Alpha-Fetoprotein as a Potential Surrogate Biomarker for Atezolizumab + Bevacizumab Treatment of Hepatocellular Carcinoma

Andrew X. Zhu[1,2], Farshid Dayyani[3], Chia-Jui Yen[4], Zhenggang Ren[5], Yuxian Bai[6], Zhiqiang Meng[7], Hongming Pan[8], Paul Dillon[9], Shivani K. Mhatre[10], Vincent E. Gaillard[11], Sairy Hernandez[12], Robin Kate Kelley[13], and Bruno Sangro[14]

ABSTRACT

Purpose: Atezolizumab + bevacizumab is the new standard of care for systemic treatment-naïve, unresectable hepatocellular carcinoma (HCC). This exploratory study investigated on-treatment alpha-fetoprotein (AFP) response as a potential surrogate biomarker of prognosis for the combination therapy.

Experimental Design: Data from Group A of the phase Ib GO30140 study were used to identify the optimal time for AFP measurement and AFP cutoffs to differentiate patients by their best confirmed response per independent review facility–assessed RECIST (IRF-RECIST) version 1.1: responders from nonresponders and patients with disease control from primary progressors. We applied these cutoffs to independent data from the atezolizumab + bevacizumab arm of the phase III IMbrave150 trial to distinguish patients based on (i) overall survival (OS) and progression-free survival (PFS) per IRF-RECIST 1.1 and (ii) best confirmed response per IRF-RECIST 1.1.

Results: We derived AFP cutoffs of ≥75% decrease and ≤10% increase from baseline at 6 weeks to identify responders and those who had disease control, respectively. These cutoffs had high sensitivity and specificity in GO30140. In IMbrave150 patients, sensitivity was 0.59 and specificity was 0.86 for the ≥75% decrease AFP cutoff; the sensitivity was 0.77 and specificity was 0.44 for the ≤10% increase AFP cutoff. Both AFP cutoffs were associated with longer OS and PFS, particularly in patients with hepatitis B virus etiology (HR < 0.5; $P < 0.01$).

Conclusions: AFP response at 6 weeks after initiating treatment is a potential surrogate biomarker of prognosis for patients with HCC receiving atezolizumab + bevacizumab.

Introduction

Alpha-fetoprotein (AFP) is the most widely tested serum biomarker in hepatocellular carcinoma (HCC) and overexpression of AFP is considered reflective of more aggressive tumor biology and burden (1). Approximately 40% of patients with unresectable HCC have baseline AFP ≥ 400 ng/mL (2). In patients with HCC, AFP ≥ 400 ng/mL is a negative prognostic factor for overall survival (OS) regardless of tumor stage and has become a stratification factor for most phase III clinical trials in unresectable HCC (1). Beyond the prognostic impact of baseline AFP levels, a variety of studies suggest that AFP changes on treatment or after treatment, across all disease stages, are surrogate biomarkers for response to both systemic treatment and locoregional treatment (3, 4). Currently, there is no optimal surrogate endpoint for OS in HCC (5). Objective response rate and progression-free survival (PFS) have been shown to have low and moderate correlations to OS, respectively, and challenges exist in the interpretation of radiologic response (4, 6). Thus, new early response surrogate biomarkers in HCC, such as on-treatment AFP response, are needed (5, 6).

Many studies have evaluated AFP response following systemic therapy using differing criteria (7–12). A 2019 meta-analysis of 11 studies, which evaluated AFP response to systemic treatment (chemotherapy, antiangiogenic therapy, or sorafenib) in 1,037 patients, identified AFP response to be associated with longer OS (HR, 0.33; $P < 0.001$; refs. 3, 13–23). More recently, data from trials on investigational therapies have emerged. Early AFP response in patients with pretreatment AFP levels of >20 ng/mL was associated with higher treatment efficacy of immune checkpoint inhibitors (programmed cell death protein 1 inhibitors and/or anti-CTL–associated antigen 4) in patients with advanced HCC (7). Compared with nonresponders, patients with an early AFP response, defined as >20% decline in serum AFP levels within the first 4 weeks of treatment initiation relative to

[1]Jiahui International Cancer Center, Jiahui International Hospital, Shanghai, P.R. China. [2]Massachusetts General Hospital Cancer Center, Harvard Medical School, Boston, Massachusetts. [3]Chao Family Comprehensive Cancer Center, University of California, Irvine, Irvine, California. [4]Department of Oncology, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan City, Taiwan. [5]Department of Hepatic Oncology, Liver Cancer Institute, Zhongshan Hospital, Fudan University, Shanghai, P.R. China. [6]Department of Gastrointestinal Oncology, Harbin Medical University Cancer Hospital, Harbin, P.R. China. [7]Department of Integrative Oncology, Fudan University Shanghai Cancer Center, Shanghai, P.R. China. [8]Department of Medical Oncology, Sir Run Run Shaw Hospital, College of Medicine, Zhejiang University, Hangzhou, P.R. China. [9]Personalized Health Care Real World Data Oncology, F. Hoffmann-La Roche Ltd., Basel, Switzerland. [10]Personalized Health Care Real World Data Oncology, Genentech, Inc., South San Francisco, California. [11]Global Product Development Medical Affairs, F. Hoffmann-La Roche Ltd., Basel, Switzerland. [12]US Medical Affairs, Genentech, Inc., South San Francisco, California. [13]Department of Medicine (Hematology/Oncology), University of California San Francisco, San Francisco, California. [14]Liver Unit, Clinica Universidad de Navarra-IDISNA and CIBEREHD, Pamplona, Spain.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Trial registration numbers: NCT02715531 (GO30140) and NCT03434379 (IMbrave150).

Corresponding Author: Andrew X. Zhu, Jiahui International Cancer Center, Jiahui International Hospital, Shanghai 200233, P.R. China. Phone: 86 (21) 5339 3217; E-mail: andrew.zhu@jiahui.com

Clin Cancer Res 2022;XX:XX–XX

doi: 10.1158/1078-0432.CCR-21-3275

Translational Relevance

Clinicians face difficulty in predicting the efficacy of hepatocellular carcinoma (HCC) treatments as there is no optimal surrogate endpoint for overall survival. Furthermore, challenges exist in the interpretation of radiologic response. Alpha-fetoprotein (AFP) is the most widely tested serum biomarker in HCC and is considered reflective of more aggressive tumor biology and burden. With atezolizumab + bevacizumab being the standard of care for patients with unresectable HCC, we explored AFP response as a potential surrogate biomarker of prognosis for the combination therapy in those with elevated baseline AFP. AFP cutoffs of ≥75% decrease or ≤10% increase measured 6 weeks after starting treatment were associated with longer overall and progression-free survival and displayed potential as surrogate biomarkers of response. Results from this exploratory analysis may be clinically useful for predicting the efficacy of atezolizumab + bevacizumab in patients with systemic treatment-naïve unresectable HCC early in treatment, helping to inform patient care.

pretreatment levels, showed longer OS [adjusted HR (aHR), 0.089; $P = 0.003$] and PFS (aHR, 0.128; $P < 0.001$; ref. 7). The phase III CELESTIAL study recently showed that the AFP control (defined as no increase from baseline levels) at 8 weeks was the best predictor of longer OS (HR, 0.50; $P < 0.001$) and longer PFS (HR, 0.48; $P < 0.001$) in patients with advanced HCC who were treated with cabozantinib (8). In the phase III trials REACH and REACH-2, AFP response was associated with a longer OS (HR, 0.45; $P < 0.001$) in patients treated with ramucirumab and was established as a predictive biomarker for therapy with ramucirumab (24). Furthermore, recent real-world data have been published, supporting AFP response as a potential surrogate biomarker for systemic treatment efficacy (9, 10).

