Published in final edited form as:

J Vasc Interv Radiol. 2010 April ; 21(4): 515–521. doi:10.1016/j.jvir.2009.11.020.

The Agreement between Competing Imaging Measures of Response of Hepatocellular Carcinoma to Yttrium-90 Radioembolization

Eugene Duke, BS¹, Jie Deng, MS^{1,4}, Saad M. Ibrahim, BS¹, Robert J. Lewandowski, MD¹, Robert K. Ryu, MD¹, Kent T. Sato, MD¹, Frank H. Miller, MD¹, Laura Kulik, MD^{1,2}, Mary F. Mulcahy, MD^{1,3}, Andrew C. Larson, PhD^{1,4}, Riad Salem, MD, MBA^{1,3}, and Reed A. Omary, MD, MS^{1,4}

¹ Department of Radiology, Northwestern University, Chicago, IL 60611

² Department of Medicine, Division of Hepatology, Northwestern University, Chicago, IL 60611

³ Department of Medicine, Division of Oncology, Northwestern University, Chicago, IL 60611

⁴ Department of Biomedical Engineering, Northwestern University, Chicago, IL 60611

Abstract

Purpose—There are currently at least 6 major competing criteria that are used to determine response to ⁹⁰Y and other liver-directed therapies including: a) RECIST, b) WHO, c) volumetric, d) 2D EASL, e) 3D EASL, f) functional diffusion weighted (DW) MR imaging. Our purpose was to evaluate the agreement between these competing tumor response classification schemes based on quantitative measurements of tumor size, necrosis, and changes in water mobility.

Materials/Methods—In this retrospective study, 20 HCC patients underwent ⁹⁰Y radioembolization. The patients' tumor burden before and 3–6 months after treatment was assessed with MRI. The percent change in size of tumors was used to classify patients into response categories. Kappa and agreement statistics were used to compare the concordance between the different criteria.

Results—Conventional size criteria (RECIST, WHO, and volumetric) all had a substantial level of agreement (kappa statistics from 0.76 to 0.78) when classifying patients into response categories. However, the conventional size criteria in relation to either 2D or 3D EASL had only slight to moderate concurrence with kappa statistics as low as 0.06. 2D EASL criteria and functional DW MR imaging resulted in the highest response rates, 11/20 (55%) and 15/20 (75%) respectively, while conventional size criteria produced lower response rates.

Conclusions—Classification of HCC response to ⁹⁰Y radioembolization was related to which of the competing criteria are used. We recommend that anatomic imaging criteria be used as the primary

Address Correspondence to: Reed A Omary, MD, MS, Department of Radiology, Northwestern University, 737 N. Michigan Ave, Suite 1600, Chicago, IL 60611, reed@northwestern.edu, Tel: (312)-695-3774, Fax: (312)-926-0826.

Presented at the 2008 annual SIR meeting, Washington, D.C. March 15–20th.

Disclosures: ED received funding from the Medical Student Summer Research Program Grant – Feinberg School of Medicine (Northwestern University). MFM receives research support from MDS Nordion. RS is a consultant for MDS Nordion. RAO partially funded by NIH 1 R01 CA126809-01A2. None of the authors have identified a conflict 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.

method to determine response and functional imaging criteria be used as a complementary secondary method.

INTRODUCTION

Yttrium-90 radioembolization (^{90}Y) is an effective treatment for unresectable hepatocellular carcinoma (HCC) (1,2). ^{90}Y is a form of brachytherapy in which ^{90}Y isotopes are delivered through a catheter directly into the hepatic arteries which supply the tumor(s).

Accurate imaging assessment of tumor burden is important in the clinical management of patients with HCC because underestimation of response to treatments could have detrimental effects on patient care. For example, underestimating response could preclude a patient from getting a liver transplant if ^{90}Y is being used as a bridging therapy or could subject patients to unnecessary ionizing radiation through extra rounds of treatment. However, there are currently at least six major competing criteria that are used to determine response to ^{90}Y and other liver-directed therapies that can confound assessment of response. Five of the imaging criteria are anatomical, including: a) uni-dimensional measurement known as the Response Evaluation Criteria in Solid Tumors (RECIST) guidelines (3), b) bi-dimensional measurement established by the World Health Organization (WHO) (4), c) volumetric measurement, d) bi-dimensional measurement taking into account reduction in viable tumor size (based on the amount of necrosis) as proposed by the European Association for the Study of the Liver (EASL) (5), and e) volumetric viable tumor size (3D EASL). The sixth method of measuring response is functional diffusion-weighted (DW) magnetic resonance imaging (MRI) assessment based upon changes in the mobility of water (6).

