Published in final edited form as:

Gastroenterology. 2011 August ; 141(2): 526–535.e2. doi:10.1053/j.gastro.2011.04.054.

Radiographic Response to Locoregional Therapy in Hepatocellular Carcinoma Predicts Patient Survival Times

Khairuddin Memon¹, Laura Kulik², Robert J Lewandowski¹, Edward Wang³, Ahsun Riaz¹, Robert K Ryu¹, Kent T Sato¹, Karen Marshall¹, Ramona Gupta¹, Paul Nikolaidis¹, Frank H Miller¹, Vahid Yaghamai¹, Seanthan Senthilnathan¹, Talia Baker³, Vanessa L Gates¹, Michael Abecassis³, Al B Benson III⁴, Mary F Mulcahy⁴, Reed A Omary¹, and Riad Salem^{1,3,4}

¹Department of Radiology, Section of Interventional Radiology, Northwestern Memorial Hospital, Robert H. Lurie Comprehensive Cancer Center, Chicago IL

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

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

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

Abstract

Background & Aims—It is not clear whether survival times of patients with hepatocellular carcinoma (HCC) are associated with their response to therapy. We analyzed the association between tumor response and survival times of patients with HCC who were treated with locoregional therapies (LRTs; chemoembolization and radioembolization).

Methods—Patients received LRTs over a 9-year period (n=463). Patients with metastases, portal venous thrombosis, or who had received transplants were excluded; 159 patients with Child-Pugh ≤ B7 were analyzed. Response (based on European Association for Study of the Liver [EASL] or World Health Organization [WHO] criteria) was associated with survival times using the landmark, risk-of-death, and Mantel-Byar methodologies. In a subanalysis, survival times of responders were compared to those of patients with stable disease (SD) and progressive disease (PD).

Results—Based on 6-month data, in landmark analysis, responders survived longer than nonresponders (based on EASL but not WHO criteria: P=0.002 and 0.0694). The risk of death was

© 2011 The American Gastroenterological Association. Published by Elsevier Inc. All rights reserved.

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, salem@northwestern.edu.

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

Author contributions:

- study concept/design: KM, RJL, AR, RS
- acquisition of data: KM, LK, RJL, AR, RKR, KTS, RG, PN, KM, RS
- analysis/interpretation of data: KM, LK, EW, AR, ABB, MFM, RS
- drafting of the manuscript-KM, LK, RS
- critical revision of the manuscript-all authors
- statistical analysis: KM, EW, AR, RS
- study supervision: KM, LK, EW, AR, SS, RAO, MA, ABB, MFM, RS

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

also lower for responders (based on EASL but not WHO criteria: $P=0.0463$ and 0.707). Landmark analysis of 12-month data showed that responders survived longer than nonresponders ($P<0.0001$ and 0.004 , based on EASL and WHO criteria, respectively). The risk of death was lower for responders ($P=0.0132$ and 0.010 , based on EASL and WHO criteria, respectively). By the Mantel-Byar method, responders had longer survival than nonresponders, based on EASL criteria ($P<0.0001$; $P=0.596$ with WHO criteria). In the subanalysis, responders lived longer than patients with SD or PD.

Conclusion—Radiographic response to LRTs predicts survival time. EASL criteria for response more consistently predicted survival times than WHO criteria. The goal of LRT should be to achieve a radiologic response, rather than to stabilize disease.

Keywords

embolization; liver cancer; radiology; treatment

Introduction

The incidence of HCC is increasing;¹ it has tripled in the United States from 1975-2005.² HCC is the 6th most common malignancy worldwide and is the third most common cause of cancer-related mortality.^{3,4} While surgical treatments (transplantation/resection) provide the best curative outcomes, many patients present at an advanced stage. Sorafenib has been shown to prolong survival in advanced HCC.^{5,6} LRTs have an important role in the management of HCC by inducing tumor necrosis, delaying progression and improving survival.⁷⁻¹⁶

One of the most important objectives of cancer treatment is to improve survival. Variables that can be used as surrogates of survival are of interest; these are of paramount importance in designing clinical trials. One such surrogate is radiographic tumor response. This straightforward endpoint is potentially the pinnacle upon which clinical decisions are made on an individual basis; patients and clinicians want to know if the tumor(s) has(ve) responded to treatment by exhibiting a decrease in size or demonstrating necrosis. Radiological tumor response by WHO, RECIST and EASL guidelines have already been established to correlate well with actual pathological necrosis.¹⁷⁻¹⁹ Unfortunately, there is a dearth of studies correlating response with survival in aggregate, particularly in HCC. Sala et al performed a study correlating initial 1-month CR to survival in patients treated with ablation.²⁰ There are no such studies using extended Landmarks reported in the setting of LRTs.

HCC, unlike other cancers, develops in the setting of cirrhosis. Liver function therefore poses a substantial confounding effect in predicting survival since patients can succumb to either cancer or cirrhosis-related complications. It is therefore of interest to investigate if tumor response can prognosticate survival irrespective of liver function. Is it intuitive that patients who respond to LRTs live longer than those who do not? If demonstrated to be true, this would be an impetus for future HCC clinical trial design to focus not only on disease control (prolonging TTP) but also tumor response.²¹

Previous investigators have commented on the statistical weakness of responder/non-responder survival analyses.²² They proposed potentially more robust alternate statistical methods that could be used to study the effect of response on survival. As a result of these recommendations, we sought to investigate this concept. The purpose of our study was to correlate radiographic response and survival in the setting of LRTs.^{10,23} We also analyzed patients with disease control (CR+PR+SD) vs. PD in order to investigate if achieving disease control translated into clinical benefit, or if there was improved survival only when

exhibiting radiographic response (CR+PR). We performed comprehensive time-dependent response analyses using Landmark, risk-of-death and Mantel-Byar methodologies.