However, there are limitations to these studies. The definition of AFP response varied across studies, with >/≥20% and >/≥50% reduction from baseline AFP level being the predominant cutoffs chosen. Furthermore, the timepoint at which AFP response was determined ranged from 1 week to 3 months following treatment initiation, with 1 month being the most common. It is also noteworthy that in most of these studies, the definition for AFP response was chosen arbitrarily (3, 4). In light of these limitations and the changing treatment landscape, there is a need to further evaluate the utility of AFP as a potential surrogate biomarker with newer methods.

Recent advances in the HCC field include the global approval of atezolizumab + bevacizumab for the treatment of unresectable HCC that has not been treated with prior systemic therapy (2). Approvals are based on the results from the phase III IMbrave150 study, in which atezolizumab + bevacizumab demonstrated statistically significant and clinically meaningful improvements over sorafenib in coprimary endpoints of OS and PFS per independent review facility–assessed RECIST (IRF-RECIST) version 1.1 (2). Compared with standard-of-care sorafenib, this combination resulted in a 42% decrease in the risk of death and a 41% decrease in the risk of disease progression or death, along with a 2.5-month improvement in median PFS at the primary analysis. Responses were observed regardless of baseline AFP levels. The treatment benefit with atezolizumab + bevacizumab versus sorafenib was maintained with an additional 12 months of follow-up (median: 15.6 months; ref. 25). These phase III data were consistent with results from the phase Ib GO30140 study that demonstrated promising antitumor activity for atezolizumab + bevacizumab (26).

We explored the use of AFP response as a potential surrogate biomarker for the efficacy of the combination of atezolizumab + bevacizumab and to inform patient care. After considering the constraints of earlier studies (i.e., limited rationale for the choices of AFP cutoff and timepoint at which AFP response was determined), we chose a two-step approach for this exploratory analysis. First, we analyzed data obtained from atezolizumab + bevacizumab–treated patients in Group A of the phase I GO30140 study to provide supportive rationale for an appropriate AFP response cutoff and timepoint for determining AFP response. We then validated the prognostic value of AFP response in patients receiving atezolizumab + bevacizumab in the phase III IMbrave150 study using the derived, optimized parameters from GO30140.

Materials and Methods

Patients

The details of both the GO30140 (NCT02715531) and IMbrave150 (NCT03434379) studies have been published previously (**Fig. 1A**; refs. 2, 26). Both of these studies were conducted in accordance with the International Conference on Harmonization guidelines for Good Clinical Practice and the principles of the Declaration of Helsinki. Protocol approval was obtained from the institutional review board or ethics committee at each site. All patients gave written informed consent. Patients from these studies were included in this exploratory analysis if they were treated with atezolizumab (1,200 mg) + bevacizumab (15 mg/kg) intravenously every 3 weeks, had an exact AFP result available at baseline (patients with AFP results expressed as "less than" or "greater than" certain thresholds were excluded), had baseline AFP levels ≥20 ng/mL, and had an exact AFP result available at 3-, 6-, and/or 9-week time windows. In addition, patients from GO30140 had to have available data on response per IRF-RECIST 1.1.

Evaluation of AFP response parameters

GO30140 data were used as the training cohort as they were available prior to IMbrave150 data. At each time window, AFP measurements were taken within 7 days before or after weeks 3, 6, and 9, respectively. Data from patients enrolled in Group A of GO30140 were used to identify the optimal time for AFP measurement and optimal AFP cutoffs to differentiate patients by their best confirmed response per IRF-RECIST 1.1, that is, differentiating (i) patients who achieved complete or partial response (CR + PR; responders) from those with stable or progressive disease (SD + PD; nonresponders) and (ii) patients who achieved CR, PR, or SD (disease control) from those with primary PD (primary progressors; **Fig. 1B**). The responder group is a subset of the disease control group. AFP response was considered achieved if the change in the patient's AFP level at the optimal time from baseline was within the range determined by the optimal cutoff value.

Subsequently, the optimal cutoffs and time window were validated using updated efficacy data from patients enrolled in the atezolizumab + bevacizumab arm of IMbrave150 (**Fig. 1B**; ref. 25). Because AFP response at each timepoint was determined on the basis of a 14-day window, patients from the IMbrave150 study were excluded from OS and PFS analyses if they experienced relevant events (i.e., death and death or progression, respectively) prior to the end of the exposure assessment period (i.e., landmark analysis). Best response per IRF-RECIST 1.1 and modified RECIST for HCC (HCC mRECIST) were used to determine the sensitivity, specificity, positive predictive value, and negative predictive value of the AFP cutoffs in differentiating responders versus nonresponders and disease control versus