Several studies have shown that the classification of therapeutic response is potentially related to which criteria are chosen to measure tumor burden (7,8). However, a rigorous comparison of these competing methods is lacking. Moreover, while recent studies agree upon the benefit of incorporating necrosis into the assessment of tumor burden (9,10), there have been relatively few studies that have included necrosis. When EASL proposed adding a necrosis criterion to evaluate response of HCC to therapies, they did not describe a detailed method to measure necrosis and their recommendation was based upon qualitative observation rather than quantitative data (5). Since EASL's recommendation, a uniform method to measure necrosis has not yet been described.

In the clinical setting, necrosis is often subjectively estimated which can result in inter-observer variability. Recently, Keppke et al. compared RECIST, WHO, and volumetric necrosis criteria to assess therapeutic response of HCC to ^{90}Y (11). However, they used computed tomography (CT) as their primary imaging modality which has reduced sensitivity to detect HCC compared to MRI (12) and does not offer the capacity to detect functional changes in water mobility as does DW-MRI. Miller et al. also compared RECIST, WHO, and volumetric necrosis criteria, correlating their results to a functional imaging measurement on PET (13). However, their study also used CT instead of MRI, and they focused on liver metastases instead of primary HCC.

The purpose of this study was to evaluate the agreement of the six major competing methods of assessing HCC response to ^{90}Y therapy using quantitative measurements of tumor size, necrosis, and changes in water mobility.

MATERIALS AND METHODS

Clinical Setting and Patients

In this retrospective Institutional Review Board-approved study, we analyzed 20 consecutive patients with a primary diagnosis of HCC treated with ^{90}Y radioembolization who met strict inclusion criteria. We selected consecutive patients treated in a single urban, academic tertiary care center between July 2005 and April 2007 who met the inclusion criteria. Patient demographics are summarized in Table 1. Inclusion criteria for this study were: a) a diagnosis of HCC by fine-needle aspiration or core biopsy or by two imaging studies (CT and MRI) with increased alpha-fetoprotein levels (>400 ng/mL), b) Eastern Cooperative Oncology Group (ECOG) performance status of 0–2, c) non-infiltrating, focal HCC, and d) available baseline anatomic and diffusion weighted (DW) MR images obtained approximately one month prior and 3 months after ^{90}Y radioembolization. Exclusion criteria included: a) patients who did not have complete MR scans before and after therapy, b) patients with non-focal, infiltrating lesions as these tumors are inherently not measurable using anatomic criteria, and c) patients who underwent any liver-directed therapies other than ^{90}Y , such as TACE, since the time course of radiographic response after ^{90}Y and other therapies may differ.

^{90}Y Treatment

Prior to ^{90}Y radioembolization, each patient underwent hepatic angiography and $^{99\text{m}}\text{Tc}$ -macroaggregated albumin evaluation to identify anatomic variants and to exclude significant lung shunting, respectively (14). If significant shunting to extrahepatic organs was detected during angiography, then protective gastrointestinal arterial embolization was performed before ^{90}Y treatment. After these preliminary procedures, patients waited 3 to 7 days before ^{90}Y was administered.

Radioembolization was performed using TheraSphere (MDS Nordion, Kanata, ON, Canada) microspheres. This radioembolic agent was delivered to liver tumors via lobar, or when possible, segmental branches of the hepatic artery. The administration of ^{90}Y and the calculation method to determine required activity level (GBq) for TheraSphere treatment were performed according to previously defined protocols (15).

All procedures were performed by three Certificate of Added Qualification (CAQ)-certified interventional radiologists who specialize in interventional oncology, each with over 5 years of experience. After the procedure, patients were monitored and discharged the same day. Patients then returned for a follow-up visit at which time their tumor burden was reassessed using MRI.