Methods

This study was compliant with the Health Insurance Portability and Accountability Act and approved by the Northwestern University Institutional Review Board. From 2000-2008, 463 HCC patients were treated at our institution with transarterial LRTs (TACE or ^{90}Y); this source population has been previously reported.⁷ Patients were eligible for transarterial therapies if they exhibited unresectable HCC (determined by transplant surgery) and bilirubin <3.0 mg/dL. To create the study population for this analysis, we excluded patients who: a) underwent transplantation (curative therapy), b) exhibited PVT or extrahepatic metastases at baseline (confounding variables) or, c) were CP>B7 (commonly used liver function cut-off for HCC studies) (Figure 1 flow chart). Patients with PVT and/or extrahepatic metastases were excluded in order to: a) create a homogeneous cohort with the fewest confounding variables, b) minimize the number of patients with negative tumor biology and, c) maximize the number of patients that would reach the 6 and 12-month Landmarks. Therefore, our study population was comprised of 159 patients whose survival outcomes were analyzed with respect to their radiological response to LRTs. 127 had died at the time of data closure. One patient subsequently received sorafenib; survival was censored at drug initiation. Two other patients received sorafenib for <4 weeks. None of the other 156 patients received sorafenib.

Evaluation/Staging

All patients underwent pretreatment assessment consisting of history, laboratory and imaging work-up. Patients were categorized as cirrhotics in the presence of nodular liver surface, varices, splenomegaly and/or thrombocytopenia. Diagnostic criteria for HCC were defined by guidelines.²¹ Baseline staging was performed by CP (liver function), UNOS TNM (tumor size/number) and BCLC classification systems (liver function, tumor size/number, cancer symptoms) (Supplementary Table 1).²¹ Patients were classified as BCLC C only if they exhibited HCC-related symptoms (e.g. tumor-related capsular or referred pain, cancer-related symptoms, weight loss).⁷

Locoregional Therapies

Chemoembolization was performed using 30 mg mitomycin, 30 mg adriamycin and 100 mg cisplatin (with lipiodol), followed by embolization using 300-500 micron particles per previous description.¹¹ Radioembolization was performed using glass microspheres per previous methodology.^{24,25}

Timing of Imaging/Laboratory Evaluations

Patients were followed using CT or MRI 1-month following treatment and at scheduled 2-3 month intervals. Liver function testing coincided with the imaging scans.

Imaging Technique

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. The technique for MR and CT is described below:

MR Imaging—Abdomen/pelvis MRI protocol for liver imaging included transverse and coronal T2-weighted half-Fourier acquisition single-shot turbo spin echo, T2-weighted turbo spin-echo with fat suppression. Unenhanced and dynamic gadolinium-enhanced T1-

weighted images were acquired by fat suppressed gradient echo using shared prepulses (arterial/venous phases). Gadopentetate dimeglumine (Magnevist; Bayer HealthCare Pharmaceuticals, Berlin, Germany) was administered at a dose of 0.1 mmol/kg, followed by 20 ml saline flush (2 ml/sec) with a power injector (Spectris; Medrad Warrendale, PA).

CT Imaging—Abdomen/pelvis CT for liver imaging included unenhanced, arterial and portal venous phases according to our standard HCC protocol. Contrast-enhanced images were obtained after 40 seconds in the arterial phase and 70 seconds in the venous phase after injection of 125 mL of iohexol (Omnipaque 350, GE Healthcare, Waukesha, WI) at a rate of 5 ml/s.

Methodology of Response Assessment

Response status was assessed using WHO and EASL necrosis criteria (Supplementary Table 2) using the index (biomarker) lesion concept.^{5,13,26-28} Patients with CR or PR were categorized as responders; those with SD or PD were categorized as nonresponders. SD was designated a non-responder because without exhibiting necrosis or a decrease in size, we could not be certain that a lesion was not stable simply because of natural history in a slowly growing tumor; this was deemed conservative. Patients developing new lesions were categorized as PD. New lesions were diagnosed by following accepted guidelines (AASLD, EASL). For suspicious new HCC lesions where imaging was not classic for HCC, a biopsy was performed.

The initial response assessment was performed by three imaging radiologists as a part of routine clinical care with reporting of official measurements entered into the medical records (PN, FHM, VY). In order to obtain a second imaging response assessment, the treating interventional radiologists (R JL, RKR, KTS, RAO, RS) re-reviewed all 159 patients, tabulating size and tumor stage for all baseline and follow-up scans. In order to maintain an element of blinding of the interventional radiologists to their own patients and their long-term outcomes, the 159 patient cohort was divided alphabetically into 5 groups and correspondingly assigned to the 5 interventional radiologists. Although there were instances where the interpreting interventional radiologists were familiar with the patient, most commonly they were not.

WHO and EASL response guidelines were used to correlate response with survival. WHO (not RECIST) guidelines were used since a high inter-method correlation between WHO and RECIST has been demonstrated.¹³ Moreover, it is oncologic gold standard to report lesion size in 2 (not 1) dimensions. mRECIST was not used since the initiation of our study predates the introduction of this concept.²⁹

For the progression analysis, establishing the exact date of progression is difficult; theoretically, this can only be determined if scans are repeated daily. Therefore, in patients categorized as nonresponders from PD (as opposed to SD), we adopted a conservative methodology explained by an example: a patient responds on imaging at 3 months (responder), and later progresses at 6 months (becoming a nonresponder). Since the patient could have theoretically progressed any-time after the 3-month scan, we chose to retrospectively adjudicate the nonresponder status (from PD) as of 3 months+1 day. This corrective measure ensured: a) conservative reporting, b) that we did not overestimate responders, c) a direct comparison of survival between true responders and nonresponders, and d) a complete elimination of variable scan time bias.