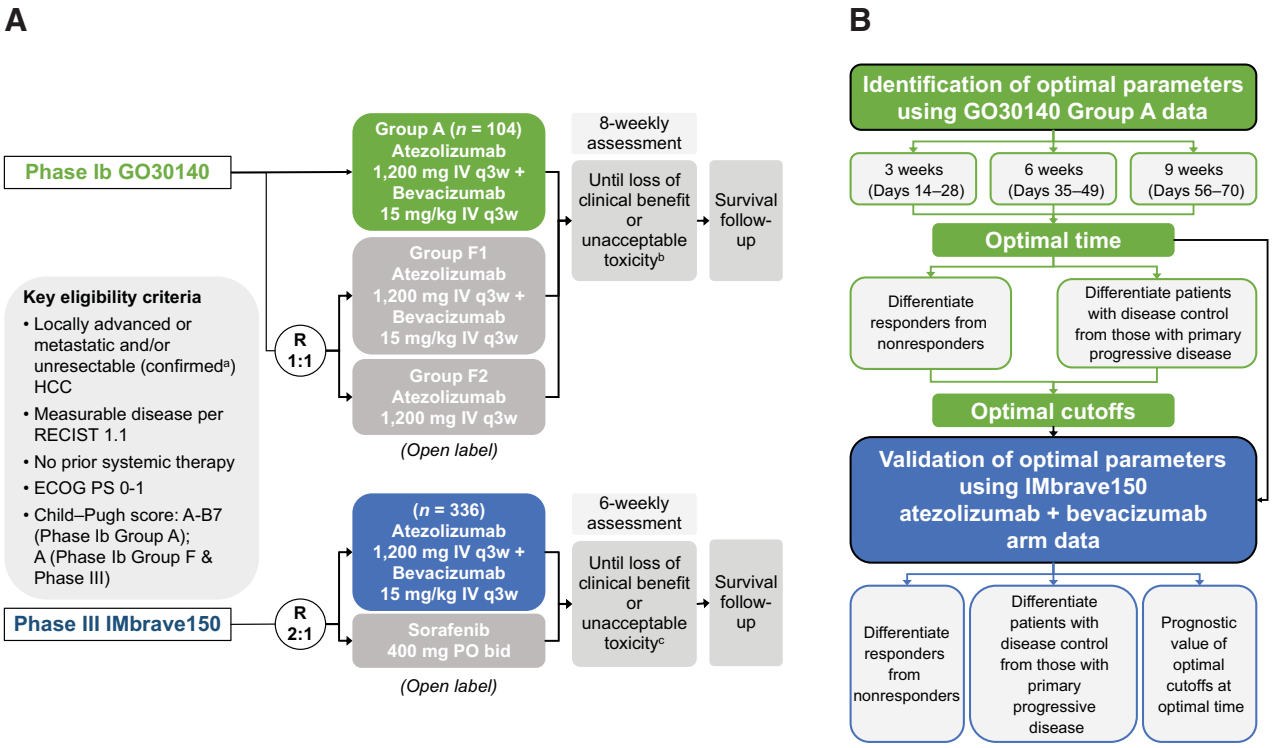

Figure 1.

Study designs. **A,** GO30140 and IMbrave150. **B,** Evaluation of AFP response. [a]Per American Association for the Study of Liver Diseases criteria. [b]Tumor assessments were done every 8 weeks for the first year and every 12 weeks thereafter until patient death, disease progression, or initiation of further systemic anticancer therapy. [c]Tumor assessments by CT or MRI were done at baseline and every 6 weeks until 54 weeks, and then every 9 weeks thereafter. bid, twice daily; ECOG PS, Eastern Cooperative Oncology Group performance status; IV, intravenously; PO, orally; q3w, every 3 weeks; R, randomization; RECIST, Response Evaluation Criteria in Solid Tumors version 1.1.

primary progressors (as defined above). AFP cutoffs were also evaluated on their ability to distinguish patients based on (i) OS (defined as the time from randomization to death from any cause) and PFS (defined as the time from randomization to disease progression per IRF-RECIST 1.1 or death from any cause, whichever occurred first) in the intention-to-treat population and on (ii) OS and PFS by etiology subgroup [hepatitis B virus (HBV), hepatitis C virus (HCV), and nonviral].

Statistical analyses

Descriptive statistics were used for summarizing the baseline characteristics of patients (i) included versus excluded from analyses, (ii) stratified by AFP cutoffs, or (iii) with AFP <20 ng/mL versus AFP ≥20 ng/mL. The ROC curve was used to determine the optimal time window and cutoffs from the GO30140 data (27). The optimal time window for AFP measurement was determined on the basis of the area under the ROC curve in differentiating patients by their best confirmed response per RECIST 1.1; the larger the area under the ROC curve at each time window for AFP measurement, the better the model can distinguish patients between the responder subgroups (27). AFP cutoffs were chosen at the optimal time window by identifying the points on the ROC curve that maximized the sum of sensitivity and specificity in differentiating patients by best confirmed response per RECIST 1.1 (27). Exact binomial confidence limits were calculated for test sensitivity, specificity, and positive and negative predictive values (28).

HRs, 95% confidence intervals (CI), and *P* values for OS and PFS were estimated with a Cox proportional hazards model using the exact ties method, which is appropriate for small sample sizes with discrete survival times that may involve tied event times (29). Landmark analysis was used to account for events that occurred before the outcome assessment window, avoiding immortality bias in survival analyses. Models were adjusted for age, baseline AFP, and etiology. *P* values are presented for descriptive purposes only. All analyses were conducted using R Core Team (2020) (R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria).

Data availability

Qualified researchers may request access to individual patient-level data through the clinical study data request platform (https://vivli.org). Further details on Roche criteria for eligible studies are available at https://vivli.org/members/ourmembers. For further details on Roche Global Policy on the Sharing of Clinical Information and how to request access to related clinical study documents, visit https://www.roche.com/research_and_development/who_we_are_how_we_work/clinical_trials/our_commitment_to_data_sharing.htm.

Results

Definition of AFP response parameters based on analyses of data from patients in Group A of GO30140

Of the 104 patients from Group A of GO30140 receiving atezolizumab + bevacizumab, 58 with baseline AFP ≥20 ng/mL were included in this exploratory analysis (Supplementary Fig. S1A). AFP responses at the 6- and 9-week time windows were statistically similar

in identifying responders and patients with disease control (ROC outputs are shown in Supplementary Table S1). The 6-week time window was chosen as it can provide an earlier indication of response to treatment. Responders were best differentiated from nonresponders when AFP decreased by ≥76.1% from baseline. Patients who achieved disease control were best distinguished from primary progressors when AFP increased by ≤9.4% from baseline. For ease of use, AFP cutoffs for the groups were rounded: (i) a ≥75% decrease to identify responders and (ii) a ≤10% increase to identify disease control.

The sensitivity and specificity of the ≥75% decrease AFP cutoff for differentiating responders from nonresponders in patients from Group A of GO30140 were 0.71 and 0.91, respectively. Using the ≤10% increase AFP cutoff, sensitivity was 0.89 and specificity was 1.00 for distinguishing disease control versus primary progressors.

Characteristics of IMbrave150 patients based on defined cutoffs

Of the 329 patients receiving atezolizumab + bevacizumab in IMbrave150, 150 were included in this exploratory analysis (Supplementary Fig. S1B); 147 of these 150 patients received two cycles in the first 6 weeks. Because the time window for AFP measurement at 6 weeks was from day 35 to 49, patients from the IMbrave150 study were excluded if they experienced progression or survival events prior to the end of this exposure assessment period at the day 49 posttreatment initiation (landmark date).

The likelihood of patients being excluded from the analysis was examined to evaluate the generalizability of the cohort included in this study. Patients were more likely to be excluded from this analysis if they were of White versus Asian descent or if they received versus did not receive prior therapy (local, radiotherapy, or surgery; $P < 0.05$; Supplementary Table S2). Patients were more likely ($P < 0.05$) to be excluded because of having baseline AFP levels <20 ng/mL if they were older versus younger, had nonviral versus viral HCC etiology, were Barcelona Clinic Liver Cancer (BCLC) stage A1/A4/B versus stage C, or received prior therapy versus did not receive prior therapy ($P < 0.05$; Supplementary Table S3).

Baseline characteristics of these patients are shown in **Table 1** based on the ≥75% decrease and ≤10% increase AFP cutoffs. Overall, patients had a median age of 62.0 years and were mostly males (81.3%) of Asian descent (62.7%). The majority of patients were classified as Child–Pugh class A5 (70.7%) and BCLC stage C (85.3%), had Eastern Cooperative Oncology Group performance score (ECOG PS) 0 (67.3%) and HBV etiology (54.0%), and had received prior locoregional therapy (63.3%).

