

Which Criteria Applied in Multi-Phasic CT Can Predict Early Tumor Response in Patients with Hepatocellular Carcinoma Treated Using Conventional TACE: RECIST, mRECIST, EASL or qEASL?

Yan Zhao^{1,2} · Rafael Duran² · Wei Bai¹ · Sonia Sahu² · Wenjun Wang¹ · Sven Kabus³ · MingDe Lin^{2,4} · Guohong Han¹ · Jean-François Geschwind^{2,5}

Received: 29 July 2017 / Accepted: 19 October 2017 / Published online: 30 October 2017

© Springer Science+Business Media, LLC and the Cardiovascular and Interventional Radiological Society of Europe (CIRSE) 2017

Abstract

Purpose Our study aimed to evaluate quantitative tumor response assessment (quantitative EASL-[qEASL]) on computed tomography (CT) images in patients with hepatocellular carcinoma (HCC) treated using conventional transarterial chemoembolization (cTACE), compared to existing 1-dimensional and 2-dimensional methods (RECIST, mRECIST, EASL).

Materials and Methods In this IRB-approved, single-institution retrospective cohort study, 52 consecutive patients with intermediate-stage HCC were consecutively included. All patients underwent contrast-enhanced CT scan at baseline and 4 weeks after cTACE.

Results Median follow-up period was 13.5 months (range 1.2–54.1). RECIST, mRECIST and EASL identified progression in 2 (4%), 1 (2%) and 1 (2%) patients, respectively, whereas qEASL identified 10 (19%) patients. qEASL was the only tumor response method able to predict survival among different tumor response groups ($P < 0.05$), whereas

RECIST, mRECIST and EASL did not ($P > 0.05$). Both EASL and qEASL were able to identify responders and non-responders and were predictive of survival ($P < 0.05$). Multivariate analysis showed that progression was an independent predictor of overall survival with hazard ratio of 1.9 ($P = 0.025$). Patients who demonstrated progression with qEASL had significantly shorter survival than those with non-progression (7.6 vs. 20.4 months, $P = 0.012$). Similar multivariate analysis using RECIST, mRECIST and EASL could not be performed because too few patients were categorized as progressive disease.

Conclusion qEASL could be applied on CT images to assess tumor response following cTACE and is a more sensitive biomarker to predict survival and identify tumor progression than RECIST, mRECIST and EASL at an early time point.

Level of Evidence Level 2a, retrospective cohort study.

Keywords Hepatocellular carcinoma · Transarterial chemoembolization · Tumor response · Survival · Three dimensional

Yan Zhao and Rafael Duran contributed equally to this work.

✉ Guohong Han
hangh@fmmu.edu.cn

¹ Department of Liver Disease and Digestive Interventional Radiology, Xijing Hospital of Digestive Diseases, Fourth Military Medical University, No. 27 West Changle Road, Xi'an 710032, China

² Department of Radiology and Biomedical Imaging, Yale University School of Medicine, New Haven, CT, USA

³ Philips Research Hamburg, Hamburg, Germany

⁴ U/S Imaging and Interventions (UII), Philips Research North America, Cambridge, MS, USA

⁵ PreScience Labs LLC, Westport, CT, USA

Abbreviations

HCC	Hepatocellular carcinoma
TACE	Transarterial chemoembolization
cTACE	Conventional TACE
BCLC	Barcelona clinic liver cancer
RECIST	Response evaluation criteria in solid tumors
mRECIST	Modified RECIST
EASL	European Association for the Study of the Liver
MR	Magnetic resonance
CT	Computed tomography
ECOG	Eastern cooperative oncology group

Introduction

Hepatocellular carcinoma (HCC) is the second leading cause of cancer-related deaths worldwide [2]. Conventional transarterial chemoembolization (cTACE) using Lipiodol is the recommended first-line therapy for intermediate-stage patients in the barcelona clinic liver cancer (BCLC) staging classification, and in real clinical setting cTACE is the most widely used treatment for unresectable HCC [1].

Commonly used response assessment criteria/guidelines include the Response Evaluation Criteria in Solid Tumors (RECIST), modified RECIST (mRECIST) and the European Association for the Study of the Liver (EASL) criteria [4, 11, 18]. RECIST only considers anatomical tumor shrinkage and so does not capture the tumor necrosis induced by TACE [12, 23]. mRECIST and EASL criteria were then proposed to address the limitations of RECIST by measuring tumor enhancement, a surrogate marker for viable tumor [13]. However, RECIST, mRECIST and EASL criteria are 1-dimensional (D) and 2-D tumor assessment methods which may not be able to capture effectively tumor necrosis induced by catheter-based locoregional therapies such as cTACE. Indeed, tumors do not grow or shrink symmetrically and undergo heterogeneous changes which would affect the reliability of RECIST, mRECIST or EASL criteria [14, 28]. The newly proposed criterion, quantitative EASL (qEASL), considers the 3-D volumetric value of enhancing tumor tissue [20]. This quantitative tumor response assessment demonstrated a high radiological-pathological accuracy with high correlation of qEASL on post-interventional magnetic resonance (MR) images with TACE-induced HCC necrosis on histopathology [6]. Moreover, qEASL predicted survival better than RECIST, mRECIST and EASL criteria on MR images in HCC patients [28].

