

The Significance of Early Alpha-Fetoprotein Level Changes in Predicting Clinical and Survival Benefits in Advanced Hepatocellular Carcinoma Patients Receiving Sorafenib

THOMAS YAU,^{a,b} T.J. YAO,^c PIERRE CHAN,^a HILDA WONG,^a ROBERTA PANG,^b SHEUNG TAT FAN,^b RONNIE T.P. POON^b

Departments of ^aMedicine and ^bSurgery and ^cClinical Trials Centre, Queen Mary Hospital, Hong Kong, China

Key Words. Hepatocellular carcinoma • Sorafenib • Alpha-fetoprotein • Surrogate

Disclosures: Thomas Yau: None; T.J. Yao: None; Pierre Chan: None; Hilda Wong: None; Roberta Pang: None; Sheung Tat Fan: None; Ronnie T.P. Poon: None.

The content of this article has been reviewed by independent peer reviewers to ensure that it is balanced, objective, and free from commercial bias. No financial relationships relevant to the content of this article have been disclosed by the authors or independent peer reviewers.

ABSTRACT

Background. The role of serum alpha-fetoprotein (AFP) changes in predicting the treatment outcomes of advanced hepatocellular carcinoma (HCC) patients to sorafenib remains unknown.

Methods. Serum AFP was collected prospectively at baseline and subsequent follow-up visits in parallel with clinical and survival outcomes. AFP response was defined as a relative drop of AFP >20% of the baseline level after 6 weeks of sorafenib. The relationship between AFP response and the treatment outcomes was first explored in patients who received sorafenib in a phase II study. Subsequently, an independent validation set of patients were obtained to validate the association of AFP response to clinical outcomes.

Results. Included in the exploration and validation sets for analysis were 41 and 53 patients, respectively, with baseline AFP level >20 µg/L. In the exploration

cohort, AFP response was significantly associated with clinical benefit (CB) rate (relative chance 3.4, 95% confidence interval [CI], 1.1–11.1), and multivariate analysis indicated that AFP response was associated with significantly better progression-free survival (PFS) (hazard ratio [HR], 0.31; 95% CI, 0.13–0.76) and marginally better overall survival (OS) (HR, 0.30; 95% CI, 0.09–1.02). When applying AFP changes in the validation set, significant associations were again found between AFP response with CB rate (relative chance, 5.5; 95% CI, 2.3–13.6) and PFS (HR, 0.12; 95% CI, 0.04–0.30) but not OS (HR, 0.61; 95% CI, 0.27–1.26).

Conclusion. Drop in AFP level at 6 weeks is an exploratory early surrogate for both CB and PFS in advanced HCC patients receiving sorafenib. *The Oncologist* 2011; 16:1270–1279

INTRODUCTION

Hepatocellular cancer (HCC) is the sixth- and eleventh-most common cancer worldwide in men and women, re-

spectively [1]. In the pivotal phase III randomized trials conducted in the Western [2] and Asian [3] patient populations, single-agent sorafenib has shown a significant

Correspondence: Ronnie Poon, F.R.C.S., Ph.D., Department of Surgery, Queen Mary Hospital, Hong Kong, China. Telephone: 852-2855-3111; Fax: 852-2816-2863; e-mail: poontp@hkucc.hku.hk Received April 1, 2011; accepted for publication June 29, 2011; first published online in *The Oncologist Express* on September 1, 2011. ©AlphaMed Press 1083-7159/2011/\$30.00/0 <http://dx.doi.org/10.1634/theoncologist.2011-0105>

survival improvement in treating advanced HCC patients.

In solid tumors, the tumor response to any systemic regime relied heavily on the objective radiological assessment as outlined by the criteria according to World Health Organization [4] or Response Evaluation Criteria in Solid Tumors (RECIST) [5]. However, most advanced HCC patients have underlying cirrhosis; in the background of cirrhosis, accurate radiological assessment of the liver lesions is difficult because of the underlying fibrosis and the presence of regeneration nodules [6, 7]. Importantly, in the era of targeted therapy, most of the targeted agents like sorafenib are rather cytostatic than cytotoxic. Thus, the tumor may respond well without significant radiological changes [8, 9]. Especially in HCC, it is well known that sorafenib can induce tumor necrosis that can initially manifest as enlarging lesion rather than shrinkage [9]. Recently, the HCC expert consensus meeting proposed the use of modified RECIST (mRECIST) criteria, which mainly measured the tumor viability in assessing advanced HCC patients on targeted therapy [10]. Nonetheless, the mRECIST criterion is still not widely adopted yet. Therefore, there is an unmet need to search for more robust ways to assess the response of the patients receiving sorafenib treatment for advanced HCC, particularly the role of serum biomarkers.