MRI

MR scans were acquired approximately one month before (mean: 24 days, range: 1 to 61 days) and three months after ^{90}Y treatment (mean: 80 days, range: 19 to 150 days) on a 1.5T MR scanner (Avanto, Siemens Medical Solutions, Erlangen, Germany) using a flexible 6-channel phased array abdominal imaging coil. Anatomic MR imaging was obtained using T2-weighted half-Fourier acquisition single-shot turbo spin-echo and contrast-enhanced T1-weighted gradient echo imaging sequences with fat suppression in arterial, venous phases (45 to 60 and 90 seconds) and delayed phases (2 to 5 minutes after contrast injection) using the following parameters: TR/TE 120–160/1.9; flip angle 70° ; slice thickness/gap 6/1.8 mm; matrix size 125×256 ; rectangular field of view 30–40 cm; 23 slices acquired in breath-hold of 20 seconds. Gadopentetate dimeglumine (Magnevist, Berlex Pharmaceuticals) was administered intravenously via a power injector (Spectris, Medrad) at a dose of 0.1 mmol/kg, followed by 20 mL of saline flush.

DW MR imaging was acquired using single-shot spin-echo echo planar imaging during one or more breath-holds with the following parameters: TR/TE 2,500/82msec; slice thickness/gap 8/4 mm; bandwidth 1.5kHz/pixel; partial Fourier factor 6/8; non-selective fat saturation; twice refocused spin-echo diffusion weighting to reduce eddy current-induced distortion with b values of 0 and 500 sec/mm².

Image Analysis

All MR images were analyzed retrospectively on an Argus image processing workstation (Siemens). Only the largest tumor in each patient was assessed for treatment response. Two fellowship-trained attending interventional radiologists, in consensus, agreed upon all tumor measurements. The readers were blinded to all previous readings of the patients' images and to any other imaging studies that the patients might have had on file.

The five anatomical imaging criteria for measuring tumor size were performed on the contrast-enhanced T1-weighted axial images. The uni-dimensional measurements of tumor size were made in accordance with RECIST guidelines by measuring the largest diameter in the axial plane. Bi-dimensional measurements of lesions were obtained according to WHO guidelines by taking the cross product between a tumor's largest diameter in the transverse plane and its largest perpendicular diameter on the same image. Volumetric measurements were performed by tracing the lesions on all of the images where they appeared, adding the areas from each slice together and multiplying by the slice thickness to obtain a volume.

Necrosis was defined as non-enhancing tissue based on comparison with surrounding normal liver tissue on arterial phase T1 weighted images. Estimation of enhancement was subjective with even minimal enhancement being considered as viable tissue. For the bi-dimensional viable tumor volume criteria (2D EASL), tumor size was first measured according to WHO criteria. Next, the tumor necrotic area was measured by taking the cross product between the necrotic center's largest diameter and its largest perpendicular diameter on the same image. If a tumor contained more than one center of necrosis, the area of each was measured separately and all of the necrotic areas were summed. Finally, the necrotic area cross product was subtracted from the WHO criteria cross product to yield a bi-dimensional viable tumor volume measurement (2D EASL).

The volumetric viable tumor volume was calculated by first measuring tumor size using standard volumetric size criteria described above. Next, the necrotic areas were traced on all images where they appeared and the necrotic areas from all of the slices were summed. The summed areas of necrosis were multiplied by the slice thickness to obtain the necrotic volume. Finally, the necrotic volume was subtracted from the total tumor volume to obtain a viable tumor volume (3D EASL).

After treatment, response was evaluated using these five anatomic imaging criteria (RECIST, WHO, volumetric, 2D EASL, and 3D EASL). We calculated percent change in size for each criterion by comparing baseline MRI with the 3-month follow-up MRI. Response was classified according to Table 2 as complete response, partial response, stable disease, or progressive disease (3).

For the functional imaging assessment of tumor response, apparent diffusion coefficient (ADC) maps were reconstructed from each series of DW images. The contrast-enhanced T1-weighted scans were used as a reference to draw a region of interest around the tumor on the corresponding DW MR scans. The region of interest was then transferred to the ADC map and a mean ADC value for the tumor was determined. As with the size criteria, ADC values were measured both on the baseline and 3 months post ⁹⁰Y MRI scans. Response by functional criteria was assessed using previously published criteria of Rhee et al. (16). An increase in

ADC of greater than 5% was considered a response while a decrease in ADC of greater than 5% was considered disease progression. Changes of less than 5% were considered — no change in ADC and were classified as stable disease based on DWMR.

Statistical Analysis

We calculated mean and standard deviations for tumor size and tumor ADC values, before and after radioembolization for each of the 6 techniques. The means before and after ^{90}Y radioembolization were compared using a two-tailed paired t-test ($\alpha = 0.05$).

