

Published in final edited form as:

J Hepatol. 2011 April ; 54(4): 695–704. doi:10.1016/j.jhep.2010.10.004.

Role of EASL, RECIST and WHO Response Guidelines Alone or in Combination for Hepatocellular Carcinoma: Radiologic-Pathologic Correlation

Ahsun Riaz, MD¹, Khairuddin Memon, MD¹, Frank H. Miller, MD¹, Paul Nikolaidis, MD¹, Laura M. Kulik, MD², Robert J. Lewandowski, MD¹, Robert K. Ryu, MD¹, Kent T. Sato, MD¹, Vanessa L. Gates, MS¹, Mary F. Mulcahy, MD³, Talia Baker, MD⁴, Ed Wang, PhD⁵, Ramona Gupta, MD¹, Ritu Nayar, MD⁶, Al B Benson 3rd, MD³, Michael Abecassis, MD⁴, Reed Omary, MD¹, and Riad Salem, MD, MBA^{1,4}

¹ Department of Radiology, Northwestern University, Chicago IL

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

³ Department of Medicine, Division of Medical Oncology, Northwestern University, Chicago, IL

⁴ Department of Surgery, Division of Transplant Surgery, Northwestern University, Chicago, IL

⁵ Department of Surgery, Section of Biostatistics, Northwestern University, Chicago, IL

⁶ Department of Pathology, Northwestern University, Chicago, IL

Abstract

Purpose—We sought to study receiver-operating characteristics (ROC) of the European Association for the Study of the Liver (EASL), Response Evaluation Criteria in Solid Tumors (RECIST) and World Health Organization (WHO) guidelines of assessing response following locoregional therapies individually and in various combinations.

Methods—Eighty-one patients with hepatocellular carcinoma underwent liver explantation following locoregional therapies. Response was assessed using EASL, RECIST and WHO. Kappa statistics were used to determine inter-method agreement. Uni/multivariate logistic regression analyses were performed to determine the variables predicting complete pathologic necrosis. Numerical values were assigned to the response classes: complete response=0, partial response=1, stable disease=2 and progressive disease=3. Various mathematical combinations of EASL and WHO were tested to calculate scores and their ROCs were studied using pathological examination of the explant as the gold standard.

Results—Median times (95% CI) to WHO, RECIST and EASL response were 5.3 (4–11.5), 5.6 (4–11.5) and 1.3 months (1.2–1.5) respectively. Kappa coefficients for WHO/RECIST, WHO/EASL and RECIST/EASL were 0.78, 0.28 and 0.31 respectively. EASL response demonstrated significant odds ratios for predicting complete pathologic necrosis on uni/multivariate analyses. Calculated areas under the ROC curves were: RECIST: 0.63, WHO: 0.68, EASL: 0.82, EASL

Corresponding Author: Riad Salem, MD, MBA, Director, Interventional Oncology, Department of Radiology, 676 N. St. Clair, Suite 800, Chicago, Illinois 60611 USA, Tel: 312-695-6371, Fax: 312-695-0654, r-salem@northwestern.edu.

Conflict of Interest: No conflicts of interest.

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

+WHO: 0.82, EASLxWHO: 0.85, EASL+(2xWHO): 0.79 and (2xEASL)+WHO: 0.85. An EASLxWHO Score of ≤ 1 had 90.2% sensitivity for predicting complete pathologic necrosis.

Conclusion—The product of WHO and EASL demonstrated better ROC than the individual guidelines for assessment of tumor response. The EASLxWHO Scoring System provides a simple and clinically applicable method of response assessment following locoregional therapies for hepatocellular carcinoma.

Keywords

Hepatocellular carcinoma; Locoregional therapies; Imaging; Pathologic correlation

INTRODUCTION

Hepatocellular carcinoma (HCC) is responsible for over one half million deaths annually and its incidence is increasing[1]. Orthotopic liver transplantation (OLT) and resection are considered curative; however, most patients do not meet selection criteria[2]. Sorafenib has shown a survival benefit for advanced HCC[3–4]. Locoregional therapies (LRTs) deliver toxic thermal/chemical/radioactive doses to tumors with minimal toxicity to normal tissue. Transarterial chemoembolization and yttrium-90 radioembolization are LRTs that have demonstrated a palliative role in the management of HCC[5–9].

Given the lack of standardization and evidence for utilizing functional imaging, morphologic methodologies such as the Response Evaluation Criteria in Solid Tumors (RECIST) are still considered standard for response assessment[10–11]. World Health Organization (WHO) guidelines and RECIST measure size of tumor irrespective of the amount of necrosis seen[12–13]. Response by size criteria are seen at 5–6 months following treatment and suggest the ability of the surrounding parenchyma to regenerate normally[14–15].

When specifically studying HCC, the European Association of the Study of the Liver (EASL) guidelines (developed in 2001) recommend measuring only enhancing tissue in the tumor[16–17]. EASL response is an early predictor of necrosis (1–2 months)[14–15]. Although the measurement of enhancing tissue has an evolving role in assessing response following treatment with cytostatic drugs (e.g. sorafenib) and/or LRTs, there have been few reports using EASL (necrosis) guidelines to assess response[5,18–20]. In fact, recent studies for hepatocellular carcinoma have employed size guidelines instead of EASL guidelines to assess response[3–4].

The question arises: what is the pathologic evidence that forms the basis for these guidelines? Furthermore, despite the fact that these size and necrosis guidelines are accepted by disciplines involved in clinical and academic activities pertaining to HCC, their accuracy has not been thoroughly investigated, particularly in relation to long-term outcomes or pathological correlates. Liver transplantation/resection for HCC provides an ideal research medium for attempting to answer this question. The limitations of current imaging response assessment guidelines have been previously discussed[21].

Recognizing the significant differences and potential advantages of size and necrosis criteria for response, we postulated that a combination of these two may be better able to predict actual tumor necrosis than either used alone. We sought to test this concept using receiver-operating characteristics (ROC) with explant pathology as the gold standard.

METHODS

Eighty-one patients with HCC who were treated with LRT prior to transplantation or resection at our institution were studied. This analysis was approved by the Institutional Review Board and was Health Insurance Portability and Accountability Act compliant.

Evaluation and Staging

Pre-treatment HCC diagnosis was confirmed by biopsy or radiographic evidence as defined by the American Association for the Study of Liver Diseases (AASLD) and EASL guidelines[16,22]. Baseline patient characteristics including alpha-fetoprotein (AFP) were analyzed; patients were staged using Child-Pugh, United Network for Organ Sharing (UNOS) TNM, and Barcelona Clinic Liver Cancer (BCLC) classification systems[23–24]. Treatment for unresectable HCC was determined at multi-disciplinary HCC conference. Radiologists performing the baseline and follow-up assessments were blinded to whether the patients had undergone OLT/resection. This was performed to minimize under-staging bias; in other words, a radiologist with preexisting knowledge of eventual transplantation may be more inclined to under-stage the disease[14].