Alpha-fetoprotein (AFP) has been widely used as the tumor marker for HCC. Recent studies had demonstrated that serial changes of AFP have an important role in predicting response in advanced HCC receiving chemotherapy [8], thalidomide [11], or a combination of various anti-angiogenic agents with metronomic oral 5-fluoropyrimidine [12]. Nevertheless, single-agent sorafenib is the only compound approved by the U.S. Food and Drug Administration and regulatory authorities worldwide for the management of advanced HCC patients [2]. There is a lack of published data regarding the potential biomarkers in predicting the tumor response or survival benefit from single-agent sorafenib in advanced HCC patients. Thus, the aim of the current study was to investigate the value of early AFP changes as a predictor of treatment outcomes to single-agent sorafenib.

PATIENTS AND METHODS

The current study was part of the study approved by the ethics committee of Queen Mary Hospital, The University of Hong Kong.

Study Population

Similar to prior studies [8], we only analyzed the patients with baseline AFP level $>20 \mu\text{g/L}$. An AFP response was defined as a relative drop of AFP $>20\%$ of the baseline

level after 6 weeks of sorafenib treatment in the current study. Twenty percent drop in AFP was chosen as the cutoff point because of prior data suggesting a significant prognostic value of this level of AFP drop in patients undergoing chemotherapy or anti-angiogenic therapy for HCC [8, 12].

(A) Exploration Set

Between November 2006 and January 2008, a phase II, open-label, single-arm sorafenib study for patients with advanced HCC was conducted at Queen Mary Hospital, The University of Hong Kong [13]. Written consents were obtained from the patients before enrollment.

Among the 51 patients enrolled in the phase II study, 41 patients had elevated baseline AFP levels ($>20 \mu\text{g/L}$) and were included in the exploration set for analysis.

(B) Validation Set

Between January 2008 and September 2009, there were 122 consecutive advanced HCC patients who received single-agent sorafenib as the routine treatment for the disease at Queen Mary Hospital, The University of Hong Kong. All the patients had signed the informed consent before taking sorafenib. Among these 122 patients, 26 patients were excluded from the analysis because they discontinued sorafenib early mainly due to treatment-related toxicities and there were no radiological assessments for proper treatment-response evaluation. Another 8 patients were also excluded because they did not have measurable disease according to RECIST 1.0 criteria. Moreover, 5 patients were not included as they had either baseline and/or 6-week AFP values missing for evaluation. Among the remaining 83 patients, who received sorafenib as a routine treatment for advanced HCC, 53 patients had elevated baseline AFP level ($>20 \mu\text{g/L}$) and were included in the validation set.

Diagnosis, Disease Assessment, and Follow-up

The initial diagnosis of HCC was based on increased serum AFP $>400 \mu\text{g/L}$ and/or typical imaging appearance according to the criteria of European Association of Study of the Liver [14]. Histology was obtained only in the case of diagnostic uncertainty. In the phase II study, advanced HCC patients with measurable disease who were not amenable to surgical resection, liver transplantation, or locoregional therapies were eligible to enter the study. All patients were initially put on single-agent sorafenib 400 mg twice daily in 4-week cycles and followed up regularly every 2 weeks. Disease assessment was performed by computed tomography around every 12 weeks. Likewise, patients treated with sorafenib outside the phase II trial had similar although less stringent eligibility criteria. Disease assessment was also performed mostly by computed to-

mography every 12–14 weeks. All the patients were followed with regular clinical examination, blood tests, and scanning until death or last follow-up.

All the radiological assessments were performed by three designated radiologists at Queen Mary Hospital, The University of Hong Kong. Responses were classified according to RECIST 1.0 criteria [5] as complete response, partial response (PR), stable disease (SD), or progressive disease. Patients were defined as having CB on sorafenib if they had a best response of complete response, PR, or SD. Only patients who had CB continued sorafenib until radiological progression, unacceptable treatment toxicities, or death. Progression-free survival (PFS) was measured from the date of treatment commencement until disease progression, death, or last follow-up. Overall survival (OS) was measured from the date of treatment commencement until death or last follow-up. The treatment outcome of patients was measured by CB rate, PFS, and OS.

AFP Analysis

AFP was collected prospectively at baseline and during each follow-up visit when the patients were treated with sorafenib in both cohorts. AFP was measured every 2 weeks when patients received sorafenib in the phase II trial and 2–3 weeks instead outside the phase II study. It was measured by the microparticle enzyme immunoassay (Abbott Laboratories, Chicago).