Statistical Analyses and Survival

Survival analysis—Previous investigators have concluded that the usual (ie. plotting survival of responders vs. nonresponders without the time variable) methods are incorrect, leading to biased estimates, invalid statistical tests and misleading conclusions.^{22,30} This is in part due to the fact that responders must live long enough for response to be observed (guarantee-time bias). This situation is especially relevant in HCC, as the effect of response on survival is confounded by coexisting mortality associated with underlying cirrhosis. Given that response is a time-dependent variable not present at baseline, survival based on response was assessed using three statistical methods correcting for guarantee-time bias.³¹ In order to investigate the confounding effect of liver function, the analyses were performed separately in the CP≤B7 and CP-A subsets.

i. Landmark Method: This method determines overall survival for patients based on their response status assessed at a specific time point after treatment. Six and 12-month Landmarks were used for this analysis; EASL and WHO response status were determined at each Landmark. Patients progressing at the Landmark were categorized as nonresponders (Supplementary Table 2). The Fisher's exact test was used to compare albumin and bilirubin levels at each Landmark between responders and nonresponders, thereby confirming that any survival difference would not be directly attributed to better liver function and that the groups were comparable at that point in time. Furthermore, in order to study the effects of disease control vs. progression, we compared survival outcomes of responders (CR+PR) vs. SD vs. PD. Survival curves were plotted by Kaplan-Meier technique and compared using the log-rank test.³²

ii. Risk-of-Death: This method compares the death rate of responders vs. non-responders in the 6 months following each Landmark. The chi-square test was used to compare death rates.

iii. Mantel-Byar Method:³³ Unlike other methods, this approach takes into consideration all patients from day 0; a cumulative risk-of-death is calculated from the individual risk-of-death determined at each time point where an event (death) occurs.

Uni/Multivariate analysis—Uni/multivariate analyses were performed using Cox proportional hazards model at each Landmark in order to capture the effects of different variables on survival, addressing whether survival was different based on response status, baseline tumor size or liver function (at the Landmark). Hazard ratio estimates were based on simultaneous analysis of all variables. Assumption of proportionality was tested using log-minus-log plot and was met. P-values <0.05 on univariate analyses were corrected for multiple comparisons using Bonferroni methodology.^{34,35} All analyses were conducted using SAS 9.2 (SAS, Cary, NC); P<.05 was considered significant.

Effect of tumor size on response and survival—Responders and nonresponders were stratified at the 6 and 12-month Landmarks by baseline tumor size to analyze the effect of tumor size on response rate and survival. The Fisher's exact test was used to compare tumor size between responders and nonresponders. Baseline tumor size was also added to the uni/multivariate analyses in order to explore its effect on survival.

Results

Baseline Characteristics

Table 1 summarizes the baseline characteristics of the 159 patients. 64 (40%) were <65 years old, 117 (74%) were male. The rates of portal hypertension, tumor distribution and cancer stage (UNOS, BCLC, Child-Pugh) are also summarized.

Survival Analyses (Table 2)

i. Landmark Method

A. Responder vs. Nonresponder: From the 6-month Landmark, median survival was 31.6 and 13.7 mo for responders and nonresponders by WHO ($P=0.0694$); for EASL it was 24.6 and 13.2 mo ($P=0.002$). From the 12-month Landmark, median survival for responders and nonresponders was 36.4 and 11.1 mo by WHO ($P=0.004$); for EASL it was 36.4 and 9.7 mo ($P<0.0001$).

B. Disease control (CR+PR+SD) vs. Progressive Disease: From the 6-month Landmark, median survival was 31.6, 21.0 and 10.6 mo for patients who were responders, SD and PD by WHO criteria ($P=0.0008$) and 24.6, 14.0 and 10.1 mo for patients who were responders, SD or PD by EASL criteria ($P=0.002$). From the 12-month Landmark, median survival was 36.4, 19.2 and 10.4 mo for patients who were responders, SD and PD by WHO criteria ($P=0.0004$) and 36.4, 11.2 and 9.7 mo for patients who were responders, SD or PD by EASL criteria ($P=0.0001$).

ii. Risk-of-Death—From the 6-month Landmark, the death rate in responders and nonresponders was 6.6% and 15.5% by WHO ($P=0.707$); it was 4.7% and 20.6% by EASL ($P=0.0463$). From the 12-month Landmark, the death rate in responders and nonresponders was 0% and 31.5% by WHO ($P=0.010$); it was 4% and 32.7% by EASL ($P=0.0132$).

iii. Mantel-Byar Method—By Mantel-Byar, survival outcomes favored EASL over WHO with $P<0.0001$ and 0.596, confirming EASL response was associated with improved survival.

Analyses i, ii, and iii on the CP-A patient subset demonstrated consistent results (Table 2).

Uni/Multivariate analyses (Table 3)

Uni/multivariate analyses confirmed the following as significant and independent prognosticators of survival at the 6-month [EASL response ($P=0.006$, HR:0.46, CI: 0.27-0.80), age ≥ 65 ($P=0.011$, HR:0.51, CI: 0.30-0.85), bilirubin ≤ 2 mg/dL ($P<0.0001$, HR: 0.13, CI:0.05-0.31)] and [EASL response ($P=0.002$, HR:0.20, CI:0.07-0.55), albumin ≥ 3.5 mg/dL ($P=0.04$, HR:0.47, CI:0.23-0.98)] at the 12-month Landmark.