However, the application of qEASL in computed tomography (CT) images remains unknown. Despite the advantage of MR scans in higher sensitivity and specificity compared with CT, helical CT remains the more widely used technique because of the higher cost, lower availability and longer image acquisition time of MR [25]. In addition, Lipiodol staining of the lesions post-cTACE makes accurate assessment of tumor response very difficult. Indeed, the attenuation caused by Lipiodol on CT images may mask underlying remaining viable contrast-enhanced tumor tissue [4]. Under this circumstance, a quantitative method to measure tumor response after cTACE is of clinical importance. Subtraction between unenhanced and contrast-enhanced images is a unique characteristic of qEASL, which mitigates contribution of background attenuation from the Lipiodol deposition. Therefore, the purpose of our study was to evaluate the

performance of qEASL, compared with existing 1-D and 2-D methods (RECIST, mRECIST, EASL criteria) on multi-phasic CT scans to assess tumor response and survival prediction at 4 weeks after cTACE.

Materials and Methods

This retrospective cohort study was compliant with Health Insurance Portability and Accountability Act and approved by the Institutional Review Board. Informed consent was waived. The study design was in agreement with the Standards for Reporting of Diagnostic Accuracy guidelines.

Patients

Between July 2010 and July 2012, 84 consecutive patients with intermediate-stage HCC (BCLC stage B) without portal vein tumor thrombosis or extrahepatic metastasis were treated with cTACE and evaluated for this study. The inclusion criteria were the following: (1) age ≥ 18 years old, (2) confirmed diagnosis of HCC according to histologic examination or clinical-radiological results of early enhancement followed by washout on dynamic liver cross-sectional imaging [4], (3) preserved liver function with Child–Pugh Class A or B (≤ 7), (4) an Eastern Cooperative Oncology Group (ECOG) performance status 0, and (4) no previous locoregional treatment. The exclusion criteria were the following: (1) infiltrative HCC (i.e., spread of multiple, minute, ill-defined tumor nodules throughout the liver with indefinite tumor border), ($n = 2$), (2) poor CT image quality with motion artifacts ($n = 3$), and (3) no baseline and/or follow-up CT scans ($n = 27$). Of note, to reflect clinical practice, the presence of beam hardening artifacts (if any) due to Lipiodol deposition was not an exclusion criterion. Thus, a total of 52 consecutive patients were included into the final analysis.

Treatment

All cTACE procedures were performed by two interventional radiologists (G.H. and W.B.) with 20 and 8 years of experience, respectively, in hepatic interventions. When bilobar disease was present, the first cTACE was performed in the liver lobe bearing the largest lesions. No whole liver cTACE was performed. An emulsion containing up to 50 mg doxorubicin (Adriamycin, Zhejiang HISUN Pharmaceutical Co. Ltd., Taizhou, China) and 2–20 ml Lipiodol (Lipiodol Ultra-Fluide, Laboratoire Guerbet, Aulnay-sous-Bois, France) was infused, followed by bland embolization with 100–300 μm polyvinyl alcohol particles (Cook Medical Inc., Bloomington, USA) until

arterial inflow was substantially reduced as seen on fluoroscopy [1, 8, 19]. In all cases, either a selective or a super-selective approach was chosen.

CT Scan Parameters

All patients were seen 4 weeks after cTACE. At that point, multi-phasic contrast-enhanced CT was performed with a multi-detector CT scanner (Aquilion, Toshiba, Tokyo, Japan). The scanning parameters were the following: 120 kVp, 350 mA; 5 mm thick sections, a 300–400 mm field of view (matrix size 512×512), reconstruction algorithm FC08. Non-ionic contrast material (Ultravist, Schering Berlin, Germany) was intravenously administered at the rate of 3 ml/s. The bolus tracking technique was used for arterial phase. All the patients underwent scans which were composed of unenhanced, arterial, portal venous and delayed phases. In addition to imaging, physical examination and relevant laboratory values were obtained during this 1-month post-procedure follow-up. Last follow-up was on May 30, 2015.

Tumor Response Evaluation

1-D and 2-D measurements of RECIST, mRECIST, and EASL criteria were independently assessed by two blinded readers to any outcome data (W.B. and Y.Z., 8 and 4 years of experience). Liver assessment was performed at 4 weeks after the first TACE instead of multiple sessions to test the ability of the different response methods used in this study to identify responders and non-responders early in the course of treatment as the ultimate goal is to impact patient care in a timely matter and identify patient who benefit from therapy. The average value for measurements was used in the analysis. Up to two primary target lesions per patient were analyzed [13, 26]. The largest tumor that was considered to be the most appropriate target for the first TACE session was selected as the target lesion [16, 24]. Only lesions with a diameter over 1 cm were considered for the analysis [29]. For each lesion, the longest diameter (RECIST) or enhancing diameter (mRECIST) and largest enhancing area (EASL criteria) were measured on the CT images obtained at the arterial phase. All measurements were taken using standardized electronic calipers by using Digital Imaging in Communications and Medicine files.

For 3-D measurement, qEASL was performed using a semiautomatic 3-D software prototype (Medisys; Philips Research, Suresnes, France) as described in detail previously [20]. Briefly, a 3-D semiautomatic tumor segmentation was performed by a radiological reader (Y.Z. with one year of experience in this software, who was not involved in the TACE procedures) on the arterial phase contrast-

enhanced CT images before and after TACE. The accuracy and reader-independent reproducibility of the semiautomatic tumor segmentation software have been shown previously [6, 27]. Moreover, the software has been used across imaging modalities (MR, cone-beam CT and multi-detector CT) [30]. From the tumor segmentation, the whole tumor volume and the enhancing portion of tumor volume (cm^3) could be obtained. To measure the enhancing volume, the CT scan obtained just before contrast medium administration (i.e., unenhanced CT scan) was registered to the arterial phase CT scan and then subtracted in order to remove any background attenuation [5, 15]. The subtraction is a key feature because it mitigates contribution of background attenuation from the Lipiodol deposition that would otherwise obscure the true enhancement from the contrast medium injection (Fig. 1). The 3-D tumor segmentation mask was then transposed onto this subtracted CT scan. The enhancing tumor volume was obtained as follows: a region of interest formed by 1 cm^3 was placed in the normal appearing liver parenchyma as a reference for normalization to calculate the relative enhancement within the tumor [6, 7]. This procedure was done by two independent readers (Y.Z. and S.S. who had each one year of experience with the software) who did not participated in the cTACE procedures, and the average values from the two readers were used in the analysis.

