

Radiologic-Pathologic Analysis of Contrast-enhanced and Diffusion-weighted MR Imaging in Patients with HCC after TACE: Diagnostic Accuracy of 3D Quantitative Image Analysis¹

Julius Chapiro, MD
 Laura D. Wood, MD, PhD
 MingDe Lin, PhD²
 Rafael Duran, MD
 Toby Cornish, MD, PhD
 David Lesage, PhD
 Vivek Charu, BS
 Rüdiger Scherthaner, MD
 Zhijun Wang, MD, PhD
 Vania Tacher, MD
 Lynn Jeanette Savic, BS
 Ihab R. Kamel, MD
 Jean-François Geschwind, MD

Purpose:

To evaluate the diagnostic performance of three-dimensional (3D) quantitative enhancement-based and diffusion-weighted volumetric magnetic resonance (MR) imaging assessment of hepatocellular carcinoma (HCC) lesions in determining the extent of pathologic tumor necrosis after transarterial chemoembolization (TACE).

Materials and Methods:

This institutional review board–approved retrospective study included 17 patients with HCC who underwent TACE before surgery. Semiautomatic 3D volumetric segmentation of target lesions was performed at the last MR examination before orthotopic liver transplantation or surgical resection. The amount of necrotic tumor tissue on contrast material-enhanced arterial phase MR images and the amount of diffusion-restricted tumor tissue on apparent diffusion coefficient (ADC) maps were expressed as a percentage of the total tumor volume. Visual assessment of the extent of tumor necrosis and tumor response according to European Association for the Study of the Liver (EASL) criteria was performed. Pathologic tumor necrosis was quantified by using slide-by-slide segmentation. Correlation analysis was performed to evaluate the predictive values of the radiologic techniques.

Results:

At histopathologic examination, the mean percentage of tumor necrosis was 70% (range, 10%–100%). Both 3D quantitative techniques demonstrated a strong correlation with tumor necrosis at pathologic examination ($R^2 = 0.9657$ and $R^2 = 0.9662$ for quantitative EASL and quantitative ADC, respectively) and a strong intermethod agreement ($R^2 = 0.9585$). Both methods showed a significantly lower discrepancy with pathologically measured necrosis (residual standard error [RSE] = 6.38 and 6.33 for quantitative EASL and quantitative ADC, respectively), when compared with non-3D techniques (RSE = 12.18 for visual assessment).

Conclusion:

This radiologic-pathologic correlation study demonstrates the diagnostic accuracy of 3D quantitative MR imaging techniques in identifying pathologically measured tumor necrosis in HCC lesions treated with TACE.

© RSNA, 2014

Online supplemental material is available for this article.

¹From the Russell H. Morgan Department of Radiology and Radiological Science, Division of Vascular and Interventional Radiology, The Johns Hopkins Hospital, Sheikh Zayed Tower, Suite 7203, 1800 Orleans St, Baltimore, MD 21287 (J.C., M.L., R.D., R.S., Z.W., V.T., L.J.S., I.R.K., J.F.G.); Department of Pathology, The Johns Hopkins Hospital, Baltimore, Md (L.D.W., T.C.); Philips Research, Medisys, Suresnes, France (D.L.); and Johns Hopkins Bloomberg School of Public Health, Baltimore, Md (V.C.). Received January 16, 2014; revision requested March 11; revision received March 17; accepted April 21; final version accepted April 29. Supported by Philips Research North America, Briarcliff Manor, NY, the French Society of Radiology (SFR), and the Rolf W. Günther Foundation for Radiological Science. Address correspondence to J.F.G. (e-mail: jfg@jhmi.edu).

²Current address: Department of Clinical Informatics, Interventional, and Translational Solutions, Philips Research North America, Briarcliff Manor, NY.