The five anatomical imaging size measurements were compared using agreement, which was defined as the percentage of cases in which two different criteria classified the patient into the same response category. Cohen's kappa statistics $\pm$ standard errors were also calculated to compare how well each of the six criteria concurred in their respective sorting of patients into response categories (17). Additionally, a two-tailed t-test ($\alpha = 0.05$) was used to evaluate whether the mean ADC values for patients classified as responders (complete or partial response by 2D EASL) was different than for the non-responders (stable or progressive disease by 2D EASL).

RESULTS

The number of patients in each response category measured by the five anatomical imaging size criteria is reported in Table 3. The response rate (patients classified as either complete or partial response) was 2/20 patients (10%) using RECIST, 2/20 patients (10%) by WHO, 2/20 patients (10%) by volumetric size, 11/20 patients (55%) by 2D EASL, and 4/20 patients (20%) by 3D EASL. The number of patients that were classified as having progressive disease was 2/20, 4/20, and 4/20 patients using RECIST, WHO, and volumetric criteria respectively. In comparison, no patients were classified as having progressive disease with either 2D or 3D EASL.

Mean tumor size measurements and mean ADC values are summarized in Table 4. From the anatomic contrast-enhanced T1-weighted MR images, tumor size did not significantly change when measured by RECIST (6.2 ± 3.5 cm to 5.9 ± 2.7 cm; $p=0.444$), WHO (41.9 ± 54.9 cm² to 33.6 ± 29.5 cm²; $p=0.300$), volumetric (2064 ± 4491 cm³ to 1333 ± 2127 cm³; $p=0.354$), or 3D EASL (1800 ± 3837 cm³ to 745 ± 1211 cm³; $p=0.192$) criteria. There was a statistically significant decrease in tumor size from 37.2 ± 46.5 cm² to 14.9 ± 12.8 cm² ($p=0.035$) when the tumor was measured by 2D EASL. Mean tumor ADC increased from 1.51×10^{-3} mm²/sec $\pm$ 0.28 at baseline to 1.96×10^{-3} mm²/sec $\pm$ 0.38 at 3 months post radioembolization. This increase was statistically significant with $p=0.001$.

Agreement data and Cohen's kappa statistic values are summarized in Table 5. The conventional anatomic imaging size criteria -- RECIST, WHO, and volumetric-- all have a substantial level of agreement when classifying patients into response categories. The agreements between RECIST and WHO, RECIST and volumetric, and between WHO and volumetric are all 18/20 (90%). When 2D EASL criteria was related to 3D EASL criteria, the concurrence for classifying patients into response categories was fair, with an agreement of 13/20 (65%) and $\kappa = 0.38 \pm 0.23$. In contrast, comparing the conventional size criteria to 2D EASL and/or 3D EASL measurements produced results with only slight to moderate concurrence. For example, there was only 8/20 (40%) agreement and $\kappa = 0.06 \pm 0.25$ when volumetric measurements were related to 2D EASL measurements. Figure 1 shows an example of a tumor that was classified as complete response by 2D and 3D EASL, but was stable by RECIST, WHO, and volumetric guidelines.

Changes in ADC values for all patients are shown in Figure 2. An increase in ADC values following ^{90}Y was seen in 15/20 patients, while 3/20 patients (15%) had no change in the ADC, and 2/20 patients (10%) had a decreased ADC. There was no statistical difference between the mean ADC values for patients classified as responders and non-responders by 2D EASL ($p=0.32$).

DISCUSSION

In our study, using rigorous quantitative analysis and 6 different response criteria, we found that the classification of HCC response to ^{90}Y radioembolization highly depended upon the criteria chosen to assess tumor response. Three months after ^{90}Y treatment, bi-dimensional viable tumor analysis that incorporated necrosis (2D EASL) was the only anatomic imaging size criterion showing a statistically significant reduction in tumor size. The remaining 4 tested anatomic imaging criteria did not show any statistically significant changes in tumor size. Although 3D-EASL did not reach a statistically significant decrease in size after ^{90}Y , the decrease in size could be clinically significant because the average patient had >50% decrease in viable tumor volume. With more patients in this study, it is possible that 3D EASL would have reached statistical significance.