Tumor response categories consisted of complete response (CR), partial response (PR), stable disease (SD) and progressive disease (PD). The criteria for the four models are summarized in Fig. 2. In addition, objective response referred to the sum of CR and PR, whereas non-response referred to the sum of SD and PD. All evaluations were based on the sum of the target lesions [4, 9, 11, 18].

Statistical Analysis

Continuous variables were summarized as means and standard deviation. Categorical variables were expressed as frequencies and percentages. Overall survival was measured from the time of treatment until the time of death from any cause. Patients lost to follow-up or alive at the end of the observation period were censored. Survival analysis was performed with the Kaplan–Meier method and the log-rank test. Cox proportional hazards model was used to examine risk factor association with survival. The two-way mixed-effects intra-class correlation (ICC) was calculated to grade inter-reader agreement as poor (ICC, < 0.5), moderate (ICC, $0.5\text{--}0.74$), good (ICC, $0.75\text{--}0.89$) or excellent (ICC, > 0.9) [3]. Inter-assessment concordance between similar categorical items of the four response criteria was measured using the k coefficient. The strength of agreement based on k values was interpreted as follows: k 0.01–0.20, slight agreement; k 0.21–0.40, fair

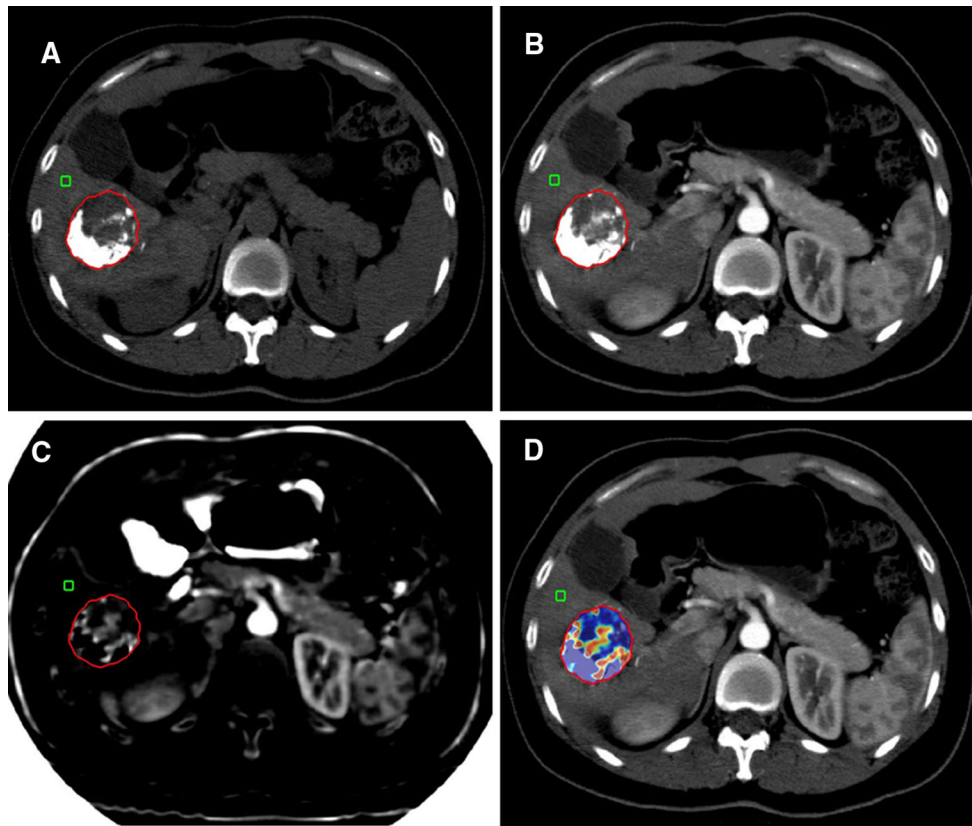
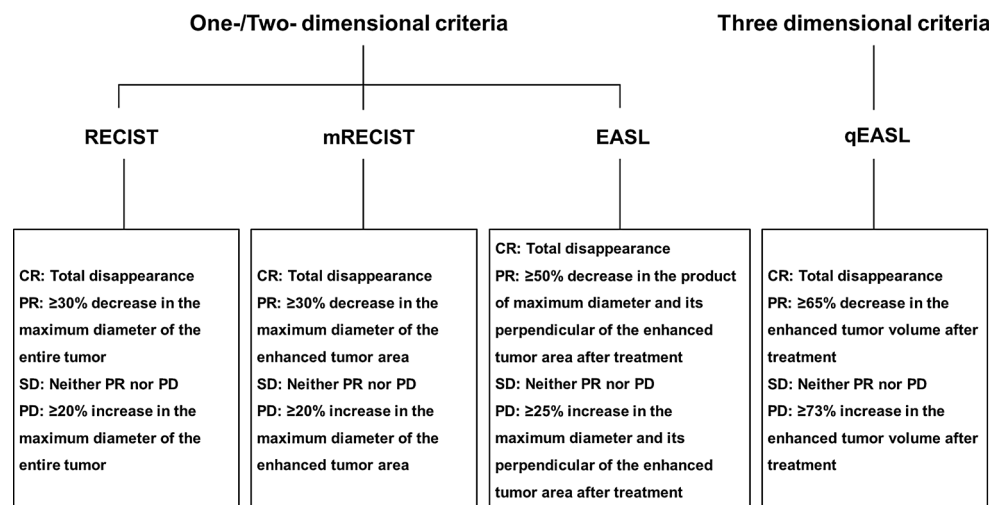


Fig. 1 CT scans obtained in a HCC patient after cTACE treatment: **A** before contrast-enhanced image; **B** after contrast-enhanced image; **C** before-enhanced image was subtracted from enhanced image in

order to remove any background signal including Lipiodol deposition; **D** qEASL color-map was overlaid on the subtracted image, which showed the enhance area of tumor

Fig. 2 Illustration of response classification based on the imaging biomarker assessment methods of RECIST, mRECIST, EASL and qEASL



agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; 0.81–0.99 almost perfect agreement [17]. A difference with a two-tailed *P* value < 0.05 was considered statistically significant. Statistical analyses were performed using SPSS (SPSS Inc., version 17.0, Chicago, IL).