Effect of tumor size on response and survival

Table 4 illustrates that responders and nonresponders were comparable by their baseline tumor size. Baseline tumor size was not a significant factor affecting survival on univariate analysis (Table 3).

Discussion

HCC is a condition where patients usually present beyond potentially curative options³⁶. In this scenario, systemic agents and LRTs have an established palliative role.^{5-7,10,11,21} Although extending patient survival is the ultimate therapeutic goal, demonstrating this

benefit has been difficult in HCC. In a disease commonly managed by multiple different therapies (resection, transplantation, ablation, LRT, systemic) with routine cross-over between treatments, identifying a survival benefit attributable to one therapy without prohibitively large studies becomes difficult. This scenario is clinically significant and unique to HCC, where survival is dependent on both the cancer and functional status of the cirrhotic liver (competing risks of death).²⁶ The underlying liver disease and its associated mortality are intricately intertwined with the treatment of HCC and can impact the endpoint of survival. The aim of this study was to investigate whether tumor response could be used as a prognosticator of survival. If so, one could design future RCTs validating radiographic response as a surrogate of survival.

Three statistical tools were used to investigate the concept of response prognosticating survival. First, the Landmark method determines patient response status at a specified time-point following treatment; survival estimates are calculated from that point forward. Patients who die before the Landmark are not included. This permits the analysis of survival in patients who have prognostically favorable biology (liver function/tumor standpoint), irrespective of their response status at the Landmark. Estimating the survival from the Landmark poses the question: “How does response to therapy impact survival?” This method decreases the probability of effects caused by unknown confounding variables on survival. Prior to analyzing the data, we selected the 6 and 12-month (+/- 1 month) Landmarks; these were deemed clinically appropriate and relevant.³⁰ Second, the risk-of-death method was used comparing death rates of responders vs. non-responders in the 6 months following each Landmark. Finally, we used the technique of Mantel-Byar. Unlike the Landmark method, this approach includes all patients from day 0; it treats response as a time-dependent covariate where all patients enter the study in ‘no-response’ state and with passage of time, some eventually respond and shift to ‘responder’ status; responders may ultimately progress and shift back to ‘non-responder’ state. At each successive death, the number of patients shifting from one response state to the other are estimated along with the pre-existing patients and the risk-of-death is estimated for each response category.³³ It removes the bias of the usual method and compares patients by their response at multiple periods of time in a dynamic fashion.

This analysis illustrates the potential utility of radiographic response, a prognosticator of survival, as a clinical trial endpoint. Survival studies can be prohibited by time/cost in early HCC, with the natural history of untreated disease >3 years.²¹ Additionally, survival can be influenced negatively or positively by crossover to alternative therapies at progression. The Landmark methodology was able to investigate survival outcomes. By selecting “healthier” HCC patients ($CP \leq B7$), the 6 and 12-month Landmarks minimized inclusion of those with poor prognostic factors (early death) due to aggressive tumor biology and/or underlying liver disease. Identification of “good biology” patients permits assessment of survival stratified by response status. Our data show that responders by WHO/EASL criteria live longer than nonresponders from the Landmark onwards, with the exception of WHO 6-month Landmark (near significance); the trend is clear. Both the Fisher's exact test and multivariate analysis adjusted for liver function (confounding variable) between responders and nonresponders (data not shown); these groups were not different at both Landmarks.

These results also illustrate the interplay between response, SD and PD. Responding patients exhibited longer survival than those who had stable disease, which in turn had better survival than those with PD. Therefore, although response and disease-control were associated with better survival, responders did outperform SD and PD. This suggests not only a correlation between response to therapy and survival, but also a hierarchy of survival benefit of response>SD>PD.

Further evidence for improved survival in responders comes from the risk-of-death analysis. Responders survived longer by both EASL and WHO. By WHO, survival did not differ significantly between responders and nonresponders at 6 months, however it did at 12 months. This may be related to the fact that the average WHO response in intra-arterial therapy often exceeds 6 months.^{7,12} EASL response at 6 and 12 months maintained a better overall survival in responders. Unlike WHO, EASL response occurs earlier post therapy; our data suggest that a lasting response at 12 months is a powerful predictor of patient outcome¹².

The Mantel-Byar method confirmed that EASL responders survived longer than nonresponders. However, WHO response did not demonstrate survival benefit by Mantel-Byar; this is potentially explained by the lower number of patients exhibiting WHO response (compared with EASL) and hence, potentially too few endpoints were observed. Alternatively, these results may imply that necrosis is a superior indicator of tumor response than size decrease; this may be related to the diminished regenerative capacity of cirrhotic livers.

Multivariate analyses revealed that EASL response was a better prognosticator at both Landmarks; WHO was not maintained at 6 months. At the 6-month Landmark, age ≥ 65 and bilirubin ≤ 2 mg/dL predict survival independently. Interestingly, both age and bilirubin became insignificant at 12-months, replaced by albumin ≥ 3.5 mg/dL. These findings were reproduced in the CP-A subset, further validating response impacting survival. Clinical trials have traditionally excluded CP>A patients in order to assess the efficacy of a therapy without the confounding risk of death from decompensated cirrhosis. Our data lend support for consideration of inclusion of CP-B7 patients in clinical trials. In order to investigate if response (and hence survival) varied according to treatment modality, this variable was included in uni/multivariate analysis. Patients treated with ^{90}Y appeared to exhibit longer survival than those treated with TACE from the 6-month Landmark; this observation was not maintained in the multivariate model.