Locoregional Therapies

Chemoembolization—Chemoembolization is a transarterial therapy delivering high doses of chemotherapeutic agents to tumor via the hepatic artery. The technical details have been discussed elsewhere[25].

Chemoembolization-Radiofrequency Ablation—In eight patients, chemoembolization was followed by radiofrequency ablation (RFA) of the tumor.

Radioembolization—Yttrium-90 radioembolization is a transarterial therapy where high radioactive doses are delivered to the tumor via the hepatic artery. The device used was glass-based (TheraSphere[®], MDS Nordion, Ottawa, Canada)[26–27]. Technical details have been discussed elsewhere[28].

Imaging Analyses

Contrast-enhanced CT or gadolinium-enhanced MRI was used as the cross-sectional imaging modality as per our protocol defined below. Although MRI was preferred, a minority of patients had contraindications and hence underwent CT.

MR Imaging Technique—The anatomic abdominal MRI protocol for imaging the liver used in our institution included transverse and coronal T2-weighted half-Fourier acquisition single-shot turbo spin echo, T2-weighted turbo spin-echo with fat suppression. Unenhanced and dynamic gadolinium-enhanced T1-weighted images were acquired by fat suppressed gradient echo using shared prepulses. T1-weighted fat suppressed gradient echo images were obtained using the following parameters: TR/TE, 120–160/1.9; flip angle, 80°; slice thickness, 6mm; gap, 1.8mm; matrix, 256×179. Dynamic gadolinium-enhanced MR images were acquired in the arterial phase (scan time based on fluoroscopy-preparation timing sequence), venous phases (45–60 and 90 seconds), and delayed phase with images obtained 2–5 minutes after contrast injection. Gadopentetate dimeglumine (Magnevist; Bayer HealthCare Pharmaceuticals, Berlin, Germany) was administered at a dose of 0.1 mmol/kg, followed by 20 ml saline flush (2ml/sec) with a power injector (Spectris; Medrad Warrendale, PA).

CT Imaging Technique—All multidetector CT scans were obtained using Siemens Somatom Sensation 64 and 16-slice scanners (Siemens Medical Solutions, Forchheim,

Germany). Unenhanced arterial and portal venous phase images were acquired according to our liver protocol. Contrast-enhanced images were obtained after approximately 40 seconds in the arterial phase and 70 seconds in the venous phase after injection of 125 mL of iohexol (Omnipaque 350, GE Healthcare, Waukesha, WI) at a rate of 5 mL/s.

Response Guidelines—In order to standardize and create a uniform methodology to calculate time-to-endpoint analyses (i.e. from time from first treatment), the focus of this study was the primary index lesion which was defined as the dominant lesion that was treated in the first treatment session. Data on the primary index lesion have been recently published[29]. Since this is only a radiologic-pathologic analysis, untreated lesions were not included. Response in the dominant lesion was assessed using size (WHO, RECIST) and necrosis (EASL) guidelines[13,16–17]. For the purposes of this study, response rate refers to the combined rate of complete response (CR) and partial response (PR). Using conservative methodology, lesions with small enhancing nodules (irrespective of how minimal) despite near complete necrosis, were considered tumor and classified as EASL PRs (not CRs).

In contradistinction to radioembolization, CT response assessment following chemoembolization may be confounded by the presence of lipiodol[14–15]. Given this issue, we adopted a conservative approach to EASL response assessment in chemoembolization patients followed with CT (n=5) including: a) the presence of any lipiodol within the tumor post-treatment was counted as enhancing tissue (and not necrosis), b) an EASL CR or PR could only be reported if the lipiodol had cleared out of the tumor and there was absent (EASL CR) or >50% reduction in contrast enhancement (EASL PR). This strict methodology would permit the true identification of reduced contrast enhancement post-treatment (and hence tumor necrosis) only in cases where the lipiodol had washed out of the treated tumor.

All calculations were performed from the date of first LRT. Imaging follow-up was performed at 1-month following treatment with subsequent scans performed at scheduled 2–3 month intervals, as per our routine standard-of-care. Four radiologists who were well-versed in assessing response following LRT performed response assessment. Any ambiguity was resolved by consensus.

The details of the RECIST, WHO and EASL are outlined in Table 1[13–14,16]. When using size criteria to assess response, one could use RECIST or WHO guidelines. The only difference is that RECIST recommends measuring one lesional dimension whereas the WHO recommends measuring two lesional dimensions. It is known that RECIST and WHO guidelines have high inter-method agreement[30]. To test our hypothesis of combining size (RECIST or WHO) and necrosis (EASL) guidelines, it was necessary to choose either WHO or RECIST. We decided to use WHO over RECIST since it is the oncologic gold-standard for radiologists to measure lesions in two and not one dimension.

To test our hypothesis that combination of size and necrosis guidelines may be better able to predict necrosis than any of these guidelines used alone, we devised numerical Scoring Systems as follows. For purposes of calculation in this analysis, the following numbers were arbitrarily assigned to the WHO classes: CR=0, PR=1, SD=2 and PD=3; similarly the EASL classes were assigned the following numbers: CR=0, PR=1, SD=2 and PD=3.

In order to maintain simplicity and clinical applicability, the following mathematical combinations of EASL and WHO were explored:

$$\begin{aligned} & \text{EASL+WHO} \\ & \text{EASL} \times \text{WHO} \\ & \text{EASL}+(2 \times \text{WHO}) \\ & (2 \times \text{EASL})+\text{WHO} \end{aligned}$$

Scores were assigned using these four formulae and are shown in Supplemental Table 1. For example, a lesion demonstrating an EASL CR and WHO PR would be assigned “0” and “1” respectively, and the EASLxWHO score assigned to this treated lesion would be “0”. A hypothetical 2x2 cm lesion with all possible imaging outcomes (WHO and EASL) following treatment is displayed schematically in Figure 1.

Although we could have explored more complex mathematical combinations of EASL and WHO, we recognized that increasing complexity would minimize the clinical applicability of the scoring system.

Pathologic Evaluation

The treated tumor (primary index lesion) was identified by the interventional radiologist; the location of this lesion was communicated to the reviewing pathologist for radiologic-pathologic correlation. The pathologist evaluated it for evidence of gross and histologic necrosis after serial sectioning using our institutional protocol (5 mm slice thickness); routine hematoxylin and eosin stains were used to prepare the slides. Percentage necrosis was semi-quantitatively classified by the pathologist into: a) Complete pathologic necrosis (CPN): absence of any viable HCC; b) >50% necrosis: significant necrosis with presence of any viable HCC; c) <50% necrosis: minimal necrosis[14]. Conservatively, the presence of any viable malignant cells histologically precluded the classification of CPN.

Statistical Analyses

Kappa statistics were used to measure inter-method agreement between RECIST, WHO and EASL[31].

Since not all patients reach a response (endpoint for time to response), median times to response were calculated using the Kaplan-Meier method[32]. Stepwise logistic regression analysis was performed to simultaneously control for baseline AFP (<200 ng/mL), AFP response (>50% decrease from baseline), age, gender, ethnicity, Child-Pugh class, UNOS stage, EASL and WHO response.