Results

Baseline patients’ demographics and clinical characteristics are presented in Table 1. Of the 52 patients, 41 (78.8%) were male and 47 (90.4%) were tested positive for hepatitis virus B/C infection. Mean patient age was 51 years (range, 21–81 years). By May 30, 2015, 45

Table 1 Patient baseline demographics and clinical characteristics ($n = 52$)

Variable	No. (%)
Age/years, mean (range)	51 (21–81)
Gender	
Male	41 (78.8%)
Female	11 (21.2%)
Etiology	
Hepatitis B/C infection	47 (90.4%)
Other	5 (9.6%)
Cirrhosis	
Yes	27 (51.9%)
No/unknown	25 (48.1%)
Child–Pugh class	
A	51 (98%)
B (7)	1 (2%)
Ascites	
No	48 (92.3%)
Yes	4 (7.7%)
No. of HCC nodules	
1–2	45 (86.5%)
≥ 3	7 (13.5%)
Tumor size (cm), mean (range)	9.5 (4.5–24.6)
AFP (ng/dL)	
< 400 ng/dL	29 (55.8%)
≥ 400 ng/dL	23 (44.2%)
Baseline laboratory values, mean (range)	
International normalized ratio (INR)	1.1 (0.9–1.5)
Albumin, g/L	39.5 (31.4–51.3)
Alanine aminotransferase (U/L)	49.6 (14–211)
Aspartate aminotransferase (U/L)	55.2 (19–166)
Total bilirubin, mg/dL	1 (0.4–2.3)

HCC hepatocellular carcinoma, AFP α -fetoprotein

(86.5%) patients had died. The median overall survival was 12.9 months (95% CI 7.2–18.6). Median follow-up period was 13.5 months (range 1.2–54.1). The 1-, 2- and 3-year survival rates were 55.8, 34.6 and 15.4%, respectively. A median of two TACE sessions (range 1–9) was performed per patient for a total of 132 procedures. The mean interval from the baseline CT scan to the TACE procedure was 2.9 ± 2.5 days. The mean interval from the cTACE procedure to the follow-up CT scan was 29.1 ± 2.2 days.

For RECIST, the mean longest tumor diameter at baseline and at 4 weeks after cTACE was 9.5 ± 3.8 and 9.1 ± 3.6 cm, respectively ($P = 0.954$). For mRECIST, the mean longest enhancing tumor diameter at baseline and at 4 weeks after cTACE was 8.4 ± 3.2 and 5.4 ± 3.6 cm, respectively ($P = 0.177$). For EASL criteria, the mean

largest enhancing tumor area at baseline and at 4 weeks after cTACE was 52.7 ± 49 and 25 ± 30 cm², respectively ($P = 0.100$). For qEASL, at baseline, the mean whole tumor volume was 392.8 ± 432 cm³ and the mean enhancing tumor volume was 110.9 ± 121.2 cm³. At 4 weeks after cTACE, the whole tumor volume did not change significantly with a mean value of 351.8 ± 345.1 cm³ ($P = 0.594$), whereas the mean enhancing tumor volume decreased to 78.7 ± 64.6 cm³ although the difference did not reach a statistical significance ($P = 0.094$). Inter-reader agreement was excellent or good for all measurements before and after TACE (ICC = 0.963 and 0.970, respectively, for RECIST, 0.912 and 0.922, respectively, for mRECIST; 0.858 and 0.919, respectively, for EASL criteria; 0.956 and 0.881, respectively, for qEASL; $P < 0.01$ for all).

Comparison of Survival According to Four Categories

The outcome of tumor response and correlated survival according to RECIST, mRECIST, EASL criteria and qEASL are shown in Table 2. Using RECIST, most patients were classified as SD (92.3%). As a consequence, the survival outcome could not be compared among the different response groups (Fig. 3A). When using mRECIST and EASL criteria, patients who were classified as SD dropped to 36.5 and 34.6%, respectively. mRECIST and EASL criteria achieved a similar response rate (CR + PR) in 61.5 and 63.4%, respectively. Both assessment methods identified 1 (2%) patient to PD. Inter-criterion agreement showed a good agreement between mRECIST and EASL criteria, as reflected by a k value of 0.776. However, the survival curves crossed or were very close to each other for mRECIST and EASL criteria-based response assessment ($P > 0.05$ for each comparison) (Fig. 3B, C). According to EASL criteria, the survival difference between PR and SD showed a trend but did not reach a statistical significance (20.4 vs. 8.6 months, respectively, $P = 0.057$).