One factor that requires further discussion includes the effect of tumor size on response and survival. Compared to nonresponders, it is possible that responders survived longer because of smaller tumor size. In order to investigate this possibility, responders and nonresponders at the 6 and 12-month Landmarks were stratified by baseline tumor size (Table 4). There was no difference between the tumor size of responders and nonresponders. On univariate analysis, baseline tumor size did not have a significant survival effect. Hence, the survival advantage noted in responders over nonresponders cannot be explained by differences in tumor size.

Ultimately, the relevance of this analysis is highlighted when challenging overall survival as the primary endpoint in HCC studies.³⁷ Despite always being quoted as the best endpoint in cancer studies, this concept potentially deserves reconsideration in HCC. First, the natural history of early untreated HCC is >3 years; designing survival studies in this setting results in prohibitively difficult trials to design and complete. This partially explains the lack of RCTs in early, resectable HCC.^{38,39} Second, the treatment algorithm for HCC is dynamic; survival is routinely confounded by multiple treatment crossovers, necessitating censoring at alternate therapies. Patients progressing on one protocol may be enrolled in another treatment protocol; it is unknown if survival benefit is incurred by the experimental treatment. Third, HCC patients have competing risks of death and can be salvaged by curative transplantation, significantly affecting survival. Furthermore, survival can be artificially prolonged by aggressive supportive measures (intensive care); this level of support is dependent on the availability of local resources. Finally, overall survival as an endpoint in isolation does not take into account the importance of quality-of-life; this is an

important discussion point as patients consider treatment options. Given these realities, the demonstration that tumor response prognosticates survival may permit more simplified clinical trial strategy, with the focus of new treatments not only being to delay progression, but also to elicit radiologic response. Although TTP is being advocated as a relevant endpoint, this may be difficult in single arm studies (no comparator) where the TTP is unknown. Hence, randomized phase 2 trials have been advocated.^{21,40}

Strengths and Limitations

There are strengths to this analysis. The Landmark methodology enabled incorporation of the clinically relevant concept of biologic “test-of-time”, theoretically minimizing uncontrolled/unknown confounders. We applied two predetermined Landmarks, thereby permitting an “early” (6-month) and “late” (12-month) biologic test-of-time; a shorter Landmark was not deemed clinically relevant. These methods corrected for responder versus nonresponder analyses plagued by guarantee-time bias. The statistical methods permitted the assessment of response as a time-dependent covariate. Only one patient received sorafenib for >4 weeks; survival was censored for that patient. One possible explanation for the lack of sorafenib use in this cohort may be that when patients progressed following LRTs, they were not candidates (elevated liver function, clinical decompensation, advanced performance status) to receive the agent; this decision was made by medical oncology, hepatology and interventional radiology. Also, this study was closed in 2008; sorafenib had not gained widespread clinical adoption at that time. Institutionally, as part of our research mission, patients with preserved liver functions and progressive disease were considered for other clinical trials. Since this drug has proven survival benefit, the exclusion of sorafenib in this study prevented a potential confounding effect on survival outcomes. As previously reported, multiple statistical methods converging to the same conclusion (as in this study) lend strength to the conclusions presented.⁴¹ Although controversial, the statistical methods included adjusted P-values and conclusions were drawn following adjustment, permitting cautious interpretation.^{34,35} There are limitations to this analysis. Nonresponders from PD were categorized as having progressed on the previous scan date+1 day. Since a certain amount of time is necessary to exhibit response, this concept may not be applicable in advance disease (survival <12 months).⁵ This study only establishes individual level association; for the purpose of trial level association validating response as a surrogate of survival, multicenter RCTs will be required. Finally, it should be considered that no matter how thorough the analysis, we cannot conclude that tumor response directly causes longer survival. Responders may survive longer than the nonresponders not because of the direct effect of response, but because response identifies patients with unknown characteristics that favor longer survival.^{22,30}

Conclusion

This is the first study that investigates the prognostic ability of tumor response in HCC patients following LRTs by using novel statistical tools. Compared to nonresponders, survival in responders was longer by 3 statistical methods; this difference could not be attributed to differences in synthetic liver function, tumor size or treatment received. Tumor response does prognosticate survival and is better than achieving SD. It also validates the concept proposed by EASL for assessment of tumor response; EASL response was significant by all statistical methodologies applied²⁶. Based on these findings, consideration should be made to develop treatments for HCC that not only prolong TTP, but also elicit radiographic tumor response. RCTs will be required to validate this concept and establish radiographic response as a surrogate of the true endpoint (survival).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Sharon Coffey, Krystina Salzig, Jenny Karp, Elizabeth Gonda and Margaret Gilbertsen for their compassionate care of our patients. We would also like to acknowledge the efforts of Saad M Ibrahim MD, Bassel Atassi MD and Steven Newman MD.

