P value	**0.008**		0.343		**0.0007**		0.376		**<0.0001**		0.052	
	HR (CI)	0.11 (0.04 – 0.29)		0.53 (0.15 – 1.83)		0.22 (0.08 – 0.58)		0.57 (0.14 – 2.29)		0.19 (0.06 – 0.56)		0.29 (0.04 – 2.09)	

Excluding Transplanted Patients													
Risk Of death In following 6 Months from Landmark		0/7 (0)	10/24 (42)	2/8 (25)	3/9 (33)	2/18 (11)	8/13 (61)	3/13 (23)	2/4 (50)	2/20 (10)	8/11 (72)	3/14 (21)	2/3 (66)
	Chi-Square	2.61		0.02		6.62		0.16		10.06		0.74	
	P value	0.106		0.875		**0.01**		0.684		**0.001**		0.388	
****Mantel Byar Method**	**N**	36				36				36			
	D	24				24				24			
	A^{+}	3				7				13			
	C^{++}	21				17				11			
	Exp A	6.20				13.50				19.65			
	Chi-Square	1.66				6.17				11.62			
	Relative Odds	0.37				0.32				0.23			
	P value	0.196				**0.013**				**0.0007**			

Abbreviations: AFP: Alpha-fetoprotein; CI: confidence interval; D: Total Deaths; EASL: European Association for the Study of the Liver; Exp A: expected deaths among responders; HR: Hazard ratio; R: responders; NC: not calculable; NR: nonresponders; N: Total Number; WHO: World Health Organization

* Survival is expressed as median (months) from the landmark. To determine actual overall survival, readers should sum the respective landmark time (3 or 6 months) to the survival times reported above.

** Mantel-Byar methodology includes all patients from baseline, not only those reaching the landmark; A^{+}, actual deaths among responders; C^{++}, actual deaths in nonresponders.

Table 4

Uni/Multivariate analysis

			UNIVARIATE (Kaplan-Meier and Logrank Test)			MULTIVARIATE (Cox Proportional Hazards Model)**	
Landmark	**Predictor**	**Category**	**Hazard Ratio (CI)**	**P-Value**	**Adjusted P Value***	**Hazard Ratio (CI)**	**P-value**
3-month	WHO	R	0.10 (0.04 - 0.27)	**0.006**	**0.042**	0.12 (0.01 – 1.26)	0.08
		NR	1.00			1.00	
	EASL	R	0.18 (0.06 – 0.50)	**0.0001**	**0.0007**	0.94 (0.16 – 5.54)	0.95
		NR	1.00			1.00	
	AFP	R	0.16 (0.05 to 0.52)	**<0.0001**	**<0.0007**	0.14 (0.02 – 0.83)	**0.03**
		NR	1.00			1.00	
	Baseline AFP	200-1000 ng/mL	0.75 (0.30 – 1.86)	0.537	-	-	-
		>1000 ng/mL	1.00			-	
	Albumin	≥3.5 g/dL	0.19 (0.06 – 0.58)	0.073	-	1.10 (0.10 –11.13)	0.93
		<3.5 g/dL	1.00			1.00	
	Bilirubin	≤1.2 mg/dL	0.56 (0.21 – 1.45)	0.182	-	0.20 (0.06 – 0.72)	**0.01**
		> 1.2 mg/dL	1.00			1.00	
	Baseline tumor Size	≤5 cm	0.76 (0.27 – 2.56)	0.659	-	-	-
		5.1 – 10 cm	0.53 (0.14 – 2.02)	0.354		-	
		>10 cm	1.00	0.639		-	-
	PVT	Absent	1.36 (0.40 – 4.58)	0.571			
		Present	1.00				
6-month	WHO	R	0.52 (0.15 – 1.83)	0.34	-	-	-
		NR	1.00			-	
	EASL	R	0.56 (0.13 – 2.25)	0.353	-	-	-
		NR	1.00			-	
	AFP	R	0.28 (0.03 – 2.07)	**0.04**	-		-
		NR	1.00			-	
	Baseline AFP	200-1000 ng/mL	0.85 (0.23 – 3.19)	0.81	-	-	-
		>1000 ng/mL	1.00			-	

			UNIVARIATE (Kaplan-Meier and Logrank Test)			MULTIVARIATE (Cox Proportional Hazards Model)**	
Landmark	**Predictor**	**Category**	**Hazard Ratio (CI)**	**P-Value**	**Adjusted P Value***	**Hazard Ratio (CI)**	**P-value**
	Albumin	≥3.5 g/dL	0.55 (0.13 – 2.19)	0.449	-	-	-
		<3.5 g/dL	1.00			-	
	Bilirubin	≤ 1.2 mg/dL	0.64 (0.17 – 2.41)	0.53	-	-	-
		> 1.2 mg/dL	1.00			-	
	Baseline tumor Size	≤5 cm	0.59 (0.09 – 3.62)	0.570	-	-	-
		5.1 – 10 cm	0.88 (0.15 – 5.01)	0.886	-	-	-
		>10 cm	1.00	0.808	-	-	-
	PVT	Absent	0.95 (0.31 – 2.79)	0.907			
		Present	1.00				

Abbreviations: AFP: Alpha-fetoprotein; CI, confidence interval; EASL, European Association for the Study of the Liver; HR, hazard ratio; PVT: Portal Vein Thrombosis; UNOS: United Network for Organ Sharing; WHO: World Health Organization

*Adjusted for multiple comparisons using Bonferroni methodology (correction factor n=7).

**Factors were included in multivariate analysis if $P < 0.25$ in univariate analysis (unadjusted for multiple comparisons).