ROC curves were studied and areas under the ROC curves (AUC) with the 95% confidence intervals (CI) were calculated. The variables of interest (Scoring Systems) were entered as the numerical values calculated using the methods described above. The classification variable under all assumptions was presence or absence of CPN. The cut-off value at which the accuracy was the maximum was chosen[33–34]. Sensitivity and specificity were calculated at that cut-off value. The comparison of the AUCs were performed according to the method described by Hanley and McNeil[35].

RESULTS

Patient Characteristics

Baseline patient characteristics are summarized in Table 2. The source of the explanted tissue was: OLT in 77 patients (95%), liver resection in three patients (4%) and autopsy in one patient (1%). This provided the gold standard explant pathology data.

The median time from last cross-sectional imaging scan to explantation was 1.2 months (CI: 0.9–1.7). The last scan prior to explantation provided the source of the response data (by WHO, RECIST and EASL), the characteristics of which were compared to explant pathology. Although we recognize that immediate pre-explantation imaging would have been ideal, this was not possible since it does not represent the standard of care.

Radiologic and Pathologic Data

The radiologic and pathologic outcome data are summarized in Table 3. 73/81 (90%) of patients had MRIs. 8/81 (10%) patients were evaluated using CT; 5/8 were treated with chemoembolization while 3/8 were treated with radioembolization. 271 cross-sectional imaging scans were reviewed, translating into 3.35 scans (range: 2–8) per patient. The median imaging follow-up time prior to explant was 3.2 months (CI: 2.3–4.6). Using Kaplan-Meier methodology, the median times (CI) to WHO, RECIST and EASL response were 5.3 (4–11.5), 5.6 (4–11.5) and 1.3 months (1.2–1.5) respectively.

Kappa Statistics

The kappa (κ) coefficient (CI) for WHO and RECIST was 0.78 (0.67–0.90). The κ coefficient (CI) for EASL and RECIST was 0.28 (0.14–0.42), while the κ coefficient for EASL and WHO was 0.31 (0.17–0.45).

Logistic Regression Analysis

WHO response had an odds ratio of 3.5 (CI: 1.4–8.9; $P=0.008$) on univariate analysis. Only EASL response was found to have a significant odds ratio of 32.7 (CI: 4.1–261.9; $P=0.001$) and 30.1 (CI: 3.8–242.1; $P=0.001$) on univariate and multivariate stepwise logistic regression analysis.

Receiver-Operating Characteristics

Table 4 presents the receiver operating characteristics of the scoring systems analyzed in this study. The sensitivity/specificity for response (CR or PR) by the three guidelines in predicting CPN were as follows: RECIST: 51%/73%; WHO: 54%/75% and EASL: 97%/45%.

The calculated AUCs were as follows: RECIST: 0.63, WHO: 0.68, EASL: 0.82, EASL +WHO: 0.82, EASL \times WHO: 0.85, EASL+(2 \times WHO): 0.79 and (2 \times EASL)+WHO: 0.85. All of these curves were statistically significant when compared to an AUC of 0.5 (by chance). The finding of a greater AUC using WHO (0.68) vs. RECIST (0.63) supported our approach of using WHO when combining size and necrosis guidelines.

Figure 2 shows the same ROC curves superimposed on each other to give a graphical representation of the differences in the areas under the curve. Supplemental Table 2 presents the actual differences in the areas under the curves. EASL had a significantly larger AUC than WHO (0.82 vs. 0.68; difference in AUC=0.14; $P=0.028$). EASL \times WHO and (2 \times EASL) +WHO were seen to have the largest AUCs (0.85). AUCs of various combinations of EASL and WHO were not significantly different in our study.

An EASL \times WHO Score of ≤ 1 had 90.2% sensitivity and 67.5% specificity for predicting complete pathologic necrosis. However, application of stricter criteria i.e. EASL \times WHO Score of 0 demonstrated a specificity of 87.5%.

DISCUSSION

Two conditions usually co-exist in patients with HCC: background cirrhosis and the malignancy. Survival, disease-free survival, progression-free survival and time-to-progression are outcomes that require lengthy follow-up for the patients to reach the endpoints. Although response by cross-sectional imaging is considered a weak outcome in the hierarchy of outcomes in trial designs for HCC, it is integral in phase I and II studies[36]. Furthermore, in clinical practice, changes in tumor size and amount of enhancing tissue are of paramount importance and aid in clinical decision making processes.

A role for the EASL (necrosis) guidelines in assessing response following systemic therapy with cytostatic agents has been suggested[36]. Our study demonstrated a high odds ratio for EASL in predicting CPN. This suggests that there is minimal chance of there being CPN unless an EASL response is observed. Furthermore, EASL response is seen quickly following treatment[14]. Recent data suggest that EASL response may not be able to predict pathologic necrosis with high accuracy in embolic therapies (such as conventional transarterial chemoembolization) since embolization may not only lead to the proximal blockade of arteries supplying the tumor with subsequent reduction in contrast enhancement (potentially artificial representation of necrosis), but also the rapid development of parasitized vessels and collaterals to the tumor[15]. On the other hand the longer time to WHO response suggests a regenerative process around the necrotic tumor that has been time tested. Response by WHO guidelines demonstrates a process that signifies cell death in the tumor with the synchronous ability of the surrounding parenchyma to regenerate normally. Since the utilization of size guidelines (WHO or RECIST) alone does not capture the amount of necrosis and the utilization of EASL guidelines alone may not capture the ability of the surrounding normal parenchyma to regenerate normally, we sought to investigate if the combination of EASL and WHO had improved ROCs when compared to either one used independently.

As investigators, we favor the use of WHO guidelines (bi-dimensional measurements) over RECIST (uni-dimensional measurements) due to: a) the low AUC when using RECIST, b) the high inter-observer agreement between RECIST and WHO, c) not all lesions are spherical and post-treatment response (especially following LRTs) does not occur in a spherical fashion and d) as we move towards technology (software) facilitating volumetric (3-D) measurements, we thought it more rational to use bi-dimensional rather than uni-dimensional guidelines.

Standard imaging guidelines that are accepted and utilized universally by all specialties involved in HCC patient care are lacking. The recently published article by Llovet *et al.* recommends using the EASL amendment of the RECIST[36]. Figure 3 is a schematic representation of cases where the measurement of a single dimension is possible. However, as shown in Figure 4, in the presence of an irregular pattern of contrast enhancement, it is difficult to determine which single enhancing-tumor dimension to measure. Thus, the use of objective methodology to measure response by EASL guidelines becomes difficult given: a) there are currently no universally available software or measuring tools to determine volumes of enhancing tissue in a time and cost-efficient manner (optimal method) and b) the determination of amount of enhancing tissue using one or two dimensions is difficult given the irregular nature of necrosis (Figure 4). We therefore suggest assessing EASL as a subjective measurement in assessing response following LRT. The rationale behind this suggestion is that: a) EASL CRs are easy to classify subjectively (there is no enhancement whatsoever), b) the distinction between the widely divergent radiologic findings of EASL PR and PD is subjectively possible given the significant difference in their appearance, leaving everything in between as SD, c) measurement of the amount of enhancing tissue is

not exact in a majority of cases due to the irregular nature of necrosis (Figure 4) and d) the combination of EASL (subjective) and WHO (objective) criteria through the use of the EASLxWHO Score may mathematically minimize the subjectivity associated with EASL. Figure 4 illustrates the commonly encountered complexities in calculating EASL response following locoregional therapy.