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

References

1. El-Serag HB. Hepatocellular carcinoma and hepatitis C in the United States. *Hepatology*. Nov; 2002 36(5 Suppl 1):S74–83. [PubMed: 12407579]
2. Alterkruse SF, M K, 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. Parkin DM, B F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin*. 2005; 55:74–108. [PubMed: 15761078]
4. Bosch FX, Ribes J, Diaz M, Cleries R. Primary liver cancer: worldwide incidence and trends. *Gastroenterology*. Nov; 2004 127(5 Suppl 1):S5–S16. [PubMed: 15508102]
5. Llovet JM, Ricci S, Mazzaferro V, et al. Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med*. Jul 24; 2008 359(4):378–390. [PubMed: 18650514]
6. Cheng AL, Kang YK, Chen Z, 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*. Jan; 2009 10(1):25–34. [PubMed: 19095497]
7. Salem R, Lewandowski RJ, Kulik L, et al. Radioembolization Results in Longer Time-to-Progression and Reduced Toxicity Compared With Chemoembolization in Patients With Hepatocellular Carcinoma. *Gastroenterology*. Feb; 2011 140(2):497–507. e492. [PubMed: 21044630]
8. Maddala YK, Stadheim L, Andrews JC, et al. Drop-out rates of patients with hepatocellular cancer listed for liver transplantation: outcome with chemoembolization. *Liver Transpl*. Mar; 2004 10(3):449–455. [PubMed: 15004776]
9. Kulik LM, Carr BI, Mulcahy MF, et al. Safety and efficacy of (90)Y radiotherapy for hepatocellular carcinoma with and without portal vein thrombosis. *Hepatology*. Nov 20; 2007 47(1):71–81. [PubMed: 18027884]
10. Salem R, Lewandowski RJ, Mulcahy MF, et al. Radioembolization for hepatocellular carcinoma using Yttrium-90 microspheres: a comprehensive report of long-term outcomes. *Gastroenterology*. Jan; 2010 138(1):52–64. [PubMed: 19766639]
11. Lewandowski RJ, Mulcahy MF, Kulik LM, et al. Chemoembolization for hepatocellular carcinoma: comprehensive imaging and survival analysis in a 172-patient cohort. *Radiology*. Jun; 2010 255(3):955–965. [PubMed: 20501733]
12. Lewandowski RJ, Kulik LM, Riaz A, et al. A comparative analysis of transarterial downstaging for hepatocellular carcinoma: chemoembolization versus radioembolization. *Am J Transplant*. Aug; 2009 9(8):1920–1928. [PubMed: 19552767]
13. Riaz A, Miller FH, Kulik LM, et al. Imaging response in the primary index lesion and clinical outcomes following transarterial locoregional therapy for hepatocellular carcinoma. *JAMA*. Mar 17; 2010 303(11):1062–1069. [PubMed: 20233824]
14. Llovet JM, Real MI, Montana X, et al. Arterial embolisation or chemoembolisation versus symptomatic treatment in patients with unresectable hepatocellular carcinoma: a randomised controlled trial. *Lancet*. May 18; 2002 359(9319):1734–1739. [PubMed: 12049862]

15. Lo CM, Ngan H, Tso WK, et al. Randomized controlled trial of transarterial lipiodol chemoembolization for unresectable hepatocellular carcinoma. *Hepatology*. May; 2002 35(5): 1164–1171. [PubMed: 11981766]
16. Salem R, Lewandowski RJ, Atassi B, et al. Treatment of unresectable hepatocellular carcinoma with use of 90Y microspheres (TheraSphere): safety, tumor response, and survival. *J Vasc Interv Radiol*. Dec; 2005 16(12):1627–1639. [PubMed: 16371529]
17. Riaz A, Lewandowski RJ, Kulik L, et al. Radiologic-pathologic correlation of hepatocellular carcinoma treated with chemoembolization. *Cardiovasc Intervent Radiol*. Dec; 2010 33(6):1143–1152. [PubMed: 19967371]
18. Riaz A, Kulik L, Lewandowski RJ, et al. Radiologic-pathologic correlation of hepatocellular carcinoma treated with internal radiation using yttrium-90 microspheres. *Hepatology*. Apr; 2009 49(4):1185–1193. [PubMed: 19133645]
19. Riaz A, Memon K, Miller FH, 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(4):695–704. [PubMed: 21147504]
20. Sala M, Llovet JM, Vilana R, et al. Initial response to percutaneous ablation predicts survival in patients with hepatocellular carcinoma. *Hepatology*. Dec; 2004 40(6):1352–1360. [PubMed: 15565564]
21. Llovet JM, Di Bisceglie AM, Bruix J, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. *J Natl Cancer Inst*. May 21; 2008 100(10):698–711. [PubMed: 18477802]
22. 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*. Aug 20; 2008 26(24):3913–3915. [PubMed: 18711176]
23. Llovet JM, Bruix J. Systematic review of randomized trials for unresectable hepatocellular carcinoma: Chemoembolization improves survival. *Hepatology*. Feb; 2003 37(2):429–442. [PubMed: 12540794]
24. 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*. Aug; 2006 17(8):1251–1278. [PubMed: 16923973]
25. Salem R, Lewandowski RJ, Gates VL, et al. Research reporting standards for radioembolization of hepatic malignancies. *J Vasc Interv Radiol*. Mar; 2011 22(3):265–278. [PubMed: 21353979]
26. Bruix J, Sherman M, Llovet JM, et al. Clinical management of hepatocellular carcinoma. Conclusions of the Barcelona-2000 EASL conference. European Association for the Study of the Liver. *J Hepatol*. Sep; 2001 35(3):421–430. [PubMed: 11592607]
27. Forner A, Ayuso C, Varela M, et al. Evaluation of tumor response after locoregional therapies in hepatocellular carcinoma: are response evaluation criteria in solid tumors reliable? *Cancer*. Feb 1; 2009 115(3):616–623. [PubMed: 19117042]
28. Therasse P, Arbuck SG, Eisenhauer EA, 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*. Feb 2; 2000 92(3):205–216. [PubMed: 10655437]
29. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. *Semin Liver Dis*. Feb; 2010 30(1):52–60. [PubMed: 20175033]
30. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. *J Clin Oncol*. Nov; 1983 1(11):710–719. [PubMed: 6668489]
31. Blade J, Lopez-Guillermo A, Bosch F, et al. Impact of response to treatment on survival in multiple myeloma: results in a series of 243 patients. *Br J Haematol*. Sep; 1994 88(1):117–121. [PubMed: 7803233]
32. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958; 53:457–481.
33. Mantel N, Byar D. Evaluation of Response-Time Data Involving Transient States: An Illustration Using Heart-Transplant Data. *Journal of the American Statistical Association*. 1974; 69(345):81–86.