Table 1. Baseline characteristics stratified by AFP cutoffs.

	≥75% decrease in AFP		≤10% increase in AFP		All patients included in analysis
	Yes (*n* = 39)	No (*n* = 111)	Yes (*n* = 107)	No (*n* = 43)	(*N* = 150)
Median age (IQR), years	61.0 (53.0–70.5)	63.0 (54.5–69.0)	61.0 (53.0–68.5)	65.0 (60.5–72.5)	62.0 (54.0–70.0)
Sex, *n* (%)					
Female	5 (12.8)	23 (20.7)	16 (15.0)	12 (27.9)	28 (18.7)
Male	34 (87.2)	88 (79.3)	91 (85.0)	31 (72.1)	122 (81.3)
Race, *n* (%)					
Asian	23 (59.0)	71 (64.0)	64 (59.8)	30 (69.8)	94 (62.7)
White	14 (35.9)	28 (25.2)	33 (30.8)	9 (20.9)	42 (28.0)
Other	0 (0.0)	4 (3.6)	4 (3.7)	0	4 (2.7)
Unknown	2 (5.1)	8 (7.2)	6 (5.6)	4 (9.3)	10 (6.7)
Median baseline AFP (IQR), ng/mL	2,907.1 (505.0–14,821.1)	615.7 (75.3–7,441.0)	1,485.8 (162.6–8,605.2)	1,193.0 (75.3–9,112.0)	1,354.5 (121.0–8,735.7)
ECOG PS, *n* (%)					
0	26 (66.7)	75 (67.6)	72 (67.3)	29 (67.4)	101 (67.3)
1	13 (33.3)	36 (32.4)	35 (32.7)	14 (32.6)	49 (32.7)
Child–Pugh class, *n* (%)					
A5	31 (79.5)	75 (67.6)	76 (71.0)	30 (69.8)	106 (70.7)
A6	8 (20.5)	36 (32.4)	31 (29.0)	13 (30.2)	44 (29.3)
Etiology, *n* (%)					
HBV	18 (46.2)	63 (56.8)	58 (54.2)	23 (53.5)	81 (54.0)
HCV	14 (35.9)	20 (18.0)	27 (25.2)	7 (16.3)	34 (22.7)
Nonviral	7 (17.9)	28 (25.2)	22 (20.6)	13 (30.2)	35 (23.3)
BCLC stage, *n* (%)					
A1/A4	0 (0.0)	5 (4.5)	4 (3.7)	1 (2.3)	5 (3.3)
B	3 (7.7)	14 (12.6)	10 (9.3)	7 (16.3)	17 (11.3)
C	36 (92.3)	92 (82.9)	93 (86.9)	35 (81.4)	128 (85.3)
Extrahepatic spread, *n* (%)	28 (71.8)	73 (65.8)	72 (67.3)	29 (67.4)	101 (67.3)
Median time from initial diagnosis (IQR), months	4.7 (1.7–9.4)	6.3 (2.1–18.2)	4.7 (1.8–11.6)	9.2 (2.1–25.2)	6.0 (1.9–15.4)
Prior locoregional therapy,[a] *n* (%)	22 (56.4)	73 (65.8)	63 (58.9)	32 (74.4)	95 (63.3)
Median number of metastatic sites (IQR), *n*	0.0 (0.0–1.0)	1.0 (0.0–1.0)	1.0 (0.0–1.0)	1.0 (0.0–1.0)	1.0 (0.0–1.0)

Abbreviations: AFP, alpha-fetoprotein; BCLC, Barcelona Clinic Liver Cancer; ECOG PS, Eastern Cooperative Oncology Group performance score; IQR, interquartile range.

[a]Includes local therapy, radiotherapy, and surgery.

Table 2. Sensitivity and specificity of AFP cutoffs in differentiating patients by response status per IRF-assessed RECIST 1.1 and HCC mRECIST.

	RECIST 1.1 ($n = 144$)		HCC mRECIST ($n = 144$)	
Point estimate (95% CI)	≥75% AFP decrease (responders vs. nonresponders)	≤10% AFP increase (disease control vs. primary progressors)	≥75% AFP decrease (responders vs. nonresponders)	≤10% AFP increase (disease control vs. primary progressors)
Sensitivity	0.59 (0.42–0.74)	0.77 (0.68–0.85)	0.56 (0.41–0.70)	0.77 (0.68–0.85)
Specificity	0.86 (0.78–0.92)	0.44 (0.27–0.62)	0.89 (0.81–0.95)	0.43 (0.26–0.61)
Positive predictive value	0.61 (0.43–0.76)	0.82 (0.73–0.89)	0.74 (0.57–0.87)	0.81 (0.72–0.88)
Negative predictive value	0.85 (0.77–0.91)	0.38 (0.23–0.54)	0.79 (0.70–0.87)	0.38 (0.23–0.54)
Area under the curve	0.78	0.66	0.75	0.67

Abbreviations: AFP, alpha-fetoprotein; HCC mRECIST, HCC-modified RECIST; IRF, independent review facility; RECIST 1.1, Response Evaluation Criteria in Solid Tumors version 1.1.

Differences in baseline characteristics between responders and nonresponders at each AFP cutoff were observed. Compared with patients who had a <75% decrease in AFP, more patients with a ≥75% decrease in AFP had higher AFP at baseline (median: 2,907 vs. 616 ng/mL; $P = 0.006$) or had viral etiology (82.1% vs. 74.8%, $P = 0.07$). In addition, compared with patients who had a >10% increase in AFP, patients with a ≤10% increase in AFP were younger (median: 61.0 vs. 65.0 years old; $P = 0.039$). Consequently, the Cox model was adjusted for age, baseline AFP, and etiology.

Association between AFP cutoffs and patient outcomes in IMbrave150

In patients from IMbrave150, sensitivity and specificity of the ≥75% decrease AFP cutoff for differentiating responders from nonresponders were 0.59 (95% CI: 0.42–0.74) and 0.86 (95% CI: 0.78–0.92), respectively (**Table 2**). The sensitivity was 0.77 (95% CI: 0.68–0.85) and specificity was 0.44 (95% CI: 0.27–0.62) when the ≤10% increase AFP cutoff was used to distinguish patients who achieved disease control versus primary progressors. As shown in **Fig. 2**, the majority of patients who had a ≥75% decrease in AFP achieved objective response, and most of those with a ≤10% increase in AFP had disease control. The HCC mRECIST criteria generated similar sensitivity and specificity estimates for both cutoffs in patients from IMbrave150. Positive and negative predictive values of each cutoff are also presented in **Table 2**. In light of the lower performance in the validation cohort, particularly for the ≤10% increase AFP cutoff to distinguish patients who achieved disease control versus primary progressors, an additional exploratory analysis was performed to evaluate the test characteristics for the commonly used cutoffs of ≥20% decrease AFP and ≥50% decrease AFP for the IMbrave150 cohort (Supplementary Table S4). On the basis of the sum of the sensitivity and specificity, the ≥75% decrease remains the optimal cutoff for distinguishing responders from nonresponders, while for differentiating disease control versus primary progressors, a ≥20% decrease appears to perform better.