A poor correlation was found between qEASL and RECIST ($k = 0.054$), mRECIST ($k = 0.130$) and EASL criteria ($k = 0.025$). When using qEASL, no patient was classified into the CR group. The number of patients who were classified into PR was 8 (15.4%). On the other hand, qEASL yielded a statistically significant or demonstrated a trend toward difference in the probability of survival across the different response categories (Fig. 3D). Notably, qEASL identified 10 patients with PD at an early time point after treatment and these patients showed a significantly shorter median overall survival compared to those without PD [7.6 months (95% CI 5.9–9.3) vs. 20.4 months (95% CI 12.5–28.3), respectively, $P = 0.012$]. Multivariate analysis

Table 2 The outcome of radiological assessments and overall survival according to the respective tumor response assessment method

	Number of patients (<i>N</i> = 52)	Median overall survival (months)
RECIST		
Complete response	0 (0)	–
Partial response	2 (4%)	–
Stable disease	48 (92.3%)	15.9 (95% CI 5.6–26.2)
Progression disease	2 (4%)	–
mRECIST		
Complete response	6 (11.5%)	22.4 (95% CI 5–39.8)
Partial response	26 (50%)	14.1 (95% CI 8.5–19.7)
Stable disease	19 (36.5%)	10.6 (95% CI 4.1–17.1)
Progression disease	1 (2%)	–
EASL		
Complete response	6 (11.5%)	22.4 (95% CI 5.0–39.8)
Partial response	27 (51.9%)	20.4 (95% CI 10.7–30.1)
Stable disease	18 (34.6%)	8.6 (95% CI 6.1–11.1)
Progression disease	1 (2%)	–
qEASL		
Complete response	0 (0)	–
Partial response	8 (15.4%)	28.2 (95% CI 17.7–58.6)
Stable disease	34 (65.3%)	12.3 (95% CI 6.7–17.9)
Progression disease	10 (19.2%)	7.6 (95% CI 5.9–9.3)

RECIST response evaluation criteria in solid tumors; *mRECIST* modified response evaluation criteria in solid tumors, *EASL* European Association for the Study of the Liver, *qEASL* quantitative European Association for the Study of the Liver

Median overall survival could not be calculated because the number of patients for this classification was too small

showed that disease progression was an independent predictor for overall survival with hazard ratio of 1.9 [95% CI 1.1–3.3], $P = 0.025$] (Table 3). Similar uni/multivariate analysis for survival prediction could not be performed using RECIST, mRECIST and EASL criteria because too few patients were categorized as progressive disease based on these methods (Table 3). Figure 4 shows a patient with different response outcomes according to these four methods.

Comparison of Survival According to Response and Non-response

With RECIST, 2 (4%) patients were classified as responders (objective response = CR + PR) and 50 (96.2%) as non-responders (SD + PD). The corresponding survival were 11.3 months (95% CI not available) versus 14.1 months (95% CI 4.4–23.8), respectively ($P = 0.444$). With mRECIST, 32 (61.5%) patients were responders and the remaining 20 (38.5%) patients were non-responders. There was no difference in median overall survival between these two groups [15.9 months (95% CI 5.5–26.3) versus 8.6 months (95% CI 1.8–15.4), respectively, $P = 0.729$]. According to EASL criteria, 33 (63.5%)

patients showed objective response and 19 (36.5%) patients were non-responders. The median overall survival was 21.6 (95% CI 12.3–30.8) months and 8.9 (95% CI 6.2–11) months for responders and non-responders, respectively ($P = 0.041$). When using qEASL, 8 (15.3%) patients showed response and 44 (84.6%) patients showed non-response. The median overall survival was 38.2 (95% CI 17.8–58.6) months and 11.3 (95% CI 8.4–14.2) months for responders and non-responders, respectively ($P = 0.045$).

Discussion

The main finding of our study is that qEASL could be applied on CT images to assess tumor response and is not only able to predict survival but also to identify tumor progression at an early time point after cTACE.

Radiological assessment of response to anticancer therapy is widely used as a surrogate marker for survival. The prognosis of patients with HCC is closely related with disease progression in the liver. Thus, it is of the utmost importance to assess response early after treatment to identify patients who benefit the most from therapy. Conventional TACE is the gold standard and the most utilized