Statistical Analysis

The relationship between AFP response and CB rate was examined by χ^2 test. The relative chance of CB for AFP responders to AFP nonresponders was estimated using logistic regression analysis. The relationship between AFP response and survival outcomes was examined by plotting the Kaplan-Meier curves and further by the Cox proportional hazards regression analysis. Other potential parameters that may affect the survival outcome were assessed as well. These parameters included age, sex, large (size >10 cm) or diffuse tumor, bilobar involvement, portal vein invasion, presence of distant metastases, hepatitis B viral status, hepatitis C viral status, the use of antiviral therapy, presence of baseline ascites, baseline liver function, and Eastern Cooperative Oncology Group performance status. First, AFP response and each of the potential factors were evaluated univariately. All prognostic factors with $p < .1$ in univariate analysis were identified and included in a multivariate Cox regression analysis [15] to study the effect of AFP response adjusting for other risk factors. All statistically significant prognostic factors were then selected by a stepwise procedure as described in SAS Procedure PHREG [16]. A variable was removed if its Wald test result was not

significant (i.e., $p > .1$). The proportional hazards assumption of the final model was assessed by the method described by Lin and Ying [17].

An independent temporal validation set of patients was obtained to validate the finding of the association of AFP response to clinical outcomes. The sample size was set largely by issues of feasibility with the minimal requirement of no smaller than the size of the exploration set. We used two methods to assess the validity of AFP response as a prognostic factor for clinical outcomes. First, the association of the AFP response and the CB rate was examined in the validation set by χ^2 test, and the association of AFP response and PFS or OS was examined by log-rank test. The multivariate model for PFS identified using the phase II trial set was fit to the validation set using Cox regression to examine the significance of AFP response adjusting for other possible covariates. Second, PSEP as an index of separation [18] between AFP responders and nonresponders was calculated. We did not intend to validate the multivariate model as a prognostic scoring tool but rather to validate the prognostic ability of AFP response. Therefore, the difference of the predicted probability of dying for a patient with and without AFP response was calculated at 12 and 16 weeks for both the exploration and validation sets.

All statistical tests were two-sided. All statistical analyses were performed using SAS 9.2 (SAS Institute, Cary, NC, 2003).

RESULTS

Demographic Data of Patients in Exploration Set and Validation Set

Table 1 shows the demographic data of the 41 and 53 patients included in the analysis of AFP in the exploration and validation sets, respectively. Overall, these two sets of patients had comparable baseline characteristics. Nevertheless, the patients in the validation set had poorer performance status than patients in the exploration set ($p < .001$) as they were nontrial patients. Moreover, the liver function of the patients in the exploration set was in general better than the patients in the validation set.

Exploration Set

Among the 41 patients studied in the exploration set, 8 patients had achieved CB with sorafenib (1 PR, 7 SD). The median PFS and OS were 14 weeks (95% confidence interval [CI], 13–15 weeks) and 22 weeks (95% CI, 18–31 weeks), respectively. Nine patients were AFP responders, 1 patient had missing 6-week AFP value for evaluation, and the remaining 31 patients were nonresponders.

Table 1. Demographic data of patients included in the exploration set and the validation set

	Exploration set	Validation set	<i>p</i>
Age (years)			
Median	55	58	
Range	28–74	19–84	.23
Sex			
Male	36 (88%)	44 (83%)	
Female	5 (12%)	9 (17%)	.57
ECOG performance status			
0	30 (73%)	15 (28%)	
1	8 (20%)	36 (68%)	
2	3 (7%)	2 (4%)	<.001
Hepatitis serology			
Hep Bs Ag positive	37 (90%)	45 (85%)	.54
Anti-HCV Ab ^a positive	2 (5%)	7 (18%)	.16
On antiviral therapy	17 (45%)	28 (53%)	.53
Child-Pugh status ^b			
A	30 (73%)	31 (60%)	
B	10 (24%)	20 (38%)	
C	1 (2%)	1 (2%)	.36
AJCC disease staging			
I	0 (0%)	3 (6%)	
II	3 (7%)	8 (15%)	
IIIA	15 (37%)	15 (28%)	
IIIB	0 (0%)	0 (0%)	
IIIC	2 (5%)	9 (17%)	
IV	21 (51%)	18 (34%)	.08
Tumor size >10 cm or diffuse	20 (49%)	19 (36%)	.29
Bilobar involvement	27 (66%)	21 (40%)	.01
Presence of ascites			
Absent	35 (85%)	38 (72%)	
Slight	3 (7%)	8 (15%)	.34
Moderate	3 (7%)	7 (13%)	
Distant metastases	21 (51%)	26 (49%)	1.0
Portal vein invasion	18 (44%)	21 (40%)	.83
Prior treatment	23 (56%)	34 (64%)	.52
Surgical treatment	9 (22%)	22 (42%)	.05
TACE	15 (37%)	21 (40%)	.83
RFA	4 (10%)	11 (21%)	.17
Chemotherapy	4 (10%)	5 (9%)	1.0
Baseline liver function (range)			
Alkaline phosphatase (μ/L)	143 (53–699)	164 (49–775)	.08
Alanine aminotransferase (μ/L)	62 (14–220)	57 (16–437)	.35
Aspartate aminotransferase (μ/L) ^b	80 (23–337)	104 (25–389)	.03
Albumin (g/L)	37 (23–47)	37 (23–49)	.89
Vilirubin (μmol/L)	19 (6–85)	17 (5–119)	.91
Prothrombin time (s) prolonged	13 (9.8–17.5)	12.8 (10.2–21)	.96
Median alpha-fetoprotein (range)	1,766 (25–1,106,800)	1,471 (22–2,193,000)	.79

^aThree missing for the phase II set, and 13 missing for the validation set.
^bOne missing for the validation set.
 Abbreviations: AJCC, American Joint Committee on Cancer; ECOG, Eastern Cooperative Oncology Group; RFA, radiofrequency ablation; TACE, transcatheter arterial chemoembolization.