At this same time period, the functional imaging measure of tumor response — an increase in ADC value on DW imaging — demonstrated statistically significant changes. Unlike a previous study which found an inverse relationship between changes in ADC values and anatomical size changes after ^{90}Y (16), our study shows that the ADC value increased after treatment for the majority of patients and that there was no correlation between a decrease in tumor size and an increase in the ADC value. The most likely explanation for this discrepancy is that the functional changes are not mirrored by the anatomical changes. However, correlation with histopathology is needed to confirm this hypothesis. From these results we believe that a functional imaging measurement of tumor burden, based upon increases in water mobility from the breakdown of necrotic cellular membranes, can serve as a separate but complementary assessment of tumor response.

We also showed that the traditional size criteria (RECIST, WHO, volumetric) had substantial agreement ($\kappa=0.76$) when classifying tumors into treatment response categories. This outcome supports previous studies showing that volumetric tumor measurements do not add significant value to those obtained by RECIST and WHO (18,19). However, our results diverge from other studies showing that volumetric measurements yield discordant and possibly more accurate results compared to those obtained by RECIST (uni-dimensional) and WHO (bi-dimensional) measurements (7,8). Given the time consuming effort required to measure tumor volume and the similar results compared to RECIST and WHO, we do not think that 3D volumetric analysis is applicable to daily clinical practice, unless software is designed to make the process more automated.

Additionally, we demonstrated that the three traditional anatomic imaging criteria (RECIST, WHO, and volumetric) underestimate response compared to the EASL anatomic criteria which incorporate necrosis. Specifically, the traditional size criteria had only slight to fair agreement ($\kappa = 0.06$ to 0.42) with 2D and 3D EASL and showed a less favorable treatment response. We found that the 2D EASL criteria produced the most favorable response to ^{90}Y . Additionally, employing 3D EASL criteria proved extremely time consuming and yielded similar results to the simpler, faster 2D EASL criteria. The fact that 3D EASL had a lower response rate compared to 2D EASL (55% vs. 20%) can be attributed to the phenomenon of necrosis typically occurring at the center of tumor before the edges. In other words, most 2D EASL measurements were made on the central slice of the tumor where necrosis was greatest. However, once we moved away from the central slices to obtain a volume, there was progressively more viable

tumor noted towards the periphery of the lesions. Additionally, there are no universally agreed upon criteria to classify response using the 3D EASL method because EASL never specified how necrosis should be defined or measured. Previously, Keppke et al. and Miller et al. used a 30% increase in necrosis to define partial response (11,13). They chose this value because it was consistent with the 30% decrease in size required by RECIST method to obtain partial response. In this manuscript we used a more rigorous 65% decrease in viable tumor in order to define response. This value is extrapolated from the 65% decrease in tumor volume required by the volume criteria (3). Had we used a 30% decrease in viable tumor volume, our response by 3D EASL would have approached response as measured by 2D EASL.

Our findings are similar to the results of Keppke et al and Miller et al (11,13) who showed that WHO and RECIST underestimated response rates compared to criteria which take tumor necrosis into account. However, unlike these two previous studies, we used MRI as our imaging modality instead of CT. This study further differed from Keppke et al (11) because we correlated tumor size with a functional measurement of tumor burden and from Miller et al (13) because we focused on HCC rather than liver metastases.

Our study has several important limitations. First, while necrosis was estimated as non-enhancing tissue for the anatomical imaging criteria, —non-enhancement may not accurately differentiate viable from necrotic tumor on histopathology. Imaging response was not correlated with histopathology because tissue sampling was not clinically warranted. Second, results were assessed only for single dominant tumors with well-defined margins. Whether the results can be generalized to multiple tumors or those without well-defined borders remains to be shown by future studies. Third, we assessed response to radioembolization and not other liver-directed therapies, such as chemoembolization or radiofrequency ablation. However, because the methodology is the same, we see no reason why the results would not be applicable to all liver-directed therapies, or even systemic chemotherapies. Fourth, there was variability in the number of days post ⁹⁰Y that the follow up MRI was obtained. In order to eliminate this variability, while maintaining 20 patients, we would have needed additional patients treated prior to July 2005 when DW MR images were not routinely available at our institution. Finally, we used a time-consuming rigorous method to outline and quantify necrotic volume. In clinical practice, necrosis is typically estimated as regions lacking enhancement. Although it remains to be shown how the results would occur in clinical practice, the use of bi-dimensional viable criteria appears most clinically applicable.