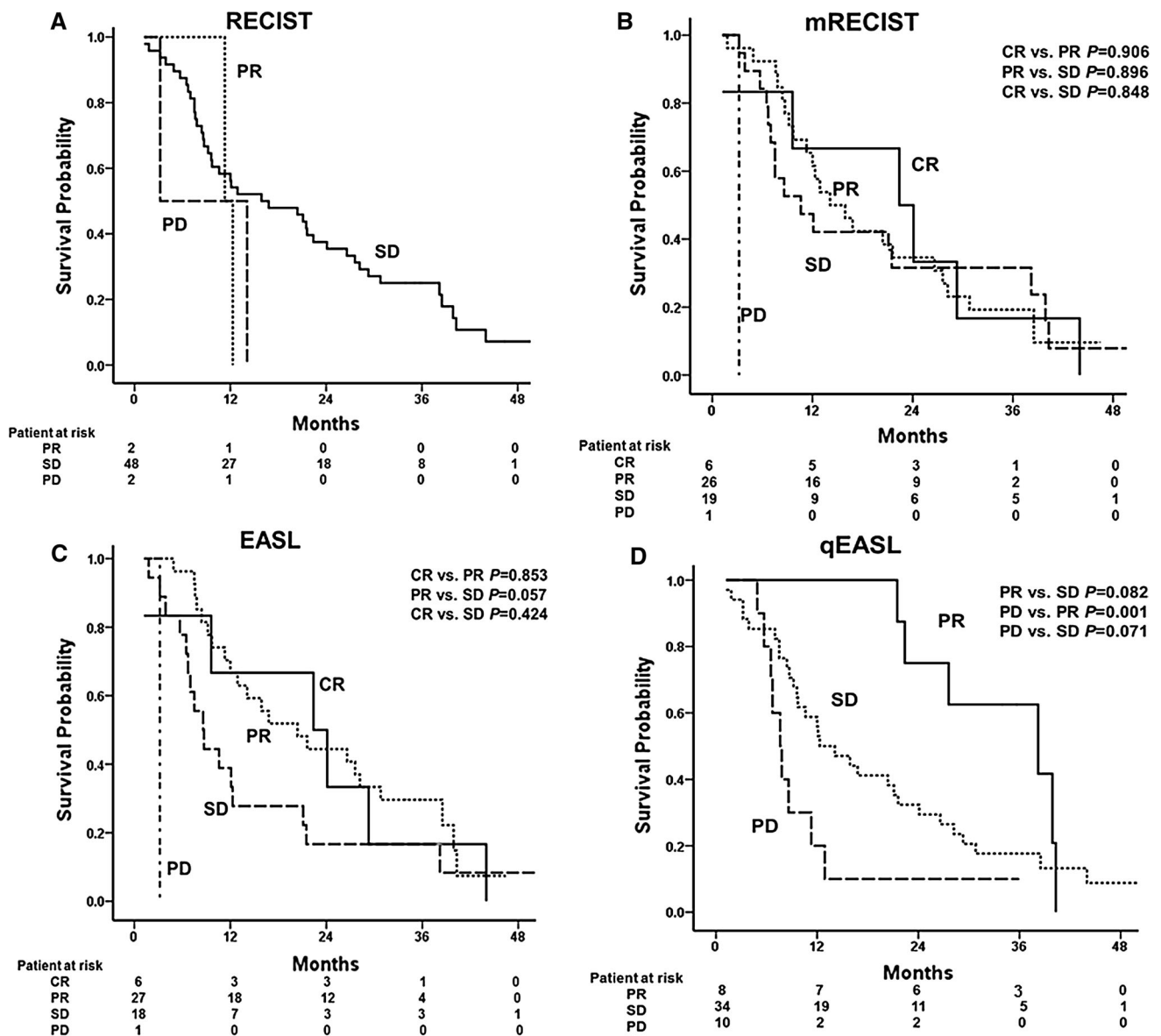


Fig. 3 Survival of 52 patients, as determined using **A** RECIST, **B** mRECIST, **C** EASL and **D** qEASL response assessment methods. Data were stratified into four response categories of complete

response (CR), partial response (PR), stable disease (SD) and progressive disease (PD). Note that qEASL was the only method capable of stratifying different tumor response correlating survival

treatment modality for intermediate-stage HCC patients [1]. Lipiodol deposition in tumor tissues limits the reliability of contrast enhancement on CT images. This is particularly true in lesions showing a high or inhomogeneous uptake of Lipiodol. However, since CT is the most frequently employed and available diagnostic modality worldwide, it is crucial to be able to accurately assess tumor response after cTACE regardless of the presence and degree of Lipiodol deposition.

Currently used tumor response evaluation methods, RECIST, mRECIST and EASL criteria, rely heavily on the radiologist’s judgment on how to perform the measure. The choice of the target lesion, the CT slice level into the lesion

where the measure is performed and how the measure is performed are the factors that are prone to error, especially when multiple readers and multiples time points are considered such as in clinical practice. Moreover, in heterogeneously enhancing tumors with a background level of attenuation (spontaneous background level of attenuation and attenuation due to Lipiodol), the use of conventional 1-/2-D-based measurements may be even more challenging. These problems are overcome when using a quantitative approach such as qEASL which analyzes the whole tumor in 3-D. Moreover, qEASL employs image subtraction between pre- and post-contrast medium injection images to mitigate the contribution of background

Table 3 Univariate and multivariate analysis considering the overall survival

Variable	Univariate analysis			Multivariate analysis		
	HR	95% CI	<i>P</i>	HR	95% CI	<i>P</i>
Age	0.976	0.956–0.996	0.019	0.977	0.956–0.998	0.032
Gender (male/female)	0.929	0.459–1.881	0.838			
Etiology (hepatitis infection/other)	1.021	0.398–2.622	0.965			
AFP (≥ 400 / < 400)	3.206	1.674–6.143	<0.001	2.947	1.427–6.086	0.003
Child–Pugh (B/A)	0.047	0–7.802	0.419			
Ascites (yes/no)	0.947	0.290–3.088	0.928			
Tumor size (≥ 5 cm/ < 5 cm)	1.662	0.919–3.007	0.093	1.065	0.432–2.627	0.892
Tumor number (≥ 3 / $1-2$)	1.555	0.648–3.732	0.323			
qEASL (PD vs. non-PD)	2.369	0.993–5.652	0.052	1.880	1.082–3.267	0.025

HR hazard ratio CI confidence interval AFP α -fetoprotein, PD progressive disease, qEASL quantitative European Association for the Study of the Liver