EASL and WHO demonstrated minimal inter-method agreement in our study (κ coefficient=0.31). This is similar to data presented by Forner *et al*[17]. EASL response is achieved earlier compared to WHO response (1.3 vs. 5.3 months). Response by EASL (necrosis) guidelines demonstrates necrosis and may demonstrate appropriate targeting of the lesion when using LRTs. EASL guidelines have an important role in the patients listed for OLT as an earlier-achieved response may demonstrate better tumor biology allowing favorable clinical decisions. They are also especially useful in patients undergoing thermally ablative procedures such as RFA in which complete targeting of tumor may leave a lesion completely necrotic (CR by EASL) despite no change or even an increase in size (SD or even PD by WHO)[37].

Mathematically, the combination of (2xEASL)+WHO and EASLxWHO showed the largest AUCs (0.85) in our analysis. An advantage of using the (2xEASL)+WHO assumption is that all numerical values between 0 and 9 have a place in this scoring system, whereas a disadvantage is that CRs by one guideline may be assigned a score of 6 because of PD status by the other guideline (e.g. WHO PD and EASL CR following RFA). EASLxWHO numerical scores of 5, 7 and 8 are mathematically impossible as seen in the Supplemental Table 1. However, since both of these mathematical combinations provided similar AUCs, we believe the simple EASLxWHO has obvious clinical applicability, including: a) all CRs are assigned the value “0”, irrespective of whether the CR was by EASL or WHO and b) EASLxWHO Scores ≤ 1 have a high accuracy for determining CPN. Hence, translated clinically, the objective of treating HCC patients would be to achieve at least an EASLxWHO Score of ≤ 1 . An EASL \times WHO score of 1 could only be possible if a lesion decreased in size and the amount of enhancing tissue decreased as well. It is interesting to note that the EASL \times WHO score of 1 (PR by both EASL and WHO guidelines) showed complete pathologic necrosis in 58% lesions as seen in Table 3. 14 (45%) of lesions showing EASL PR had complete pathologic necrosis. An explanation of this would be the utilization of conservative yet strict methodology, where any amount of enhancing tissue (even a small enhancing nodule despite near complete necrosis), was counted as tumor and hence not an EASL CR (Figure 5). Hence, we probably underreported EASL CRs. Furthermore, an EASL \times WHO score of 0 (stricter criteria) would be even more specific for predicting complete pathologic necrosis. For this analysis, we opted to explore the clinical implications of using the EASLxWHO in Figure 3 as this was the simpler calculation; it also had a similar AUC to that of (2xEASL)+WHO.

Guidance on the clinical utility of the EASLxWHO Scoring System and its time-dependence are presented in the form of the examples shown in Figure 3. Based on our earlier discussion, an EASLxWHO Score ≤ 1 has a high accuracy of determining CPN. Such a simple binomial categorical (yes/no) application of this scoring system may be of interest. In other words, these types of analyses may suggest that response should be categorized in a binomial fashion (yes/no) as opposed to a continuous numerical variable representing change in size/amount of enhancing tissue. Furthermore, the utility of the EASLxWHO score at different time points following treatment is also interesting and shown in Figure 3.

What is the objective of differentiating EASL from WHO response when looking at HCC? It has been shown that EASL response is achieved earlier (1–2 months) and WHO response is achieved later (5–6 months)[14–15]. Using only one guideline at various time points

following treatment could miss response evidenced by the other guideline. The EASLxWHO Scoring System allows concomitant representation of both guidelines at different time points and easy numerical comparison to previous scans. It is also important to stress that subsequent scans are compared to the maximum response post-treatment and not to the baseline (pre-treatment) scan, in accordance with guidelines[13].

Functional imaging may have an increasing role in the near future for response assessment following therapy for HCC[38–39]. However, we should recognize that these technologies will not be universally available to all clinicians and will depend on local resources. Hence, until the widespread availability of technology able to provide functional information and/or determine degree of necrosis, simple morphologic guidelines should continue to play a complementary role in response assessment, despite their limitations.

There are limitations to this analysis. Minimal differences between AUCs of combinations of EASL and WHO limited the ability to attain statistical significance upon comparison. However, it should be acknowledged that 81 patients with RECIST-WHO-EASL-Pathology correlates represent a large number considering the specific patient population with HCC requiring liver explantation. The utilization of this scoring system is of value in patients contraindicated for MRI following chemoembolization where lipiodol deposition may limit the applicability of EASL guidelines[15]. A conservative reporting methodology was adopted for the 5 patients that were followed with CT. This study was performed at the lesional level; future scoring systems for response assessment and determination of time-to-progression should incorporate the appearance of new lesions, development of PVT and development of extrahepatic metastases. Changes in alpha-fetoprotein levels may also have a role in a composite scoring system for response assessment following LRT of HCC[40]. Eight patients in this study received RFA after chemoembolization; however, the focus of this study is on the EASL-WHO imaging finding in the index lesion irrespective of treatment modality. The scoring system is of value in these patients as well (all EASL CRs are assigned a score of 0) and may have a role in patients treated with ablative LRTs. Although at first glance, the inclusion of patients having undergone three treatment modalities as part of this analysis could be seen as a potential limitation, this would be contrary to the guidelines; these were intended to be applicable to all treatment modalities. Otherwise, EASL, RECIST and WHO would have to provide distinct response guidelines specifically for transarterial, ablative and systemic therapies and their various combinations.