We acknowledge the potential for bias with parts of the study design such as lesion selection, using consensus readings, or choosing readers who are familiar with the disease process and ⁹⁰Y therapy. However, the goal of this study was not to establish the amount of necrosis caused by ⁹⁰Y or to evaluate intra- or inter-observer variability. Instead, it was to show how different imaging criteria leads to different response rates. We selected patients with single or dominant lesions with well defined margins because measuring these lesions would offer the highest likelihood for reproducibility between readers. Also, by definition, lesions without measurable margins cannot be assessed using RECIST or WHO imaging criteria.

In conclusion, of the 5 competing anatomic imaging analysis approaches, bi-dimensional viable tumor size (2D EASL) criteria provide the greatest anatomic response to radioembolization. However, in the absence of correlative histopathology or longitudinal studies, we can only speculate that the 2D EASL approach is the preferred anatomic approach, while traditional size criteria underestimate response. For functional assessment of water mobility changes, DW MR imaging can serve as a complementary secondary method to assess response. However, it is not a substitute for primary anatomic measurements because we did not find a correlation between these methods. Correlation with clinical outcome is the subject of future study.

Acknowledgments

The authors would like to thank Ann Ragin for her critical reading of the manuscript and the nursing members of the Interventional Oncology team at Northwestern Memorial Hospital for their dedicated efforts in patient care (Karen Marshall, Sharon Coffey, Krystina Sajdak, Margaret Gilbertsen).

References

1. Kulik LM, Carr BI, Mulcahy MF, et al. Safety and efficacy of 90Y radiotherapy for hepatocellular carcinoma with and without portal vein thrombosis. *Hepatology* 2008;47:71–81. [PubMed: 18027884]
2. 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:S223–229. [PubMed: 12354840]
3. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 2000;92:205–216. [PubMed: 10655437]
4. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981;47:207–214. [PubMed: 7459811]
5. 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;35:421–430. [PubMed: 11592607]
6. Deng J, Miller FH, Rhee TK, et al. Diffusion-weighted MR imaging for determination of hepatocellular carcinoma response to yttrium-90 radioembolization. *J Vasc Interv Radiol* 2006;17:1195–1200. [PubMed: 16868174]
7. Prasad SR, Jhaveri KS, Saini S, Hahn PF, Halpern EF, Sumner JE. CT tumor measurement for therapeutic response assessment: comparison of unidimensional, bidimensional, and volumetric techniques initial observations. *Radiology* 2002;225:416–419. [PubMed: 12409574]
8. Hopper KD, Kasales CJ, Eggli KD, et al. The impact of 2D versus 3D quantitation of tumor bulk determination on current methods of assessing response to treatment. *J Comput Assist Tomogr* 1996;20:930–937. [PubMed: 8933793]
9. Llovet JM, Beaugrand M. Hepatocellular carcinoma: present status and future prospects. *J Hepatol* 2003;38 (Suppl 1):S136–149. [PubMed: 12591191]
10. Sato K, Lewandowski RJ, Bui JT, et al. Treatment of unresectable primary and metastatic liver cancer with yttrium-90 microspheres (TheraSphere): assessment of hepatic arterial embolization. *Cardiovasc Intervent Radiol* 2006;29:522–529. [PubMed: 16729228]
11. Keppke AL, Salem R, Reddy D, et al. Imaging of hepatocellular carcinoma after treatment with yttrium-90 microspheres. *AJR Am J Roentgenol* 2007;188:768–775. [PubMed: 17312067]
12. Burrel M, Llovet JM, Ayuso C, et al. MRI angiography is superior to helical CT for detection of HCC prior to liver transplantation: an explant correlation. *Hepatology* 2003;38:1034–1042. [PubMed: 14512891]
13. Miller FH, Keppke AL, Reddy D, et al. Response of liver metastases after treatment with yttrium-90 microspheres: role of size, necrosis, and PET. *AJR Am J Roentgenol* 2007;188:776–783. [PubMed: 17312068]
14. 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;17:1251–1278. [PubMed: 16923973]
15. Dancy JE, Shepherd FA, Paul K, et al. Treatment of nonresectable hepatocellular carcinoma with intrahepatic 90Y-microspheres. *J Nucl Med* 2000;41:1673–1681. [PubMed: 11037997]
16. Rhee TK, Naik NK, Deng J, 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;19:1180–1186. [PubMed: 18656011]
17. Cohen J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* 1960;20:37–46.