34. Gyorffy B, Gyorffy A, Tulassay Z. The problem of multiple testing and solutions for genome-wide studies. *Orv Hetil.* Mar 20; 2005 146(12):559–563. [PubMed: 15853065]
35. Cleophas TJ, Zwiderman AH. Clinical trials are often false positive: a review of simple methods to control this problem. *Curr Clin Pharmacol.* Jan; 2006 1(1):1–4. [PubMed: 18666373]
36. Mazzaferro V, Regalia E, Doci R, et al. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. *N Engl J Med.* Mar 14; 1996 334(11):693–699. [PubMed: 8594428]
37. Sargent DJ, Hayes DF. Assessing the measure of a new drug: is survival the only thing that matters? *J Clin Oncol.* Apr 20; 2008 26(12):1922–1923. [PubMed: 18421044]
38. Okuda K, Ohtsuki T, Obata H, et al. Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients. *Cancer.* Aug 15; 1985 56(4):918–928. [PubMed: 2990661]
39. Llovet JM, Bustamante J, Castells A, et al. Natural history of untreated nonsurgical hepatocellular carcinoma: rationale for the design and evaluation of therapeutic trials. *Hepatology.* Jan; 1999 29(1):62–67. [PubMed: 9862851]
40. Gan H, Grothey A, Pond G, et al. Randomized phase II trials: inevitable or inadvisable? *J Clin Oncol.* 2010; 15:2641–2647. [PubMed: 20406933]
41. Green E, Yothers G, Sargent DJ. Surrogate endpoint validation: statistical elegance versus clinical relevance. *Stat Methods Med Res.* Oct; 2008 17(5):477–486. [PubMed: 18285438]

Abbreviations

AASLD	American Association for the Study of Liver Disease
BCLC	Barcelona Clinic Liver Cancer
CP	Child-Pugh
CR	Complete Response
CI	95% Confidence Interval
CT	triphasic contrast-enhanced computerized tomography
ECOG	Eastern Cooperative Oncology Group
EASL	European Association for the Study of the Liver
HR	Hazard Ratio
HCC	Hepatocellular Carcinoma
LRT	Locoregional therapy
mo	Months
MRI	gadolinium-enhanced magnetic resonance imaging
mRECIST	Modified Response Evaluation Criteria in Solid Tumors
PVT	Portal venous thrombosis
PD	Progressive Disease
PR	Partial Response
RCT	Randomized controlled trial
RECIST	Response Evaluation Criteria in Solid Tumors
SD	Stable Disease
TACE	Transarterial Chemoembolization

TTP	time-to-progression
UNOS	United Network for Organ Sharing
WHO	World Health Organization
⁹⁰Y	Yttrium-90 radioembolization

*Number of patients is higher than the total excluded (many patients met 1 or more exclusion criteria)

Table 1
Baseline Patient Characteristics

Characteristic	Category	Patients N=159
Treatment	⁹⁰ Y Radioembolization	92 (58)
	Chemoembolization	67 (42)
Age (years)	<65	64 (40)
	≥65	95 (60)
Ethnic Group	Caucasian	117 (74)
	Asian	18 (11)
	Hispanic	11 (7)
	African-American	13 (8)
Gender	Male	117 (74)
	Female	42 (26)
Etiology	Alcohol	21 (13)
	Hepatitis C Virus (HCV)	54 (34)
	HCV + Alcohol	4 (3)
	Hepatitis B Virus	21 (13)
	Nonalcoholic Steatohepatitis	1 (1)
	Cryptogenic	34 (21)
	Autoimmune Hepatitis	2 (1)
	Primary Biliary Cirrhosis	1 (1)
	Hemochromatosis	5 (3)
	Hemochromatosis + HCV	1 (1)
	Unknown	15 (9)
Method of Diagnosis	Imaging	77 (48)
	Biopsy	82 (52)
Portal Hypertension	Present	102 (64)
	Absent	57 (36)
Ascites	Present	8 (5)
	Absent	151 (95)
Cirrhosis	Present	131 (82)
	Absent	28 (18)
Lobar Distribution	Unilobar	87 (55)
	Bilobar	72 (45)
Distribution	Solitary	65 (41)
	Multifocal	94 (59)
Largest Tumor Size (cm)	Median	4.9
	Range	1.1-20.2
Distribution of tumor size	≤ 5 cm	84 (53)
	>5-10 cm	51 (32)

Characteristic	Category	Patients N=159
	> 10	24 (15)
*AFP (ng/mL)	>200	48 (31)
	≤200	110 (69)
ECOG	0	101 (63)
	1	57 (36)
	2	1 (1)
UNOS	T1	3 (2)
	T2	43 (27)
	T3	50 (31)
	T4a (≥4 tumors)	63 (40)
BCLC	A	43 (27)
	B	95 (60)
	C	21 (13)
Child-Pugh	A	110 (69)
	B7	49 (31)

* Baseline AFP was not available in one patient

Table 2

Survival Analysis by Response Status

	Statistical Method		WHO						EASL					
			6-month			12-month			6-month			12-month		
			R	NR	PD	R	SD	PD	R	SD	PD	R	SD	PD
Child-Pugh ≤B7	Landmark	R vs. NR (SD+PD)												
			N											
			Survival*											
			P-value			0.0694			0.0020			<0.0001		
	Landmark	DC (R+SD) vs. PD												
			N											
			Survival*											
			P-value			0.0008			0.002			0.0001		
	6-month risk of death following the Landmark													
			N											
Child-Pugh A	Landmark	R vs. NR (SD+PD)												
			N											
			Survival*											
			P-value			0.0759			0.0584			<0.0001		
	Mantel-Byar**													
			N											
			D											
			A+											
	Mantel-Byar**													
			C++											
			Exp A											
			Chi-Square											
	Mantel-Byar**													
			Relative Odds											
			P-value			0.596			<0.0001					