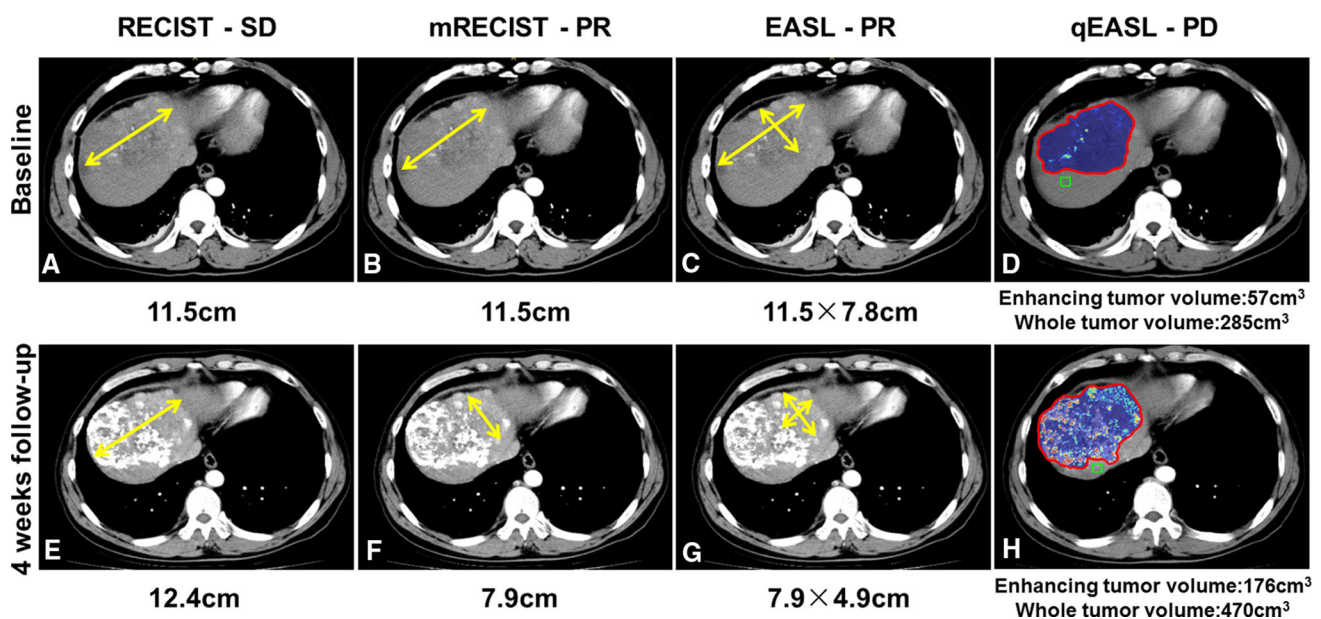


Fig. 4 The different outcomes of tumor response according to the four criteria in a patient with a poor survival time of 4.9 months: stable disease (SD) for RECIST, partial response (PR) for RECIST and mRECIST and progression disease (PD) for qEASL. A–D The

baseline scans according to the RECIST, mRECIST, EASL and qEASL, respectively. (E–H) The follow-up scans at 4 weeks after TACE treatment according to these four criteria

attenuation (Fig. 1). This allows for accurate tumor response assessment of remaining enhancing viable tumor independently of the pattern of enhancement and the degree of Lipiodol deposition.

Our study showed that RECIST is not a suitable response assessment method following TACE, confirming previously published results [9, 10, 12, 13, 21]. Indeed when using RECIST, more than 90% patients were classified as SD in our study, thus no stratification between responders and non-responders was possible and no survival data could be calculated. mRECIST failed to predict survival across the different response categories and in the objective response analysis. Interestingly, EASL criteria

performed better than mRECIST, despite a similar response rate [CR + PR (in 61.5 and 63.5% for mRECIST and EASL criteria, respectively)]. EASL criteria was able to predict survival between responders and non-responders ($P = 0.041$). Taken together, these results highlight the difficulty to assess tumor response when using mRECIST and EASL criteria as the identification of underlying viable enhancing tumor may be hampered by Lipiodol deposition as shown by a high number of patients classified for both methods as responders. This directly translates into a low number of patients classified as PD. On the other hand, qEASL was not only able to accurately predict survival between responders and non-responders but also achieved

better survival curves separation when compared to mRECIST and EASL criteria (Fig. 3). Previous validation works have reported high inter-reader reproducibility of this method and its radiological-pathological accuracy [6, 7, 28]. More importantly, qEASL was the only method able to identify disease progression which demonstrated to be associated with poor survival on multivariate analysis. Similar multivariate analysis could not be performed when using the other response criteria, in particular EASL criteria, because too few patients were categorized as PD.

Our study has some limitations. First, potential bias may exist because of the retrospective nature and relatively small sample size. However, we included a prospectively collected and homogeneous patient cohort with intermediate-stage HCC and well-preserved liver function, which minimized the impacts of other factors on prognosis. Further studies with a prospective design and larger sample size should be conducted to confirm the advantage of qEASL over the other response assessment methods on CT images following cTACE. Second, histopathology was not available when the radiological assessments were performed. However, a previous study on the rabbit model showed that tumor volume measured using segmentation software on CT images (similar software than the one used in our study) had a strong correlation with the actual tumor volume measured on pathology [22]. Moreover, it was demonstrated that 3-D segmentation of HCC lesions treated with TACE provided high volumetric concordance across MR, cone-beam CT and multi-detector CT images [30]. Further studies are needed to explore the performance of qEASL on CT images in determining the extent of pathological tumor necrosis. Third, our study was designed to investigate response on CT early after therapy (i.e., after the first cTACE). Future studies are needed to investigate the survival prediction ability of qEASL after multiple treatments. Fourth, 3-D quantitative tumor response assessment was performed using one platform. Further studies comparing the prediction abilities of different 3-D quantitative platforms are needed.

In conclusion, qEASL could be applied on CT images to assess tumor response in HCC patients following cTACE and is a more sensitive biomarker to predict survival and identify tumor progression than RECIST, mRECIST and EASL criteria at an early time point after cTACE.

Funding Our study was funded by NIH/NCI R01 CA160771, Philips Research North America, Cambridge, MA and National Natural Science Foundation of China 81172145 and 81420108020.

Author's Contributions All authors were involved in study conception and design. Yan Zhao, Rafael Duran and Wei Bai, Wenjun Wang helped in collection and analysis of data. Sven Kabus and MingDe Lin contributed to imaging technique support. Yan Zhao and Rafael Duran involved in manuscript writing. All authors helped in

revision of the manuscript. All authors were involved final approval of the version to be published.