In conclusion, combinations of EASL and WHO (e.g. EASLxWHO) demonstrate improved ROCs compared to the individual guidelines (either EASL, WHO or RECIST alone) in the response assessment following LRTs. The Scoring System is a simple way of combining size and necrosis characteristics in a temporal manner following treatment with an EASL × WHO Score of 0 (high specificity for predicting CPN) being the optimal clinical objective. The trend of the EASLxWHO Score assigned to a treated lesion at subsequent follow-up scans may assist in the clinical decision-making process after LRT for HCC. More data are required to validate these concepts and to expand the use of such a scoring system to assess response following other therapies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. El-Serag HB. Hepatocellular carcinoma and hepatitis C in the United States. *Hepatology*. 2002; 36(5 Suppl 1):S74–83. [PubMed: 12407579]

2. Mazzaferro V, Regalia E, Doci R, Andreola S, Pulvirenti A, Bozzetti F, et al. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. *N Engl J Med.* 1996; 334(11):693–699. [PubMed: 8594428]
3. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, et al. Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med.* 2008; 359(4):378–390. [PubMed: 18650514]
4. Cheng AL, Kang YK, Chen Z, Tsao CJ, Qin S, Kim JS, et al. Efficacy and safety of sorafenib in patients in the Asia-Pacific region with advanced hepatocellular carcinoma: a phase III randomised, double-blind, placebo-controlled trial. *Lancet Oncol.* 2009; 10(1):25–34. [PubMed: 19095497]
5. Salem R, Lewandowski RJ, Mulcahy MF, Riaz A, Ryu RK, Ibrahim S, et al. Radioembolization for hepatocellular carcinoma using Yttrium-90 microspheres: a comprehensive report of long-term outcomes. *Gastroenterology.* 2010; 138(1):52–64. [PubMed: 19766639]
6. Maddala YK, Stadheim L, Andrews JC, Burgart LJ, Rosen CB, Kremers WK, et al. Drop-out rates of patients with hepatocellular cancer listed for liver transplantation: outcome with chemoembolization. *Liver Transpl.* 2004; 10(3):449–455. [PubMed: 15004776]
7. Lewandowski RJ, Kulik LM, Riaz A, Senthilnathan S, Mulcahy MF, Ryu RK, et al. A comparative analysis of transarterial downstaging for hepatocellular carcinoma: chemoembolization versus radioembolization. *Am J Transplant.* 2009; 9(8):1920–1928. [PubMed: 19552767]
8. Llovet JM, Real MI, Montana X, Planas R, Coll S, Aponte J, et al. Arterial embolisation or chemoembolisation versus symptomatic treatment in patients with unresectable hepatocellular carcinoma: a randomised controlled trial. *Lancet.* 2002; 359(9319):1734–1739. [PubMed: 12049862]
9. Lo CM, Ngan H, Tso WK, Liu CL, Lam CM, Poon RT, et al. Randomized controlled trial of transarterial lipiodol chemoembolization for unresectable hepatocellular carcinoma. *Hepatology.* 2002; 35(5):1164–1171. [PubMed: 11981766]
10. Choi H, Charnsangavej C, Faria SC, Macapinlac HA, Burgess MA, Patel SR, et al. Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria. *J Clin Oncol.* 2007; 25(13):1753–1759. [PubMed: 17470865]
11. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer.* 2009; 45(2):228–247. [PubMed: 19097774]
12. WHO. WHO Handbook for Reporting Results of Cancer Treatment. WHO offset publication; Geneva, Switzerland, WHO: 1979.
13. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst.* 2000; 92(3):205–216. [PubMed: 10655437]
14. Riaz A, Kulik L, Lewandowski RJ, Ryu RK, Giakoumis Spear G, Mulcahy MF, et al. Radiologic-pathologic correlation of hepatocellular carcinoma treated with internal radiation using yttrium-90 microspheres. *Hepatology (Baltimore, Md).* 2009; 49(4):1185–1193.
15. Riaz A, Lewandowski RJ, Kulik L, Ryu RK, Mulcahy MF, Baker T, et al. Radiologic-Pathologic Correlation of Hepatocellular Carcinoma Treated with Chemoembolization. *Cardiovasc Intervent Radiol.* 2009
16. Bruix J, Sherman M, Llovet JM, Beaugrand M, Lencioni R, Burroughs AK, et al. Clinical management of hepatocellular carcinoma. Conclusions of the Barcelona-2000 EASL conference. European Association for the Study of the Liver. *J Hepatol.* 2001; 35(3):421–430. [PubMed: 11592607]
17. Forner A, Ayuso C, Varela M, Rimola J, Hessheimer AJ, de Lope CR, et al. Evaluation of tumor response after locoregional therapies in hepatocellular carcinoma: are response evaluation criteria in solid tumors reliable? *Cancer.* 2009; 115(3):616–623. [PubMed: 19117042]
18. Malagari K, Chatzimichael K, Alexopoulou E, Kelekis A, Hall B, Dourakis S, et al. Transarterial chemoembolization of unresectable hepatocellular carcinoma with drug eluting beads: results of an

- open-label study of 62 patients. *Cardiovasc Intervent Radiol.* 2008; 31(2):269–280. [PubMed: 17999110]
19. Lammer J, Malagari K, Vogl T, Pilleul F, Denys A, Watkinson A, et al. Prospective Randomized Study of Doxorubicin-Eluting-Bead Embolization in the Treatment of Hepatocellular Carcinoma: Results of the PRECISION V Study. *Cardiovasc Intervent Radiol.* 2009
 20. Malagari K, Pomoni M, Kelekis A, Pomoni A, Dourakis S, Spyridopoulos T, et al. Prospective Randomized Comparison of Chemoembolization with Doxorubicin-Eluting Beads and Bland Embolization with BeadBlock for Hepatocellular Carcinoma. *Cardiovasc Intervent Radiol.* 2009
 21. Jaffe CC. Measures of response: RECIST, WHO, and new alternatives. *J Clin Oncol.* 2006; 24(20): 3245–3251. [PubMed: 16829648]
 22. Bruix J, Sherman M. Management of hepatocellular carcinoma. *Hepatology.* 2005; 42(5):1208–1236. [PubMed: 16250051]
 23. Llovet JM, Fuster J, Bruix J. The Barcelona approach: diagnosis, staging, and treatment of hepatocellular carcinoma. *Liver Transpl.* 2004; 10(2 Suppl 1):S115–120. [PubMed: 14762851]
 24. El-Serag HB, Marrero JA, Rudolph L, Reddy KR. Diagnosis and treatment of hepatocellular carcinoma. *Gastroenterology.* 2008; 134(6):1752–1763. [PubMed: 18471552]
 25. Bruix J, Sala M, Llovet JM. Chemoembolization for hepatocellular carcinoma. *Gastroenterology.* 2004; 127(5 Suppl 1):S179–188. [PubMed: 15508083]
 26. TheraSphere Yttrium-90 microspheres package insert. MDS Nordion; Kanata, Canada: 2004.
 27. Salem R, Thurston KG, Carr BI, Goin JE, Geschwind JF. Yttrium-90 microspheres: radiation therapy for unresectable liver cancer. *J Vasc Interv Radiol.* 2002; 13(9 Pt 2):S223–229. [PubMed: 12354840]
 28. Salem R, Thurston KG. Radioembolization with 90Yttrium Microspheres: A State-of-the-Art Brachytherapy Treatment for Primary and Secondary Liver Malignancies: Part 1: Technical and Methodologic Considerations. *J Vasc Interv Radiol.* 2006; 17(8):1251–1278. [PubMed: 16923973]
 29. Riaz A, Miller FH, Kulik LM, Nikolaidis P, Yaghamai V, Lewandowski RJ, et al. Imaging response in the primary index lesion and clinical outcomes following transarterial locoregional therapy for hepatocellular carcinoma. *JAMA.* 2010; 303(11):1062–1069. [PubMed: 20233824]
 30. Park JO, Lee SI, Song SY, Kim K, Kim WS, Jung CW, et al. Measuring response in solid tumors: comparison of RECIST and WHO response criteria. *Jpn J Clin Oncol.* 2003; 33(10):533–537. [PubMed: 14623923]
 31. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977; 33(1):159–174. [PubMed: 843571]
 32. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc.* 1958; 53:457–481.
 33. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem.* 1993; 39(4):561–577. [PubMed: 8472349]
 34. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982; 143(1):29–36. [PubMed: 7063747]
 35. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology.* 1983; 148(3):839–843. [PubMed: 6878708]
 36. Llovet JM, Di Bisceglie AM, Bruix J, Kramer BS, Lencioni R, Zhu AX, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. *J Natl Cancer Inst.* 2008; 100(10):698–711. [PubMed: 18477802]
 37. Sala M, Llovet JM, Vilana R, Bianchi L, Sole M, Ayuso C, et al. Initial response to percutaneous ablation predicts survival in patients with hepatocellular carcinoma. *Hepatology.* 2004; 40(6): 1352–1360. [PubMed: 15565564]
 38. Rhee TK, Naik NK, Deng J, Atassi B, Mulcahy MF, Kulik LM, et al. Tumor response after yttrium-90 radioembolization for hepatocellular carcinoma: comparison of diffusion-weighted functional MR imaging with anatomic MR imaging. *J Vasc Interv Radiol.* 2008; 19(8):1180–1186. [PubMed: 18656011]
 39. Kamel IR, Reyes DK, Liapi E, Bluemke DA, Geschwind JF. Functional MR imaging assessment of tumor response after 90Y microsphere treatment in patients with unresectable hepatocellular carcinoma. *J Vasc Interv Radiol.* 2007; 18(1 Pt 1):49–56. [PubMed: 17296704]