Table 2. Univariate and multivariate analysis of potential prognostic factors for progression-free survival in the exploration set

	Univariate analysis		Multivariate analysis		
	<i>p</i>	Hazard ratio	<i>p</i>	Hazard ratio	95% CI
Age	.91	1.00			
Male	.63	1.29			
Alpha-fetoprotein response	.09	0.49	.01	0.31	0.13–0.76
Tumor size >10 cm/diffuse	.06	1.95	<.01	3.34	1.53–7.30
Bilobar involvement	.12	0.57			
Portal vein invasion	.50	1.25			
Distant metastases	.30	1.43			
HBsAg +ve	.09	0.34	.02	0.21	0.06–0.80
Anti-HCV +ve ^a	.63	1.64			
Antivirus therapy	.66	0.86			
Baseline ascites	.51	0.80			
Baseline liver function					
Alanine aminotransferase	.11	1.01			
Aspartate aminotransferase	.07	1.01			
Albumin	.48	0.98			
Bilirubin	.81	1.00			
Alkaline phosphatase	.14	1.00			
Prothrombin time	.65	1.06			
ECOG performance status	.19	1.47			

^aThree missing for the phase II set, and 13 missing for the validation set.
^bOne missing for the validation set.
 Abbreviations: CI, confidence interval; ECOG, Eastern Cooperative Oncology Group; HCV, hepatitis C virus.

(A) Relationship between AFP Response and Clinical Benefit Rate at 12 Weeks

AFP response ($p = .04$) was significantly associated with CB rate at 12 weeks. The relative chance of achieving CB for AFP responders (44.4%) to AFP nonresponders (12.9%) was estimated to be 3.4 (95% CI, 1.1–11.1).

(B) Relationship between AFP Response and Survival Benefits

Table 2 shows the results of univariate and multivariate analyses of potential prognostic factors for PFS. The multivariate analysis indicated that AFP response ($p = .01$; hazard ratio [HR], 0.31; 95% CI, 0.13–0.76) was one of the independent prognostic factors significantly associated with better PFS.

Table 3 shows the results of univariate and multivariate analyses of potential prognostic factors for OS. The multivariate analysis indicated AFP response ($p = .05$; HR, 0.30; 95% CI, 0.09–1.02) was the only independent prognostic factor marginally associated with better OS.

Validation Set

Among the 53 patients in the independent validation set, 14 patients had achieved CB with sorafenib (1 PR, 13 SD). The median PFS and OS were 12 weeks (95% CI, 10–13 weeks) and 22 weeks (95% CI, 17–31 weeks), respectively. Overall, 13 (24.5%) of 53 patients had AFP response.

(A) Relationship between AFP Response and Clinical Benefit Rate at 12 Weeks

The association between AFP response and CB rate was significant ($p < .001$); 9 of 13 (69.2%) AFP responders had CB, and only 5 of 40 (13%) non-AFP responders had CB. The relative chance of radiological CB for AFP responders to AFP nonresponders in the validation set was estimated to be 5.5 (95% CI, 2.3–13.6).

(B) Relationship between AFP Response and Survival Benefits

All patients have progressed or died; only two patients were alive without progression, and they were both AFP

Table 3. Univariate and multivariate analysis of potential prognostic factors for overall survival in the exploration set

	Univariate analysis		Multivariate analysis		
	<i>p</i>	Hazard ratio	<i>p</i>	Hazard ratio	95.0% CI
Age	.40	0.98			
Male	.69	1.35			
Alpha-fetoprotein response ^a	.09	0.35	.05	0.3	0.09–1.02
Tumor size >10 cm/diffuse	.01	2.74	<.01	4.26	1.53–11.83
Bilobar involvement	.09	2.13			
Portal vein invasion	.23	1.61			
Distant metastases	.18	1.72			
AJCC stage					
HBsAg +ve	.32	2.79			
Anti-HCV +ve ^a	.99	0.00			
Antivirus therapy	.21	1.67			
Baseline ascites	.45	1.32			
Baseline liver function					
Alanine aminotransferase	<.01	1.02			
Aspartate aminotransferase	<.01	1.01	<.01	1.01	1.01–1.02
Albumin	.20	0.96			
Bilirubin	.04	1.03			
Alkaline phosphatase	<.01	1.01			
Prothrombin time	.06	1.29			
ECOG PS	.53	1.23			

Abbreviations: AJCC, American Joint Committee on Cancer; CI, confidence interval; ECOG, Eastern Cooperative Oncology Group; HCV, hepatitis C virus.

responders. The Kaplan-Meier estimates of PFS and OS at 12 and 16 weeks by AFP response are listed in Table 4.