18. Sohaib SA, Turner B, Hanson JA, Farquharson M, Oliver RT, Reznick RH. CT assessment of tumour response to treatment: comparison of linear, cross-sectional and volumetric measures of tumour size. *Br J Radiol* 2000;73:1178–1184. [PubMed: 11144795]
19. Dachman AH, MacEneaney PM, Adedipe A, Carlin M, Schumm LP. Tumor size on computed tomography scans: is one measurement enough? *Cancer* 2001;91:555–560. [PubMed: 11169938]

Figure 1. 59 year-old male with primary HCC

- a) Arterial phase contrast-enhanced, T1-weighted MRI with fat suppression, obtained before ^{90}Y shows a focal HCC (arrow) within the lateral segment of the left lobe of the liver.
- b) Arterial phase contrast-enhanced, T1-weighted MRI with fat suppression, obtained after ^{90}Y therapy shows that the tumor (arrow) has not significantly changed in size when measured by RECIST, WHO, or volumetric criteria, but that is now completely necrotic by 2D EASL and 3D EASL criteria.

Figure 2.

Functional DW MRI assessments based upon changes in mobility of water after ⁹⁰Y therapies are shown. An increase in ADC values is seen for patients 1–7, 9, 13–24, and 16–20, no change in ADC values is seen for patients 10–12, and a decrease in ADC is seen for patients 8 and 15.

Table 1

Patient Demographics

Age (years)		
	Mean	65
	Range	50–80
Gender		Number of Patients
	Males	16
	Females	4
Etiology		
	Hepatitis C	8
	Cryptogenic	6
	Alcohol	4
	Hepatitis B	1
	Autoimmune Hepatitis	1
ECOG status		
	ECOG 0	16
	ECOG 1	3
	ECOG 2	1
Child Pugh		
	A	9
	B	11
Liver Burden		
	0–25%	18
	26–50%	1
	50–75%	1

Table 2

Treatment Response Categories and Tumor Measurement Techniques

Category	RECIST	WHO	Volumetric	Bi-dimensional viable size (2D EASL)	Volumetric viable size (3D EASL)
Complete response	Tumor disappearance	Tumor disappearance	Tumor disappearance	Tumor disappearance	Tumor disappearance
Partial response	>30% reduction in diameter	>50% reduction in cross product	>65% reduction in volume	>50% reduction in cross product	>65% reduction in volume
Stable disease	Size between that for partial response and progressive disease	Size between that for partial response and progressive disease	Size between that for partial response and progressive disease	Size between that for partial response and progressive disease	Size between that for partial response and progressive disease
Disease progression	>20% increase in diameter	>25% increase in cross product	>44% increase in volume	>25% increase in cross product	>44% increase in volume

Table 3

Number of Patients in Treatment Response Categories (n = 20 patients)

	Complete	Partial	Stable	Progression
RECIST	0	2	16	2
WHO	0	2	14	4
VOL	0	2	14	4
2D EASL	1	10	9	0
3D EASL	1	3	16	0

Table 4

Tumor Response Post 90Y Therapy Based on MRI and ADC Measurements

	Baseline	3 month Post 90Y	P value
RECIST	6.2±3.5 cm	5.9 ± 2.7 cm	0.444
WHO	41.9± 54.9 cm ²	33.6±29.5 cm ²	0.300
Volumetric	2064±4491 cm ³	1333±2127 cm ³	0.354
2D EASL	37.2±46.5 cm ²	14.9±12.8 cm ²	0.035
3D EASL	1800±3837 cm ³	745±1211 cm ³	0.192
ADC Value	1.51×10 ⁻³ mm ² /sec ± 0.28	1.96×10 ⁻³ mm ² /sec ± 0.38	0.001

Table 5

Agreement and Kappa Statistic Data (n = 20 patients)

	Agreement	% Agreement	Cohen's Kappa $\pm$ SE
RECIST and WHO	18	90%	0.76 $\pm$ 0.16
RECIST and VOL	18	90%	0.76 $\pm$ 0.16
RECIST and 2D EASL	10	50%	0.15 $\pm$ 0.26
RECIST and 3D EASL	16	80%	0.42 $\pm$ 0.23
WHO and VOL	18	90%	0.78 $\pm$ 0.15
WHO and 2D EASL	8	40%	0.06 $\pm$ 0.25
WHO and 3D EASL	14	70%	0.29 $\pm$ 0.23
VOL and 2D EASL	8	40%	0.06 $\pm$ 0.25
VOL and 3D EASL	14	70%	0.29 $\pm$ 0.23
2D EASL and 3D EASL	13	65%	0.38 $\pm$ 0.23