Both AFP cutoffs were associated with OS and PFS in the overall population (**Fig. 3**). Patients in IMbrave150 whose AFP levels decreased by ≥75% from baseline at 6 weeks had not reached median OS at the time of data cutoff, whereas patients whose AFP levels decreased by <75% had a median OS of 14.2 months (aHR, 0.36; 95% CI: 0.20–0.66; $P < 0.001$; **Fig. 3A**). For patients whose AFP levels had increased by ≤10% versus >10%, median OS was 23.7 and 10.6 months, respectively (aHR, 0.45; 95% CI: 0.29–0.70; $P < 0.001$; **Fig. 3B**).

Median PFS in patients with a ≥75% decrease in AFP was 13.2 months compared with 6.7 months in patients with a <75% decrease in AFP (aHR, 0.46; 95% CI: 0.28–0.77; $P < 0.001$; **Fig. 3C**). For patients with a ≤10% versus >10% increase in AFP, median PFS was 9.9 and 5.5 months, respectively (aHR, 0.40; 95% CI: 0.24–0.66; $P < 0.001$; **Fig. 3D**).

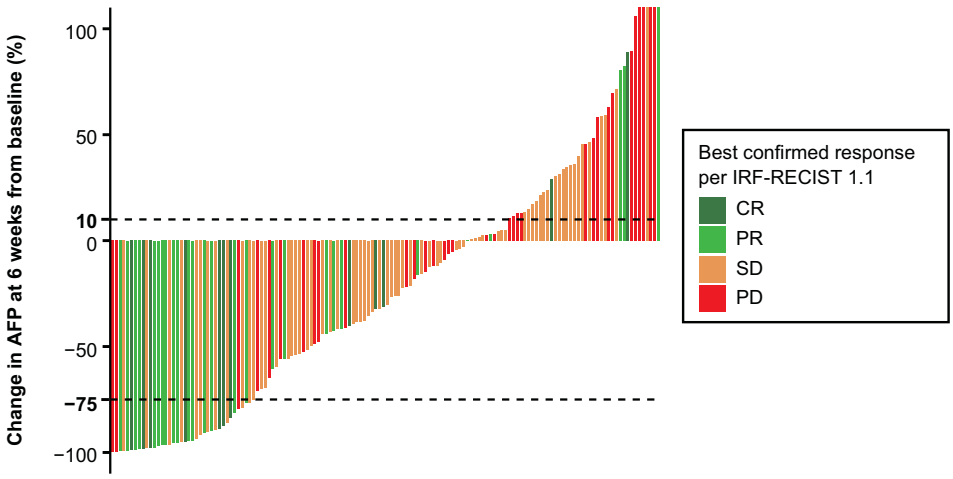

Figure 2.
Waterfall plot of change in AFP at 6 weeks from baseline and best confirmed response per IRF-RECIST 1.1. AFP, alpha-fetoprotein; CR, complete response; IRF-RECIST, independent review facility–assessed Response Evaluation Criteria in Solid Tumors version 1.1; PD, progressive disease; PR, partial response; SD, stable disease.

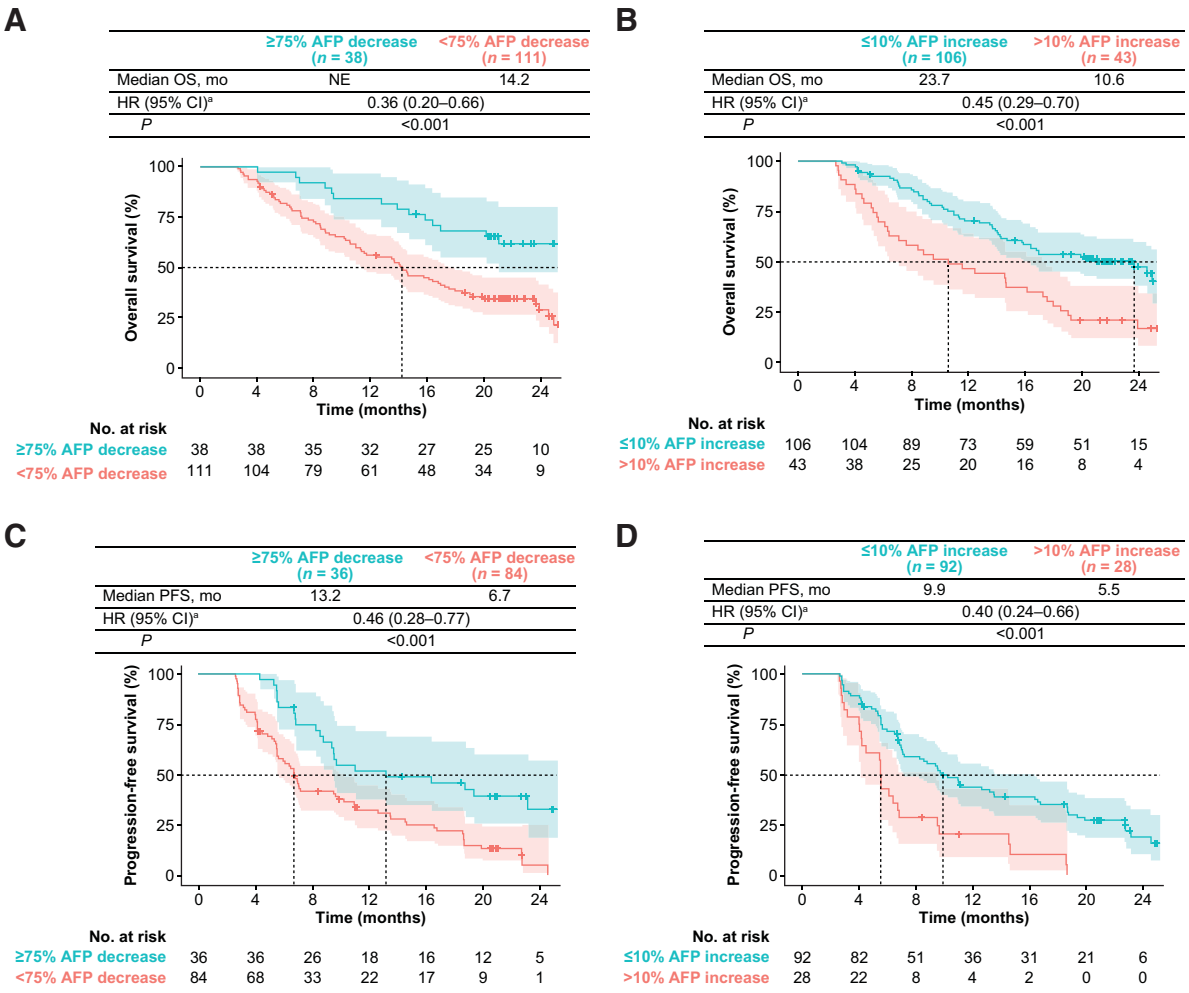

Figure 3.