The AFP response was highly associated with PFS (*p* < .001) (Figure 1). The relative hazard of progression for AFP responders to nonresponders was estimated to be 0.12 (95% CI, 0.04–0.30, *p* < .001), adjusted for hepatitis B and large or diffused tumor. This adjusted estimate of HR was quite different from the estimate from the exploration set, but this nonetheless gave supporting evidence that AFP response was predictive of longer time to progression. Notably, the PSEP values for PFS in the validation set were much larger than in the exploration set, suggesting that overoptimism was quite unlikely.

On the other hand, AFP response was not a significant predictor for OS in these 53 patients. The relative hazard from the univariate Cox regression was estimated to be 0.61 (95% CI, 0.27–1.26, *p* = .20) adjusting for baseline aspartate aminotransferase and large or diffused tumor. The result indicated that it was inconclusive whether AFP had predictive effect of OS.

(C) Predictive Value of 6-Week Drop in AFP

When applying our current definition of AFP level and AFP drop, AFP level >20 μg/L at baseline and >20% drop in AFP after 6 weeks of sorafenib treatment in predicting the CB rate in the validation set, the positive and negative predictive values were 69.2% and 87.5%, respectively. Likewise, when the same definition was used in predicting the PFS at 16 weeks in the validation set, the positive and negative predictive values were 77% and 93%, respectively.

Relationship between Hepatitis Status, Use of Antiviral Therapy, AFP Response, and PFS

The majority of the analyzed patients in the exploration (90%) and validation set (85%) were hepatitis B carriers. Moreover, 45% of patients in the exploration set and 53% of patients in the validation set received antiviral treatment during sorafenib treatment for advanced HCC. Notably, the use of antiviral therapy did not seem to affect the prognostic value of AFP response on PFS: In the exploration set, 4 of 17 (23.5%) and 5 of 20 (25%) patients with or without antiviral therapy, respectively, had AFP response. The dif-

Table 4. Summary of median survival, probability, and PSEP of the exploration and validation sets

	Exploration set (n = 41)	Validation set (n = 53)
Progression-free survival, wk(s)		
AFP responder		
Median	15.9	21.7
95% CI	11.0–22.9	13.0–30.4
AFP nonresponder		
Median	13.6	10.9
95% CI	12.3–14.1	9.0–12.2
Progression-free survival probability, %		
AFP responder		
At 12 wk(s)	88.9	84.6
At 16 wk(s)	44.4	76.9
AFP nonresponder		
At 12 wk(s)	83.3	42.5
At 16 wk(s)	24.3	7.5
PSEP, %		
At 12 wk(s)	5.6	42.1
At 16 wk(s)	20.1	69.4
Overall survival, wk(s)		
AFP responder		
Median	Undefined	30.4
95% CI	15.1-undefined	13.0–53.2
AFP nonresponder		
Median	21.4	21.0
95% CI	16.9–27.7	16.7–32.2
Overall survival probability, %		
AFP responder		
At 18 wk(s)	88.9	61.5
At 24 wk(s)	63.5	61.5
AFP nonresponder		
At 18 wk(s)	60.9	64.0
At 24 wk(s)	39.4	41.7
PSEP, %		
At 18 wk(s)	20.0	–2.5
At 24 wk(s)	24.1	19.8

Abbreviations: AFP, alpha-fetoprotein, CI, confidence interval; PSEP, a simple index of separation.

ference was not significant ($p = .92$). Likewise, in the validation cohort, 6 of 28 (21.4%) and 7 of 25 (28%) patients with or without antiviral therapy, respectively, had AFP response. Again, the difference was not significant ($p = .58$). Moreover, when we pooled all the patients on antiviral ther-

apy in both the exploration and validation sets for analysis, the AFP responders ($n = 10$) still had significantly better median PFS than nonresponders ($n = 35$) (23 weeks, 95% CI, 15–30 weeks, versus 12 weeks, 95% CI, 10–12 weeks, $p < .001$) (Figure 2).