40. Riaz A, Ryu RK, Kulik LM, Mulcahy MF, Lewandowski RJ, Minocha J, et al. Alpha-fetoprotein response after locoregional therapy for hepatocellular carcinoma: oncologic marker of radiologic response, progression, and survival. *J Clin Oncol*. 2009; 27(34):5734–5742. [PubMed: 19805671]

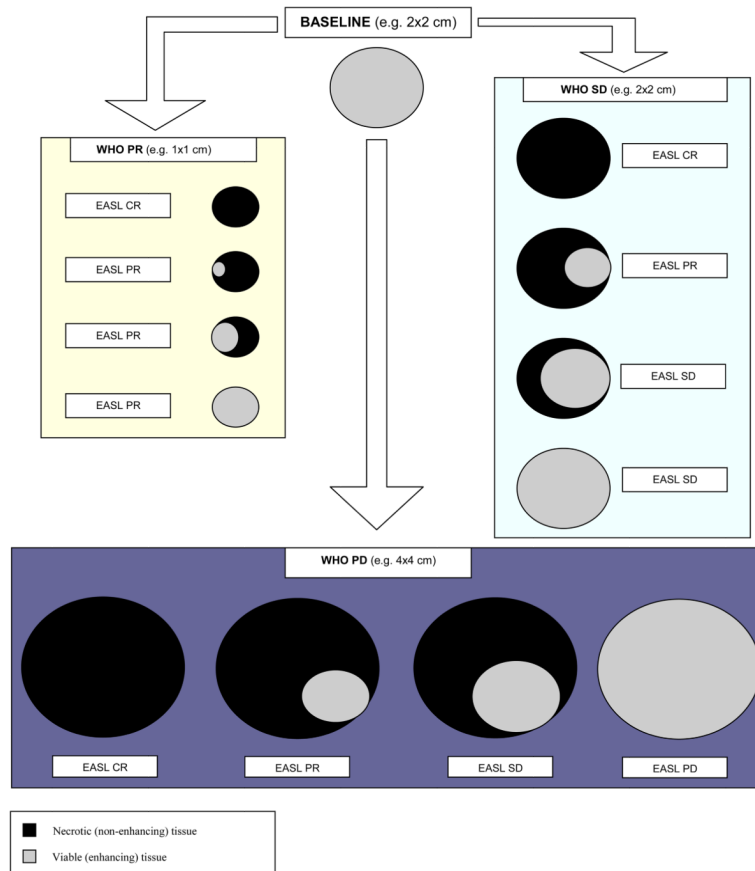


Figure 1. Schematic representation of the World Health Organization (WHO) and European Association for Study of the Liver (EASL) guidelines. The WHO guidelines suggest measurement of the whole tumor irrespective of the amount of necrosis (black) seen. The EASL guidelines suggest measurement of the enhancing tissue only (grey). This figure represents the discordance between the two guidelines.

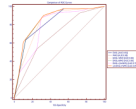


Figure 2.

Figure demonstrating comparison of receiver-operating characteristics (ROC) curves using the various scoring systems. The following scoring systems were analyzed: Response Evaluation Criteria for Solid Tumors (RECIST); World Health Organization (WHO) guidelines; European Association for Study of the Liver (EASL); EASL + WHO; EASL \times WHO; EASL + (2xWHO); (2xEASL) + WHO.

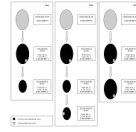


Figure 3.

Examples using EASLxWHO scoring systems. A patient is diagnosed with a $2 \times 2 \text{ cm}^2$ HCC. The baseline EASLxWHO Score assigned to this lesion would be 4 [SD (2) by both guidelines; $2 \times 2 = 4$]. **3(a):** The one-month post-treatment scan shows that the size of the lesion remains the same but the lesion shows EASL PR (EASLxWHO Score= $1 \times 2 = 2$). This decrease in the EASLxWHO Score at the one-month scan would demonstrate response in the treated lesion. Clinically, this would translate into observation of the patient with routine scheduled follow-up. The four-month scan shows that the size of the lesion is now $1 \times 1 \text{ cm}^2$ and the lesion now shows an EASL CR (EASLxWHO Score= $0 \times 1 = 0$). In this patient the EASLxWHO Scores allocated to this patient were $4 \rightarrow 2 \rightarrow 0$ at baseline, one-month and four-month follow-up scans, respectively; the odds of this lesion having complete pathologic necrosis are high. This decrease in the EASLxWHO Score would again translate into observation of the patient and routine follow-up. **3(b):** In the same patient as the one represented in Figure 3a, the seven-month scan shows that the lesion is still $1 \times 1 \text{ cm}^2$ but now shows reappearance of enhancing tissue which would be classified as EASL PD (EASLxWHO Score= $1 \times 3 = 3$). In this patient with an initial response, there is subsequent progression as seen in the EASLxWHO Scores of $4 \rightarrow 2 \rightarrow 0 \rightarrow 3$ at baseline, one-month, four-month and seven-month follow-up scans, respectively. Now that the score is >1 , this would signify progression and may translate into clinical interventions such as re-treatment. **3(c):** Another clinical scenario may be presented as follows using the same patient example as above with similar findings on the one-month post-treatment scan (EASLxWHO Score= $1 \times 2 = 2$). However, the findings on the four-month post-treatment scan show that the lesion remained $2 \times 2 \text{ cm}^2$ and there is still enhancing tissue in the lesion (EASL PR), the EASLxWHO Score remains 2. In this patient the allocated EASLxWHO Scores were $4 \rightarrow 2 \rightarrow 2$ at baseline, one-month and four-month follow-up scans, respectively. This stability in the score between the one-month and the four-month scan may demonstrate treatment failure due to incomplete targeting of the tumor.