OS of patients in IMbrave150 treated with atezolizumab + bevacizumab stratified by AFP cutoffs at 6 weeks. **A,** ≥75% AFP decrease. **B,** ≤10% AFP increase. PFS of patients in IMbrave150 treated with atezolizumab + bevacizumab stratified by AFP cutoffs at 6 weeks. **C,** ≥75% AFP decrease. **D,** ≤10% AFP increase. [a]Estimated using a Cox proportional hazards model adjusted for baseline AFP, age, and viral etiology. AFP, alpha-fetoprotein; HR, hazard ratio; mo, months; NE, not estimable; OS, overall survival; PFS, progression-free survival.

Association between AFP cutoffs and improvement in survival based on etiology

In patients with HBV etiology, longer OS was seen in patients with a ≥75% decrease (aHR, 0.22; 95% CI: 0.07–0.68; $P = 0.008$) and ≤10% increase in AFP (aHR, 0.38; 95% CI: 0.21–0.69; $P = 0.002$) versus the rest of this subgroup (patients with a <75% decrease and >10% increase in AFP, respectively). Longer PFS was also observed in patients with a ≥75% decrease (aHR, 0.46; 95% CI: 0.22–0.95; $P = 0.037$) and ≤10% increase in AFP (aHR, 0.38; 95% CI: 0.18–0.81; $P = 0.012$; **Table 3**) compared with the rest of this subgroup. In patients with HCV or nonviral etiologies, neither cutoff was associated with longer OS or PFS ($P > 0.05$).

Discussion

Using data from patients in the single-arm cohort of GO30140, we derived AFP cutoffs of a ≥75% decrease and a ≤10% increase from baseline at 6 weeks to identify responders and patients with disease control, respectively. These cutoffs were subsequently validated in patients receiving atezolizumab + bevacizumab from IMbrave150 with baseline AFP levels of ≥20 ng/mL, showing the potential for AFP as a surrogate and early biomarker of response. Compared with nonresponders, more patients who achieved AFP response based on either of the cutoffs were younger, had viral etiology, and had higher AFP at baseline. In addition, both AFP cutoffs were associated with longer OS and PFS, particularly in patients with HBV etiology (aHR <0.5; $P < 0.01$). Notably, the proportion of patients with HBV at baseline was higher among AFP nonresponders than responders, which was also reported in another AFP study (24). This association between HBV etiology and AFP response suggests that AFP cutoffs may be more selective in patients with baseline HBV compared with patients with HCV or nonviral etiology, potentially explaining the longer OS and PFS in AFP responders with HBV. Nonetheless, a trend for longer OS and PFS was also observed in the HCV and nonviral subgroups but further analysis in a larger sample of these subgroups would be required to confirm such conclusions. Results from this exploratory analysis can be clinically relevant to assess the prognosis of patients receiving the atezolizumab + bevacizumab combination at

Table 3. OS and PFS per AFP cutoffs and HCC etiology.

	OS HR (95% CI); P[a]			PFS per IRF-assessed RECIST 1.1 HR (95% CI); P[a]		
Etiology	***n***	**≥75% AFP decrease**	**≤10% AFP increase**	***n***	**≥75% AFP decrease**	**≤10% AFP increase**
HBV	80	0.22 (0.07–0.68) $P = 0.008$	0.38 (0.21–0.69) $P = 0.002$	61	0.46 (0.22–0.95) $P = 0.037$	0.38 (0.18–0.81) $P = 0.012$
HCV	34	0.56 (0.19–1.63) $P = 0.288$	0.34 (0.10–1.15) $P = 0.083$	29	0.44 (0.15–1.30) $P = 0.139$	0.29 (0.09–0.91) $P = 0.034$
Nonviral	35	0.45 (0.14–1.43) $P = 0.174$	0.74 (0.33–1.66) $P = 0.471$	30	0.49 (0.15–1.64) $P = 0.248$	0.52 (0.21–1.31) $P = 0.164$

Abbreviations: AFP, alpha-fetoprotein; IRF, independent review facility; OS, overall survival; PFS, progression-free survival; RECIST 1.1, Response Evaluation Criteria in Solid Tumors version 1.1.

[a]Estimated using a Cox proportional hazards model adjusted for baseline AFP and age.

6 weeks after starting treatment. In particular, the threshold for a >10% increase in AFP levels may be most useful for identifying, carefully monitoring, and managing patients who may be less likely to achieve disease control and/or experience prolonged OS from treatment with atezolizumab + bevacizumab, keeping in mind that a patient showing an early AFP increase may still experience a radiologic response at a later timepoint. Of note, 5 of the 43 patients who had a >10% increase in AFP at 6 weeks showed an objective response (PR or CR) at some point during the treatment. Conversely, an early AFP drop of ≥75% may increase both the treating physician's and their patient's confidence in the treatment choice.

The main strength of this study is the two-step study design, which makes our work the first to provide supportive rationale for the chosen AFP response cutoffs and time window and the first to validate the prognostic value of AFP response in a larger, external sample of patients. Of note, none of the previously reported cutoffs emerged from our analysis of GO30140 as being the optimal thresholds. Only one other study has used ROC curve analyses to identify an optimal cutoff for AFP response. The cutoff derived was a ≥40% decrease in AFP levels 1 month after lenvatinib initiation, with sensitivity and specificity levels of 1.00 and 0.78, respectively, in distinguishing patients with disease control from primary progressors (9). Although the performance of this cutoff was comparable with the cutoff derived in our exploratory analysis, which was a ≤10% increase in AFP at 6 weeks with a sensitivity of 0.89 and specificity of 1.00 in identifying disease control, this study did not provide justification for the chosen time window nor test their cutoff in an independent validation cohort (9). In the absence of a validation cohort test, it is difficult to directly compare these findings to those reported in our study. Moreover, to the authors' knowledge, no other study has explicitly reported the test characteristics of their chosen cutoff, preventing further comparisons on the performance of our model. One study conducted an exploratory analysis using maximally selected rank statistics to identify an optimal AFP cutoff that provided the strongest association with OS (8). However, this cutoff was not tested in a validation cohort and maximally selected rank statistics cannot be compared directly with our approach. The only other study that included an exploratory and validation cohort applied an arbitrarily defined AFP cutoff and time window rather than leveraging an exploratory cohort to identify an optimal AFP response definition (22). Furthermore, strengths of this study include the use of a landmark analysis which explicitly excluded patients who experienced progression or survival outcomes prior to AFP response classification, thereby avoiding immortality bias in survival analyses.

A limitation of this analysis is that the results may be less relevant for patient populations who were not part of the study cohort, for example, patients who were Child–Pugh class B or C; with AFP levels <20 ng/mL; and who experienced progression or survival outcomes prior to day 50 in the landmark analysis. The latter two subgroups were excluded from this analysis to improve response identification and reduce immortality bias, respectively. However, in the new era of personalized healthcare, our approach may be considered a strength by analyzing the outcomes of a biologically relevant subgroup. Furthermore, most studies of AFP response to systemic treatments to date have been limited to similar patient cohorts.