DISCUSSION

There is an absence of published data in the literature regarding the trend of AFP changes in patients receiving targeted therapy treatment for advanced HCC [6]. Llovet et al. reported the baseline level of AFP as one of the prognostic indicators for survival in Sorafenib HCC Assessment Randomized Protocol (SHARP) [2] trial; however, the serial AFP response to sorafenib was not provided. Similarly, in Asian sorafenib trials [3], the serial change of AFP during sorafenib treatment was not measured either. Despite the established role of using AFP in screening, diagnosis, and surveillance of HCC, its use as a surrogate marker for treatment response remains controversial [6]. More recently, there is renewed interest in exploring the role of serial AFP changes as a serological surrogate for treatment response in HCC patients receiving systemic treatment. Chan et al. [8] reported that a 20% decrease in serum AFP was strongly associated with radiological response and overall survival benefit in advanced HCC patients receiving systemic chemotherapy. Similarly, a retrospective analysis by Vora et al. [6] suggested that >50% AFP decrease from the baseline during treatment may serve as a useful surrogate marker for clinical outcome in patients with advanced HCC receiving systemic therapy. Interestingly, in this analysis, patients were recruited from five different clinical trials: two chemotherapy trials, one thalidomide trial, and two targeted therapy trials. The result did show significant correlation with radiological response but no overall survival benefit in this heterogeneous study population. Moreover, Shao et al. [12] retrospectively analyzed patients with advanced HCC who had been enrolled in three prospective phase 2 clinical trials that evaluated either sorafenib, bevacizumab, or thalidomide in combination with a metronomic, oral 5-fluoropyrimidine as first-line systemic therapy for advanced HCC. In this study, an early AFP response was again defined as a decline >20% from baseline after 2–4 weeks of sorafenib treatment. The results suggested that early AFP response is a useful surrogate marker to predict treatment response and prognosis in advanced HCC patients who received the targeted therapy together with metronomic chemotherapy. Notably, patients included in these three trials are mostly patients either receiving systemic chemotherapy alone or in combination with various targeted agents for advanced HCC. Nevertheless, no chemotherapy regime thus far shows any significant survival advantage in treating patients with advanced HCC, and sorafenib is now the standard in treating advanced HCC patients who are not amenable to sur-

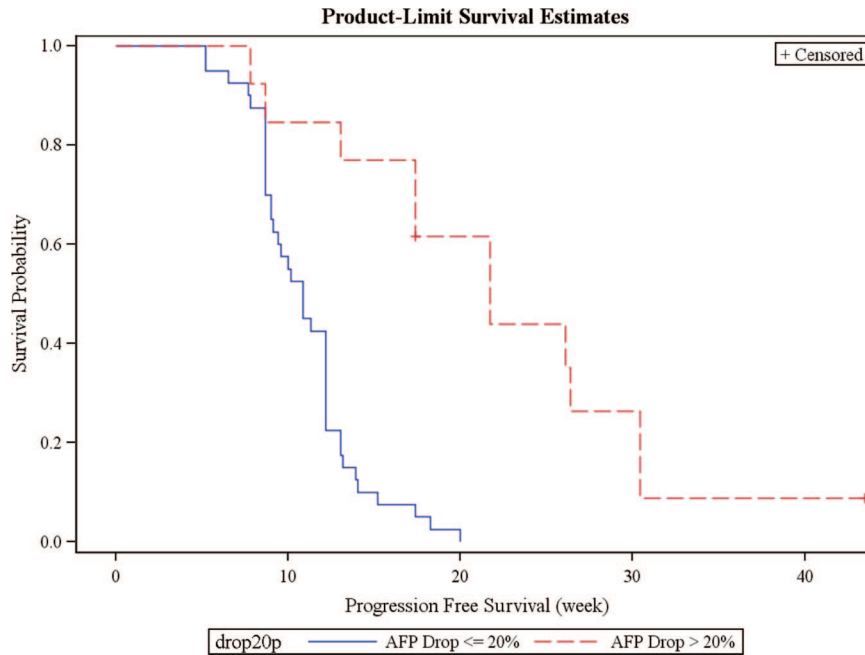


Figure 1. Kaplan-Meier PFS of AFP responders and nonresponders in the validation set. Abbreviations: AFP, alpha-fetoprotein; PFS, progression-free survival.