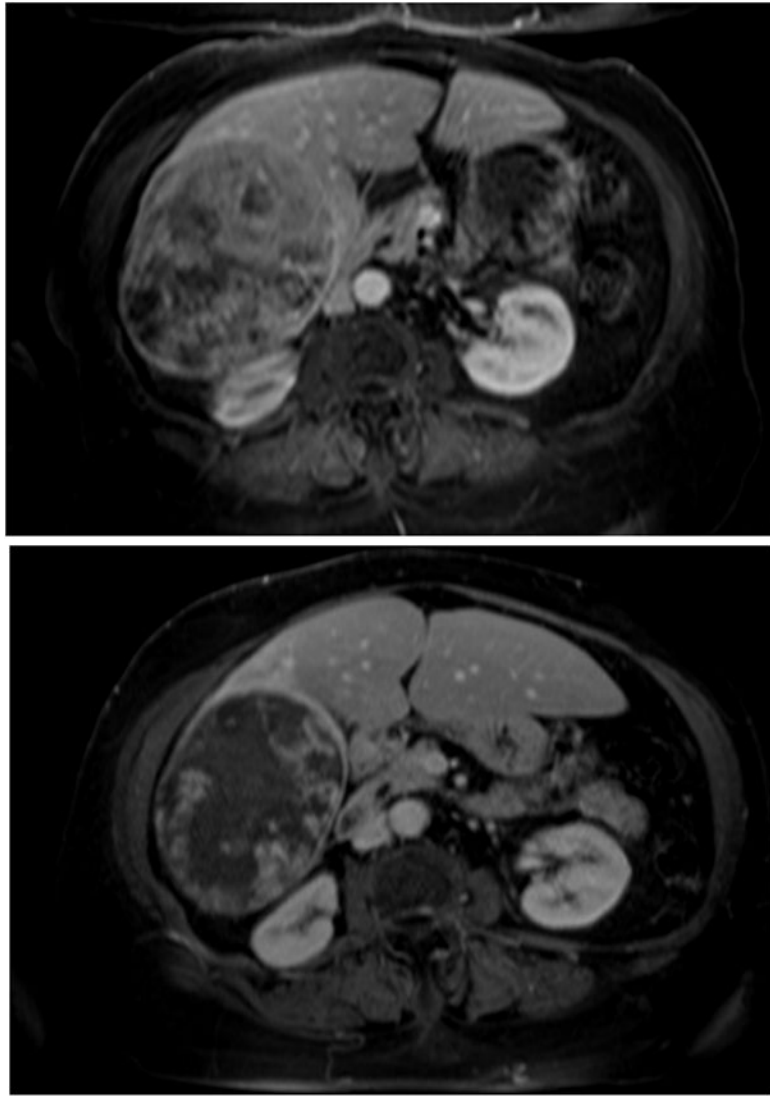


Figure 4.
4(a) Contrast-enhanced MRI demonstrates a large right lobe biopsy-proven HCC. Although the baseline scan clearly demonstrates a viable tumor, there is a moderate amount of necrosis at baseline. Quantifying the amount of enhancing tissue (viable hepatocellular carcinoma) at baseline without using computer software applying EASL guidelines would be difficult given the irregular and ill-defined interface between viable disease and necrotic tissue. **4(b)** Post-treatment MRI demonstrates a decrease in enhancing tissue representing a clear EASL response by subjective assessment; however the exact quantification of necrosis versus viable tissue is again difficult because of the persistent irregularity in pattern of contrast enhancement.