Statistical Method			WHO								EASL							
		N	10	31	32	16	6	31	30	19	24	21	4	28				
		Survival*	31.6	22	11.4	36.4	NC	7.6	22.3	14	12.1	36.4	18.4	7.6				
		P-value	0.011				0.0001		0.086		0.0001							
			R	NR	R	NR	R	NR	R	NR	R	NR						
	6-month risk of death following the Landmark		1/10 (10)	7/63 (11.1)	0/16 (0)	14/37 (37.8)	2/30 (6.6)	6/43 (13.9)	1/21 (4.8)	13/32 (40.6)								
Mantel-Byar**	Chi-Square	0.194		6.396		0.360		6.649										
	P-value	0.659		0.001		0.548		0.009										
	N	110		110		110												
	D	84		84		84												
	A ⁺	10		10		16												
	C ⁺⁺	74		74		68												
	Exp A	8.22		8.22		30.47												
	Chi-square	0.22		0.22		10.17												
	Relative Odds	1.2504		1.2504		0.409												
	P-value	0.6375		0.6375		0.0014												

Abbreviations : CP, Child-Pugh; CI, confidence interval; D, Total Deaths; DC, disease control; EASL, European Association for Study of the Liver; Exp A, expected deaths among responders; R, responders; NC, calculable; NR, not nonresponders; N, Total Number; PD, progressive disease; SD, stable disease; WHO, World Health Organization

* Survival is expressed as median (months);

** 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 3

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	
6-month	WHO	R	0.526 (0.263-1.052)	0.0694		0.91 (0.42-2.006)	0.832	
		NR	1.00	1.00				
	EASL	R	0.487 (0.309-0.769)	0.0020	0.016	0.46 (0.27-0.80)	0.006	
		NR	1.00		1.00			
	Age	≥65	1.687 (1.086-2.622)	0.0199	0.16	0.51 (0.30-0.85)	0.011	
		<65	1.00		1.00			
	Bilirubin	≤2 mg/dL	0.165 (0.077-0.354)	<.0001	0.0008	0.13 (0.05-0.31)	<0.0001	
		>2 mg/dL	1.00		1.00			
	Albumin	≥3.5 mg/dL	0.617 (0.380-1.003)	0.0513		0.75 (0.44-1.31)	0.32	
		<3.5 mg/dL	1.00		1.00			
	INR	≤1.2	0.739 (0.465-1.172)	0.1986		0.87 (0.51-1.49)	0.628	
		>1.2	1.00		1.00			
	Baseline Tumor Size	≤5 cm	1.009 (0.57-1.78)	0.976				
		>5-10 cm	0.86 (0.45-1.66)	0.668				
		>10 cm	1.00					
Treatment	90Y	0.61 (0.39 – 0.97)	0.02	0.16	0.68 (0.41-1.10)	0.124		
	TACE	1.00			1			
WHO	R	0.38 (0.22-0.66)	0.004	0.032	0.68 (0.24-1.92)	0.475		
	NR	1.00			1.00			
ESAL	R	0.251 (0.129-0.487)	<.0001	0.0008	0.20 (0.07-0.55)	0.002		
	NR	1.00			1.00			
Age	≥ 65	1.791 (1.033-3.105)	0.0378	0.30	0.67 (0.34-1.33)	0.260		
	< 65	1.00			1.00			
Bilirubin	≤2 mg/dL	0.374 (0.158-0.888)	0.0258	0.206	0.65 (0.21-1.9)	0.46		
	>2 mg/dL	1.00			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 mg/dL	0.412 (0.223-0.763)	0.0048	0.038	0.47 (0.23-0.98)	0.04	
		<3.5 mg/dL	1.00			1.00		
	INR	≤1.2	0.711 (0.403-1.255)	0.2394		0.63 (0.30-1.29)	0.21	
		>1.2	1.00			1.00		
	Baseline Tumor Size	≤5 cm	0.89 (0.45-1.75)	0.736				
		>5-10 cm	0.75 (0.34-1.64)	0.474				
		>10 cm	1.00					
	Treatment	⁹⁰ Y	0.61 (0.34 – 1.09)	0.07	0.56	0.82 (0.41-1.64)	0.588	
		TACE	1.00			1.00		

Abbreviations: CI, confidence interval; EASL, European Association for the Study of the Liver; HR, hazard ratio; INR, International Normalized ratio; TACE, Transarterial chemoembolization; WHO, World Health Organization; ⁹⁰Y, Yttrium-90 radioembolization

* Adjusted for multiple comparisons using Bonferroni methodology (correction factor n=8).

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

Table 4
Distribution of responders and nonresponders based on baseline tumor size

Landmark	Size	WHO			EASL		
		R	NR	P value	R	NR	P value
6 months	≤5 cm	11	49	0.36	27	33	0.27
	>5-10 cm	2	26		11	17	
	>10 cm	2	15		4	13	
12 months	≤5 cm	11	30	0.90	14	27	0.96
	>5-10 cm	6	14		7	13	
	>10 cm	3	10		4	9	

Abbreviation: EASL, European Association for the Study of the Liver; NR, nonresponders; R, responders; WHO, World Health Organization