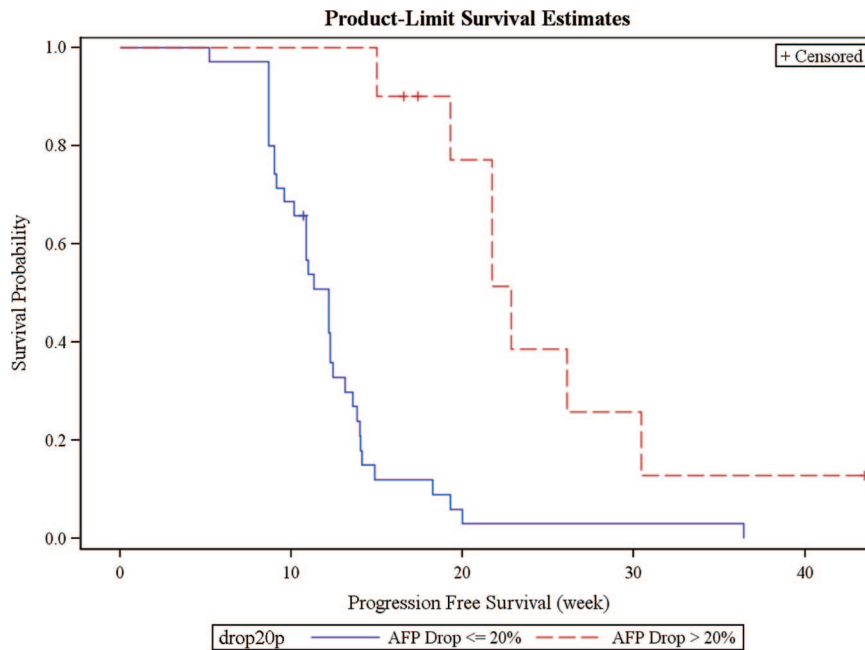


Figure 2. Kaplan-Meier PFS of AFP responders and nonresponders in the patients on antiviral therapy in both exploration and validation sets.

Abbreviations: AFP, alpha-fetoprotein; PFS, progression-free survival.

gical intervention according to the latest National Comprehensive Cancer Network guideline [19]. Therefore, it is important to establish the role of AFP changes with the treatment outcome of the patients on sorafenib.

In this study, we consistently demonstrated that AFP drops >20% after 6 weeks of sorafenib treatment signifi-

cantly correlated with CB rate and also served as an early surrogate marker for PFS benefit in advanced HCC patients receiving sorafenib. Thus, clinicians may use this early AFP change to stratify the patients who have poor AFP response for earlier radiological assessment and consideration to switch to other systemic treatment or participation

in clinical trials. This is of particular importance in view of the high cost and significant side effects associated with sorafenib treatment. Notably, the measurement of AFP is a simple, quick, and cheap test. It is operator-independent [20] and routinely available in most parts of the world. This may have a particular value in countries where reliable radiological assessment is not readily accessible.

There are several unsettled issues worth addressing before recommending the use of AFP response as a surrogate endpoint in routine clinical practice and/or clinical trial. First, there is no consensus yet regarding the magnitude of decrease in AFP drop in defining AFP response. Notably, different studies used different arbitrarily defined cutoff values for defining AFP response [6, 8, 12]. One study suggested that the specificity of using serial AFP changes as a surrogate was enhanced if the cutoff point was defined as $>50\%$ [6]. Although the use of a more stringent criterion may potentially improve the specificity in using AFP response to predict treatment outcomes, it will become less useful as fewer patients can be classified as having AFP response because patients on sorafenib rarely have such a dramatic drop in AFP. Second, studies showed that there was a higher specificity in using AFP response to predict treatment outcomes if patients with higher baseline AFP were recruited because patients with lower baseline AFP level are more prone to have false-positive AFP response. Thus, some studies [11] suggested that only patients with higher baseline AFP such as $>200 \mu\text{g/L}$ should be analyzed. However, when applying this criterion in the present analysis, the results may be biased as only a few patients in both the exploratory and validation sets had such a high AFP level. Therefore, further studies are urgently needed to provide a more uniform definition of AFP response and to guide defining the desired baseline level of AFP for analysis. Third, there are laboratory variations in using different assays of AFP values. Thus, cautious interpretation should be exercised when trying to compare different AFP values measured from different laboratories.

The major limitations of the study are the relatively small number of patients analyzed and that it is retrospective in nature. Nonetheless, both the exploratory and validation sets are based on the database with all the clinical parameters, AFP levels, radiological assessment results, and survival outcome prospectively collected for analysis. This helps to minimize the selection bias by just analyzing a nontrial retrospective dataset. The other limitation is that the change in AFP level may also be affected by the patient's underlying liver disease. Especially in the areas of

endemic hepatitis infection, active hepatitis may lead to a modest increase in AFP level and confuse the interpretation of AFP response. Notably, although there was a high prevalence of hepatitis B carriers in the exploration and the validation cohorts, our finding demonstrated that the use of antiviral therapy did not significantly affect the prognostic value of AFP response on PFS. It is not known whether there is any benefit in using antiviral therapy in advanced HCC patients receiving sorafenib and with underlying hepatitis B or C infection. Nevertheless, the use of antiviral therapy may lead to optimal control of underlying hepatitis and thus provide stable and normalized viral titers for a more reliable interpretation of AFP response in advanced HCC patients with active hepatitis.

In the coming years, the use of serial AFP changes as a surrogate endpoint should be revisited rigorously in other treatment modalities of HCC, such as post-hepatectomy or other local ablative procedures. Recently, in fact, serial AFP changes have also been shown to correlate to radiologic response, progression, and survival in patients undergoing chemo-embolization or radioembolization [20]. Furthermore, the use of other more sophisticated serologic markers such as fucosylated fraction of AFP [21] or Des-gamma carboxyprothrombin [21] may be used together with serial AFP changes to more accurately predict the response to the targeted therapy in HCC patients in the future.