Another limitation is that the sensitivity and specificity of the AFP cutoffs in differentiating patients per RECIST 1.1 were lower in IMbrave150 than in GO30140. There was a decrease in specificity for the ≤10% cutoff from 1.00 in GO30140 to 0.44 in IMbrave150. Further exploratory analyses indicate that the 75% cutoff remains the optimal cutoff in distinguishing responders from nonresponders, while for differentiating disease control versus primary progressors a 20% decrease appears to perform better. One possible reason for the decrease in specificity for the ≤10% cutoff is that tumor assessments occurred at different frequencies in these two studies (2, 11). In GO30140, tumors were assessed every 8 weeks for the first year and every 12 weeks thereafter (11), whereas in IMbrave150, tumors were assessed every 6 weeks until week 54 and every 9 weeks thereafter (2). Considering that the AFP cutoffs were derived from GO30140 data, it is not entirely unexpected that the same cutoffs would not be equally sensitive and specific in IMbrave150. One possible way to increase the sensitivity and specificity of the AFP cutoffs in predicting treatment response in both datasets is to increase the baseline AFP threshold (<20 ng/mL) used in the exclusion criteria. By doing so, even more clinically meaningful changes in AFP levels will be needed to meet the AFP cutoffs, enabling a better prediction of response per RECIST 1.1. However, this will also result in a larger exclusion of patients with low AFP at baseline and would require careful consideration.

Of note, the ≥75% decrease and ≤10% increase AFP cutoffs were derived on the basis of tumor response data in GO30140 and then applied to survival data in IMbrave150. As different factors might impact tumor response versus PFS and OS, these cutoffs, while robust, may not be optimal for predicting longer PFS and/or OS. Sensitivity analyses including further key prognostic factors such as BCLC stage, liver function, extrahepatic spread, and ECOG PS, in multivariate models did not change conclusions (Supplementary Table S5). Future studies could consider using our approach to identify and test an optimal AFP response definition in exploratory and validation cohorts

but leverage maximally selected rank statistics to optimize the cutoffs with regard to OS or PFS. Alternatively, further investigation could combine AFP response and radiologic evaluation of response to predict OS and PFS, as well as compare the prognostic ability of the combined approach with radiologic evaluation alone.

Despite these limitations, this exploratory analysis provides supportive rationale for appropriate AFP response cutoffs and tested the prognostic value of AFP response, presenting clinicians with relevant data that can be used for informing therapy management with atezolizumab + bevacizumab. To the best of the authors' knowledge, this is the first study to evaluate AFP as a surrogate marker in the current standard of care, atezolizumab + bevacizumab, in first-line unresectable HCC. Furthermore, this is the also the first study to use two different datasets to identify and test the AFP cutoffs and assessment windows.

Conclusions

The use of AFP changes as a surrogate biomarker of response has long been of interest for the improvement of HCC patient management. With the approval of atezolizumab + bevacizumab as the new standard of care for patients with unresectable HCC, there are new possibilities to analyze and interpret such data. On the basis of data from IMbrave150 patients with baseline AFP levels of >20 ng/mL, a ≥75% decrease or ≤10% increase in AFP levels measured 6 weeks after starting treatment shows an association with longer OS and PFS. These AFP cutoffs also displayed potential as surrogate biomarkers of response to this combination. Results from this exploratory analysis may be clinically useful for predicting the efficacy of atezolizumab + bevacizumab in patients with systemic treatment-naïve unresectable HCC early in the course of treatment.

Authors' Disclosures

A.X. Zhu reports personal fees from I-mab, Lilly, Merck, Exelixis, Roche, Eisai, and Bayer outside the submitted work. F. Dayyani reports grants and personal fees from AstraZeneca and Exelixis; personal fees from Eisai, Genentech, and Sirtex; and grants from BMS and Merck outside the submitted work. In addition, F. Dayyani's spouse is employed by Roche Diagnostics. Z. Ren reports personal fees from AstraZeneca, Merck Sharp & Dohme, and Roche outside the submitted work. P. Dillon is an employee of F. Hoffmann-La Roche, the study sponsor. S.K. Mhatre reports other support from F. Hoffmann-La Roche AG outside the submitted work. V.E. Gaillard reports other support from F. Hoffmann-La Roche during the conduct of the study as well as other support from F. Hoffmann-La Roche outside the submitted work. S. Hernandez reports personal fees from Genentech during the conduct of the study as well as personal fees from Genentech outside the submitted work. In addition, S. Hernandez is an employee of Genentech. R.K. Kelley reports personal fees from Genentech/Roche during the conduct of the study as well as grants from Agios, AstraZeneca, Bayer, BMS, Eli Lilly, EMD Serono, Exelixis, Genentech/Roche, Merck, Novartis, QED, Relay Therapeutics, Surface Oncology, and LOXO Oncology and personal fees from Gilead, Exact Sciences, and Kinnate outside the submitted work. B. Sangro reports grants from BMS, Sirtex Medical, and Roche and personal fees from Adaptimmune, AstraZeneca, Bayer, BMS, Boston Scientific, Eisai, Incyte, Ipsen, Roche, Sirtex Medical, and Terumo outside the submitted work. No disclosures were reported by the other authors.

Authors' Contributions

A.X. Zhu: Conceptualization, investigation, writing–original draft, writing–review and editing, interpretation. **F. Dayyani:** Investigation, writing–review and editing. **C.-J. Yen:** Investigation, writing–review and editing. **Z. Ren:** Investigation, writing–review and editing. **Y. Bai:** Investigation, writing–review and editing. **Z. Meng:** Investigation, writing–review and editing. **H. Pan:** Investigation, writing–review and editing. **P. Dillon:** Data curation, software, formal analysis, validation, investigation, visualization, methodology, writing–original draft, writing–review and editing, interpretation. **S.K. Mhatre:** Investigation, writing–review and editing. **V.E. Gaillard:** Conceptualization, supervision, investigation, visualization, writing–original draft, writing–review and editing, interpretation. **S. Hernandez:** Investigation, writing–review and editing. **R.K. Kelley:** Investigation, writing–review and editing. **B. Sangro:** Conceptualization, investigation, writing–review and editing, interpretation.

Acknowledgments

This study was funded by F. Hoffmann-La Roche Ltd. The sponsor was involved in the study design; collection, analysis, and interpretation of data; writing of the report; and the decision to submit the article for publication. The authors would like to thank the patients who participated in the trial, the patients' families, and the investigators and staff at all clinical study sites. Third-party medical writing assistance for this article was provided by Bena Lim, PhD, of MediTech Media and funded by F. Hoffmann-La Roche Ltd.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received September 14, 2021; revised December 15, 2021; accepted April 13, 2022; published first April 18, 2022.