Compliance with Ethical Standards

Conflicts of interest Jean-François Geschwind: Consultant: Bristol Myers, Guerbet, BTG, Philips, Merck, Terumor, Bayer. Grant Support: NIH, Boston Scientific, Guerbet, Philips. Founder and CEO PreScience Labs, LLC. MingDe Lin is a Philips employee and has grant support from NIH-R01. The other authors declared no conflict of interest.

References

1. European Association For The Study Of The Liver. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol.* 2012;56(4):908–43.
2. World Health Organization. Global battle against cancer won't be won with treatment alone—effective prevention measures urgently needed to prevent cancer crisis. *Cent Eur J Public Health.* 2014;22(1):23–8.
3. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1(8476):307–10.
4. 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.* 2001;35(3):421–30.
5. Buerger C, Senegas J, Kabus S, et al. Comparing nonrigid registration techniques for motion corrected MR prostate diffusion imaging. *Med Phys.* 2015;42(1):69–80.
6. Chapiro J, Wood LD, Lin M, et al. Radiologic-pathologic analysis of contrast-enhanced and diffusion-weighted MR imaging in patients with HCC after TACE: diagnostic accuracy of 3D quantitative image analysis. *Radiology.* 2014;273(3):746–58.
7. Chockalingam A, Duran R, Sohn JH, et al. Radiologic-pathologic analysis of quantitative 3D tumour enhancement on contrast-enhanced MR imaging: a study of ROI placement. *Eur Radiol.* 2016;26(1):103–13.
8. de Baere T, Arai Y, Lencioni R, et al. Treatment of liver tumors with lipiodol TACE: technical recommendations from experts opinion. *Cardiovasc Intervent Radiol.* 2016;39(3):334–43.
9. Duran R, Chapiro J, Frangakis C, et al. Uveal melanoma metastatic to the liver: the Role of quantitative volumetric contrast-enhanced mr imaging in the assessment of early tumor response after transarterial chemoembolization. *Transl Oncol.* 2014;7(4):447–55.
10. Edeline J, Boucher E, Rolland Y, et al. Comparison of tumor response by response evaluation criteria in solid tumors (RECIST) and modified RECIST in patients treated with sorafenib for hepatocellular carcinoma. *Cancer.* 2012;118(1):147–56.
11. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer.* 2009;45(2):228–47.
12. 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.* 2009;115(3):616–23.
13. Gillmore R, Stuart S, Kirkwood A, et al. EASL and mRECIST responses are independent prognostic factors for survival in

- hepatocellular cancer patients treated with transarterial embolization. *J Hepatol.* 2011;55(6):1309–16.
14. Gonzalez-Guindalini FD, Botelho MP, Harmath CB, et al. Assessment of liver tumor response to therapy: role of quantitative imaging. *Radiogr A Rev Publ Radiol Soc N Am Inc.* 2013;33(6):1781–800.
 15. Kabus S, Klinder T, Murphy K, van Ginneken B, van Lorenz C, Pluim JP. Evaluation of 4D-CT lung registration. *Medical image computing and computer-assisted intervention. MICCAI Int Conf Med Image Comput Assist Interv.* 2009;12(1):747–54.
 16. Kim BK, Kim SU, Kim MJ, et al. Number of target lesions for EASL and modified RECIST to predict survivals in hepatocellular carcinoma treated with chemoembolization. *Clin Cancer Res Off J Am Assoc Cancer Res.* 2013;19(6):1503–11.
 17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159–74.
 18. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. *Semin Liver Dis.* 2010;30(1):52–60.
 19. Liapi E, Geschwind JF. Transcatheter arterial chemoembolization for liver cancer: is it time to distinguish conventional from drug-eluting chemoembolization? *Cardiovasc Intervent Radiol.* 2011;34(1):37–49.
 20. Lin M, Pellerin O, Bhagat N, et al. Quantitative and volumetric European Association for the Study of the Liver and response evaluation criteria in solid tumors measurements: feasibility of a semiautomated software method to assess tumor response after transcatheter arterial chemoembolization. *J Vasc Interv Radiol JVIR.* 2012;23(12):1629–37.
 21. Liu L, Wang W, Chen H, et al. EASL- and mRECIST-evaluated responses to combination therapy of sorafenib with transarterial chemoembolization predict survival in patients with hepatocellular carcinoma. *Clin Cancer Res Off J Am Assoc Cancer Res.* 2014;20(6):1623–31.
 22. Pellerin O, Lin M, Bhagat N, Ardon R, Mory B, Geschwind JF. Comparison of semi-automatic volumetric VX2 hepatic tumor segmentation from cone beam CT and multi-detector CT with histology in rabbit models. *Acad Radiol.* 2013;20(1):115–21.
 23. 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. *J Hepatol.* 2011;54(4):695–704.
 24. 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.* 2010;303(11):1062–9.
 25. Sahani DV, Kalva SP. Imaging the liver. *Oncologist.* 2004;9(4):385–97.
 26. Shim JH, Lee HC, Won HJ, et al. Maximum number of target lesions required to measure responses to transarterial chemoembolization using the enhancement criteria in patients with intrahepatic hepatocellular carcinoma. *J Hepatol.* 2012;56(2):406–11.
 27. Tacher V, Lin M, Chao M, et al. Semiautomatic volumetric tumor segmentation for hepatocellular carcinoma: comparison between C-arm cone beam computed tomography and MRI. *Acad Radiol.* 2013;20(4):446–52.
 28. Tacher V, Lin M, Duran R, et al. Comparison of existing response criteria in patients with hepatocellular carcinoma treated with transarterial chemoembolization using a 3D quantitative approach. *Radiology.* 2016;278(1):275–84.
 29. 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(3):205–16.
 30. Wang Z, Chapiro J, Scherthaner R, et al. Multimodality 3D Tumor Segmentation in HCC Patients Treated with TACE. *Acad Radiol.* 2015;22(7):840–5.