In conclusion, drop in serum AFP $>20\%$ in the first 6 weeks may be a useful early surrogate endpoint for CB and PFS benefits in advanced HCC patients receiving sorafenib treatment. Nevertheless, because of the relatively small sample size included in the analysis, our finding should still be regarded as exploratory rather than confirmatory. Moreover, clinicians should be cautious in applying AFP response in predicting treatment outcome, especially in patients with active hepatitis as well as those with low baseline AFP values. Further prospective studies in large cohorts of HCC patients treated with sorafenib are warranted to confirm the result of this study.

AUTHOR CONTRIBUTIONS

Conception/Design: Thomas Yau, Pierre Chan, Ronnie T.P. Poon
Provision of study material or patients: Pierre Chan, Sheung Tat Fan, Ronnie T.P. Poon
Collection and/or assembly of data: Thomas Yau, Pierre Chan, Hilda Wong, Sheung Tat Fan, Ronnie T.P. Poon
Data analysis and interpretation: Thomas Yau, T.J. Yao, Pierre Chan, Hilda Wong, Roberta Pang, Sheung Tat Fan, Ronnie T.P. Poon
Manuscript writing: Thomas Yau, T. J. Yao, Pierre Chan, Hilda Wong, Roberta Pang, Sheung Tat Fan, Ronnie T.P. Poon
Final approval of manuscript: Thomas Yau, T.J. Yao, Pierre Chan, Hilda Wong, Roberta Pang, Sheung Tat Fan, Ronnie T.P. Poon

REFERENCES

- 1 Hussain SA, Ferry DR, El-Gazzaz G et al. Hepatocellular carcinoma. *Ann Oncol* 2001;12:161–172.
- 2 Llovet JM, Ricci S, Mazzaferro V et al. Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med* 2008;359:378–390.
- 3 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* 2009;10:25–34.
- 4 World Health Organization. WHO Handbook for Reporting Results of Cancer Treatment. WHO offset publication No. 48. Geneva: WHO, 1979.
- 5 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* 2000;92:205–216.
- 6 Vora SR, Zheng H, Stadler ZK et al. Serum alpha-fetoprotein response as a surrogate for clinical outcome in patients receiving systemic therapy for advanced hepatocellular carcinoma. *The Oncologist* 2009;14:717–725.
- 7 Choi BI, Lee JM. Advancement in HCC imaging: diagnosis, staging and treatment efficacy assessments: imaging diagnosis and staging of hepatocellular carcinoma. *J Hepatobiliary Pancreat Sci* 2010;17:369–373.
- 8 Chan SL, Mo FK, Johnson PJ 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.
- 9 Abou-Alfa GK, Schwartz L, Ricci S et al. Phase II study of sorafenib in patients with advanced hepatocellular carcinoma. *J Clin Oncol* 2006;24:4293–4300.
- 10 Llovet JM, Di Bisceglie AM, Bruix J et al. Design and endpoints of clinical trials in hepatocellular carcinoma. *J Natl Cancer Inst* 2008;100:698–711.
- 11 Chen LT, Liu TW, Chao Y et al. Alpha-fetoprotein response predicts survival benefits of thalidomide in advanced hepatocellular carcinoma. *Aliment Pharmacol Ther* 2005;22:217–226.
- 12 Shao YY, Lin ZZ, Hsu C et al. Early alpha-fetoprotein response predicts treatment efficacy of antiangiogenic systemic therapy in patients with advanced hepatocellular carcinoma. *Cancer* 2010;116:4590–4596.
- 13 Yau T, Chan P, Ng KK et al. Phase 2 open-label study of single-agent sorafenib in treating advanced hepatocellular carcinoma in a hepatitis B-endemic Asian population: presence of lung metastasis predicts poor response. *Cancer* 2009;115:428–436.
- 14 Llovet JM, Bru C, Bruix J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. *Semin Liver Dis* 1999;19:329–338.
- 15 Collett D. Modelling survival data in medical research. Chapman & Hall/CRC Texts in Statistical Science. Boca Raton, FL: CRC Press, 2003.
- 16 The PHREG Procedure; Variable Selection Methods. Cary, NC: SAS Institute Inc. Available at http://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_phreg_sect001.htm, accessed 2009.
- 17 Lin DY, Wei LJ, Ying Z. Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* 1993;80:557–572.
- 18 Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000;19:453–473.
- 19 NCCN Clinical Practice Guidelines in Oncology In Edition, 2010. Available at http://www.nccn.org/professionals/physician_gls/f_guidelines.asp, accessed 2010.
- 20 Riaz A, Ryu RK, Kulik LM 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.
- 21 Tamura Y, Igarashi M, Suda T et al. Fucosylated fraction of alpha-fetoprotein as a predictor of prognosis in patients with hepatocellular carcinoma after curative treatment. *Dig Dis Sci* 2010;55:2095–2101.