References

1. Galle PR, Foerster F, Kudo M, Chan SL, Llovet JM, Qin S, et al. Biology and significance of alpha-fetoprotein in hepatocellular carcinoma. Liver Int 2019;39:2214–29.
2. Finn RS, Qin S, Ikeda M, Galle PR, Ducreux M, Kim TY, et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. N Engl J Med 2020;382:1894–905.
3. He C, Peng W, Liu X, Li C, Li X, Wen TF. Post-treatment alpha-fetoprotein response predicts prognosis of patients with hepatocellular carcinoma: a meta-analysis. Medicine 2019;98:e16557.
4. Xu XS, Qu K, Liu C, Zhang YL, Liu J, Song YZ, et al. Highlights for α-fetoprotein in determining prognosis and treatment monitoring for hepatocellular carcinoma. World J Gastroenterol 2012;18:7242–50.
5. European Association for the Study of the Liver. EASL Clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 2018;69:182–236.
6. Llovet JM, Montal R, Villanueva A. Randomized trials and endpoints in advanced HCC: role of PFS as a surrogate of survival. J Hepatol 2019;70:1262–77.
7. Shao YY, Liu TH, Hsu C, Lu LC, Shen YC, Lin ZZ, et al. Early alpha-fetoprotein response associated with treatment efficacy of immune checkpoint inhibitors for advanced hepatocellular carcinoma. Liver Int 2019;39:2184–9.
8. Kelley RK, Meyer T, Rimassa L, Merle P, Park JW, Yau T, et al. Serum alpha-fetoprotein levels and clinical outcomes in the phase III CELESTIAL study of cabozantinib versus placebo in patients with advanced hepatocellular carcinoma. Clin Cancer Res 2020;26:4795–804.
9. Saeki I, Yamasaki T, Yamashita S, Hanazono T, Urata Y, Furutani T, et al. Early predictors of objective response in patients with hepatocellular carcinoma undergoing lenvatinib treatment. Cancers 2020;12:779.
10. Sun X, Mei J, Lin W, Yang Z, Peng W, Chen J, et al. Reductions in AFP and PIVKA-II can predict the efficiency of anti-PD-1 immunotherapy in HCC patients. BMC Cancer 2021;21:775.
11. Kuzuya T, Asahina Y, Tsuchiya K, Tanaka K, Suzuki Y, Hoshioka T, et al. Early decrease in α-fetoprotein, but not des-γ-carboxy prothrombin, predicts sorafenib efficacy in patients with advanced hepatocellular carcinoma. Oncology 2011;81:251–8.
12. Vora SR, Zheng H, Stadler ZK, Fuchs CS, Zhu AX. Serum alpha-fetoprotein response as a surrogate for clinical outcome in patients receiving systemic therapy for advanced hepatocellular carcinoma. Oncologist 2009;14:717–25.
13. 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;27:446–52.

14. Chen LT, Liu TW, Chao Y, Shiah HS, Chang JY, Juang SH, et al. Alpha-fetoprotein response predicts survival benefits of thalidomide in advanced hepatocellular carcinoma. Aliment Pharmacol Ther 2005;22:217–26.
15. Kawaoka T, Aikata H, Murakami E, Nakahara T, Naeshiro N, Tanaka M, et al. Evaluation of the mRECIST and α-fetoprotein ratio for stratification of the prognosis of advanced-hepatocellular-carcinoma patients treated with sorafenib. Oncology 2012;83:192–200.
16. Kuzuya T, Ishigami M, Ishizu Y, Honda T, Hayashi K, Katano Y, et al. Early clinical response after 2 weeks of sorafenib therapy predicts outcomes and anti-tumor response in patients with advanced hepatocellular carcinoma. PLoS One 2015;10:e0138776.
17. Lee S, Kim BK, Kim SU, Park JY, Kim dY, Ahn SH, et al. Early α-fetoprotein response predicts survival in patients with advanced hepatocellular carcinoma treated with sorafenib. J Hepatocell Carcinoma 2015;2:39–47.
18. Nakazawa T, Hidaka H, Takada J, Okuwaki Y, Tanaka Y, Watanabe M, et al. Early increase in α-fetoprotein for predicting unfavorable clinical outcomes in patients with advanced hepatocellular carcinoma treated with sorafenib. Eur J Gastroenterol Hepatol 2013;25:683–9.
19. Personeni N, Bozzarelli S, Pressiani T, Rimassa L, Tronconi MC, Sclafani F, et al. Usefulness of alpha-fetoprotein response in patients treated with sorafenib for advanced hepatocellular carcinoma. J Hepatol 2012;57:101–7.
20. Shao YY, Lin ZZ, Hsu C, Shen YC, Hsu CH, Cheng AL. Early alpha-fetoprotein response predicts treatment efficacy of antiangiogenic systemic therapy in patients with advanced hepatocellular carcinoma. Cancer 2010;116:4590–6.
21. Sánchez AIP, Roces LV, García IZ, López EL, Hernandez MAC, Parejo MIB, et al. Value of α-fetoprotein as an early biomarker for treatment response to sorafenib therapy in advanced hepatocellular carcinoma. Oncol Lett 2018;15:8863–70.
22. Yau T, Yao TJ, Chan P, Wong H, Pang R, Fan ST, et al. The significance of early alpha-fetoprotein level changes in predicting clinical and survival benefits in advanced hepatocellular carcinoma patients receiving sorafenib. Oncologist 2011;16:1270–9.
23. Chou WC, Lee CL, Yang TS, Huang CY, Teng W, Tseng YT, et al. Changes in serum α-fetoprotein level predicts treatment response and survival in hepatocellular carcinoma patients and literature review. J Formos Med Assoc 2018;117:153–63.
24. Zhu AX, Finn RS, Kang YK, Yen CJ, Galle PR, Llovet JM, et al. Serum alpha-fetoprotein and clinical outcomes in patients with advanced hepatocellular carcinoma treated with ramucirumab. Br J Cancer 2021;124:1388–97.
25. Finn R, Qin S, Ikeda M, Galle P, Ducreux M, Kim T-Y, et al. IMbrave150: Updated overall survival (OS) data from a global, randomized, open-label phase III study of atezolizumab (atezo) + bevacizumab (bev) versus sorafenib (sor) in patients (pts) with unresectable hepatocellular carcinoma (HCC). J Clin Oncol 2021;39(suppl 3):abstr 267.
26. Lee MS, Ryoo B-Y, Hsu C-H, Numata K, Stein S, Verret W, et al. Atezolizumab with or without bevacizumab in unresectable hepatocellular carcinoma (GO30140): an open-label, multicentre, phase 1b study. Lancet Oncol 2020;21:808–20.
27. Bewick V, Cheek L, Ball J. Statistics review 13: receiver operating characteristic curves. Crit Care 2004;8:508–12.
28. Collett D. Modelling binary data. 1st ed. Boca Raton, FL: Chapman and Hall/CRC; 1999. p. 369.
29. Therneau TM, Lumley T, Atkinson E, Crowson C. Package 'Survival' Version 3.2-7. Available from: https://cran.r-project.org/web/packages/survival/survival.pdf.