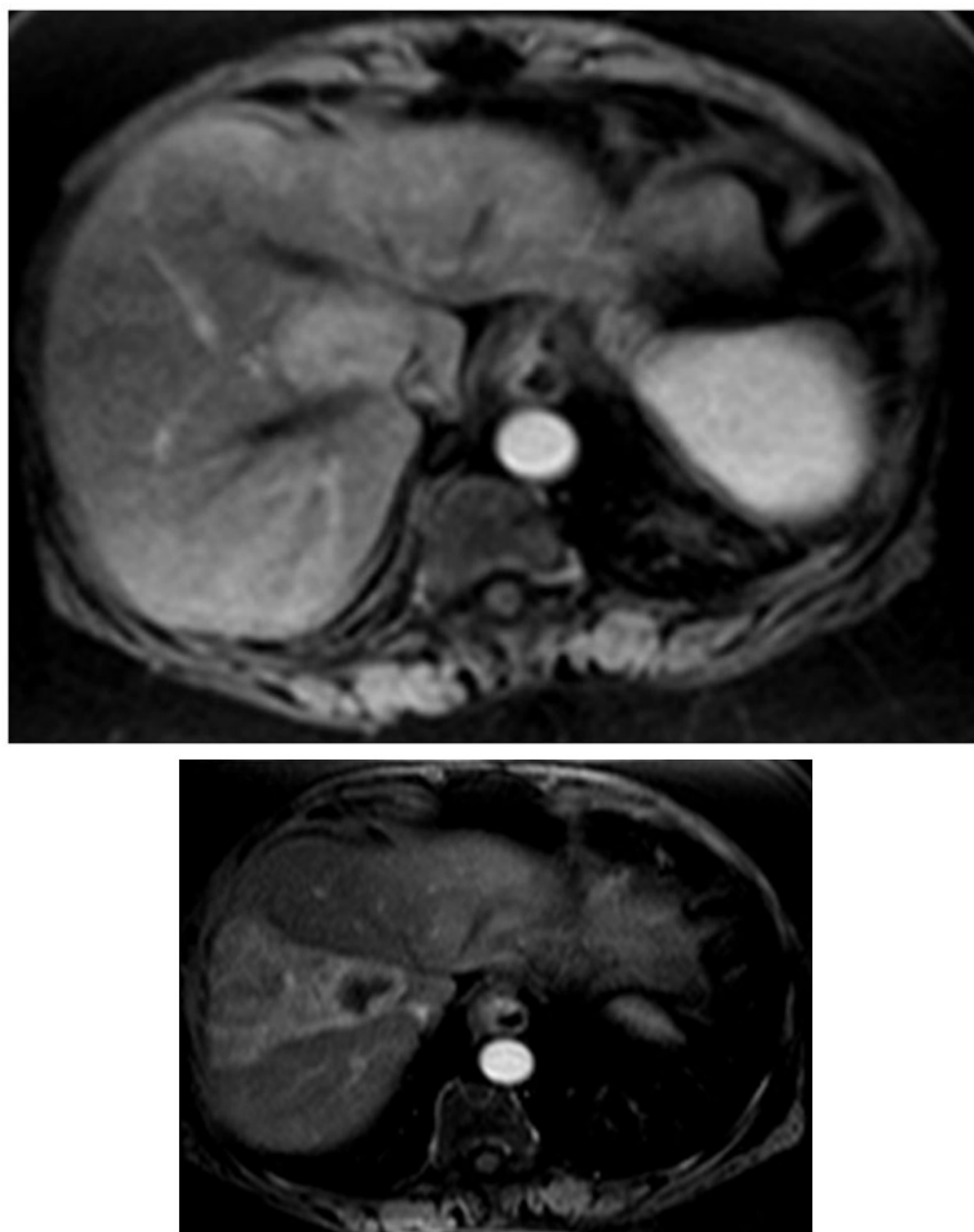


Figure 5. **5(a)** This is an enhancing hepatoma in segment 5/8. **5(b)** Following treatment, one can clearly see the radiation effect along the vascular planes of segment 5/8. Because of the slight irregular mural enhancement, this was labeled as EASL PR, illustrating the conservative methodology used throughout the analysis.

TABLE 1

Imaging Response Guidelines for Hepatocellular Carcinoma

System	Classification	Definition
World Health Organization (WHO) Guidelines	CR	100% decrease in cross-product of target lesion(s)
	PR	≥ 50% decrease in cross-product of target lesion(s)
	SD	<50% decrease to ≤25% increase in cross-product of target lesion(s)
	PD	>25% increase from maximum response of target lesion(s)
Response Evaluation Criteria for Solid Tumors (RECIST) Guidelines	CR	100% decrease in maximum diameter of target lesion(s)
	PR	≥30% decrease in maximum diameter of target lesion (s)
	SD	<30% decrease to ≤20% increase in maximum diameter of target lesion(s)
	PD	>20% increase from maximum response of target lesion(s)
European Association for Study of the Liver (EASL) Guidelines*	CR	100% decrease in amount of enhancing tissue in target lesion(s)
	PR	≥50% decrease in amount of enhancing tissue in target lesion(s)
	SD	<50% decrease in amount of enhancing tissue in target lesion(s)
	PD	>25% increase in amount of enhancing tissue in target lesion(s) and/or New enhancement in previously treated lesions warranting further LRT

LRT=locoregional therapy, CR=Complete response, PR=Partial response, SD=Stable disease, PD=Progressive disease

* To report EASL response conservatively, the presence of lipiodol within the tumor was counted as enhancing tissue rather than necrosis in patients who were treated using chemoembolization and evaluated using CT (n=5)

TABLE 2

BASELINE PATIENT CHARACTERISTICS

Demographics		Total N=81
Age	<65 years	59 (69)
	≥65 years	22 (31)
Gender	Male	65 (80)
	Female	16 (20)
Ethnicity	Caucasian	53 (65)
	Asian	12 (15)
	Hispanic	9 (11)
	African American	7 (9)
Etiology	HCV	41 (51)
	HBV	10 (12)
	Alcohol	13 (16)
	Autoimmune	3 (4)
	Cryptogenic	6 (7)
	Other	8 (10)
Distribution	Solitary	40 (49)
	Multifocal	41 (51)
Method of Diagnosis	Imaging (CT/MRI)	50 (62)
	Biopsy	31 (38)
Treatment	Chemoembolization	36 (44)
	Chemoembolization+RFA	8 (10)
	Radioembolization	37 (46)
Staging System	Stage	
UNOS	T1	1 (1)
	T2	40 (49)
	T3	22 (27)
	T4a	12 (15)
	T4b	5 (7)
	N/M	1 (1)
BCLC	A	40 (49)
	B	31 (38)
	C	9 (12)
	D	1 (1)
Child Pugh Class	A	39 (48)
	B	41 (51)
	C	1 (1)

UNOS=United Network for Organ Sharing, BCLC=Barcelona Clinic for Liver Cancer

TABLE 3

RADIOLOGIC AND PATHOLOGIC DATA

.Imaging Guidelines	Total N=81	Pathologic Class			P value*	
		<50%	51-99%	100%		
EASL	CR	31 (38)	0	5 (16)	26 (84)	<.001
	PR	31 (38)	4 (13)	13 (42)	14 (45)	
	SD	12 (16)	9 (75)	3 (25)	0	
	PD	7 (9)	5 (72)	1 (14)	1 (14)	
WHO	CR	1 (1)	0	0	1 (100)	0.015
	PR	31 (38)	2 (6)	8 (26)	21 (68)	
	SD	35 (44)	9 (26)	10 (29)	16 (45)	
	PD	14 (17)	7 (50)	4 (29)	3 (21)	
RECIST	CR	1 (1)	0	0	1 (100)	0.057
	PR	31 (38)	2 (6)	9 (29)	20 (65)	
	SD	38 (47)	11 (29)	11 (29)	16 (42)	
	PD	11 (14)	5 (45)	2 (18)	4 (37)	
EASLxWHO	0	31 (38)	0	5 (16)	26 (84)	<.001
	1	19 (23)	2 (10)	6 (32)	11 (58)	
	2	8 (10)	1 (12.5)	5 (62.5)	2 (25)	
	3	4 (5)	1 (25)	2 (50)	1 (25)	
	4	11 (14)	8 (73)	3 (27)	0	
	6	1 (1)	1 (100)	0	0	
	9	7 (9)	5 (72)	1 (14)	1 (14)	

EASL=European Association for Study of the Liver, WHO=World Health Organization, RECIST=Response Evaluation Criteria for Solid Tumors, CR=Complete response, PR=Partial response, SD=Stable disease, PD=Progressive disease

* P value calculated using the Fisher's Exact Test.

TABLE 4
REPRESENTATION OF RECEIVER OPERATING CHARACTERISTICS

Scoring System	Possible Scores	Area Under Curve (95% Confidence Interval)	P Value *	Maximum Accuracy		
				Score	Sensitivity (%)	Specificity (%)
RECIST	0, 1, 2, 3	0.63 (0.52-0.73)	0.037	≤ 1	51.2	72.5
WHO	0, 1, 2, 3	0.68 (0.57-0.78)	0.002	≤ 1	53.7	75
EASL	0, 1, 2, 3	0.82 (0.72-0.90)	<.001	≤ 0	63.4	87.5
EASL+WHO	0, 1, 2, 3, 4, 5, 6	0.82 (0.72-0.90)	<.001	≤ 2	87.8	70
EASL×WHO	0, 1, 2, 3, 4, 6, 9	0.85 (0.75-0.92)	0.002	≤ 1	90.2	67.5
EASL+(2×WHO)	0, 1, 2, 3, 4, 5, 6, 7, 8, 9	0.79 (0.68-0.87)	<.001	≤ 4	87.8	70
(2×EASL)+WHO	0, 1, 2, 3, 4, 5, 6, 7, 8, 9	0.85 (0.76-0.92)	<.001	≤ 3	90.2	67.5

* P value demonstrates difference of the curve to an AUC of 0.5 (by chance)