Hepatocellular carcinoma (HCC) is the sixth most common neoplasm in the world (1). With more than 700 000 newly diagnosed cases per year and a dismal prognosis, it continues to be the leading cause of cancer-related death in patients with liver cirrhosis (2,3). In many cases, imaging-guided intraarterial therapies such as transarterial chemoembolization (TACE) represent the only therapeutic option for controlling the disease and improving patient survival rates (4,5). Although overall survival is the ultimate end point in cancer research, other imaging-based surrogate end points, such as time to progression and tumor response rates, have become indispensable for both clinical trials and everyday therapeutic decisions (6,7). Most intraarterial therapies involve the element of embolization to induce tumor infarction, which leads to tissue necrosis without immediate effects on the overall lesion size. Because of that, anatomic imaging biomarkers such as Response Evaluation Criteria in Solid Tumors and World Health Organization criteria have repeatedly failed to accurately identify pathologic tumor response at cross-sectional imaging in

patients treated with these techniques (8–12). Thus, measurements of enhancing tumor tissue on the basis of European Association for the Study of the Liver (EASL) guidelines as well as modified Response Evaluation Criteria in Solid Tumors evolved as important parameters for quantitative tumor response assessment (1,6).

A strong correlation of imaging-based measurements of tumor response to TACE with the standard of reference of tumor pathology would provide a crucial foundation to link radiologic findings with actual therapeutic effects. This validation step can then help substantiate the clinical use of the previously mentioned methods. As such, numerous radiologic-pathologic studies were conducted with the goal of comparing enhancement-based uni- and bidimensional (one- and two-dimensional) criteria with pathologically estimated tumor necrosis (8,13–15). An important benefit of these techniques is the simplicity and reproducibility of measurements, specifically when using modified Response Evaluation Criteria in Solid Tumors (16). However, by nature of one- and two-dimensional measurements, these methods assume that three-dimensional (3D) tumor growth or response to treatment occurs in a symmetric, spheric manner. This assumption is not representative of asymmetric changes in tumor progression and thus cannot fully represent

tumor biology (9). The advent of new 3D methods of tumor assessment addresses the discordance between lesion diameter and nonspheric volume of the viable tumor tissue (17). On the basis of semiautomated or automated tumor segmentation, these volumetric techniques include enhancement-based and diffusion-weighted magnetic resonance (MR) imaging and are predictive of patient survival (18,19). However, to our knowledge, no study has addressed the correlation of radiologic and pathologic findings for the evolving 3D techniques.

We performed this retrospective study to evaluate the diagnostic accuracy of 3D quantitative enhancement-based and diffusion-weighted volumetric MR imaging assessment of HCC lesions in determining the extent of pathologic tumor necrosis after TACE. The aim of the radiologic-pathologic correlation was to confirm the diagnostic value of

Advances in Knowledge

- Semiautomated, three-dimensional (3D) quantitative assessment of hepatocellular carcinoma lesions at MR imaging after transarterial chemoembolization (TACE) shows high overall accuracy (100%, 17 of 17 patients) in the prediction of pathologic tumor necrosis.
- Enhancement-based and apparent diffusion coefficient–based MR imaging techniques are predictive of pathologic tumor necrosis ($R^2 = 0.9657$ and $R^2 = 0.9662$ for quantitative European Association for the Study of the Liver [qEASL] criteria and quantitative apparent diffusion coefficient [qADC], respectively) and show a high intermethod agreement ($R^2 = 0.9585$) when used for 3D quantitative analysis.

Implication for Patient Care

- The high accuracy and intermethod agreement of 3D quantitative techniques in the assessment of tumor necrosis after TACE is clinically relevant and underlines the shift away from one- and two-dimensional tumor response criteria and toward automated and semiautomated 3D approaches; the high diagnostic performance of qEASL criteria and qADC might have implications regarding clinical decisions as to whether a patient should undergo repeat treatment after a TACE session.

Published online before print

10.1148/radiol.14140033 Content code: GI

Radiology 2014; 273:746–758

Abbreviations:

ADC = apparent diffusion coefficient
 CI = confidence interval
 EASL = European Association for the Study of the Liver
 HCC = hepatocellular carcinoma
 qADC = quantitative ADC
 qEASL = quantitative EASL
 ROI = region of interest
 RSE = residual standard error
 3D = three-dimensional
 TACE = transarterial chemoembolization

Author contributions:

Guarantors of integrity of entire study, J.C., J.F.G.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; literature research, J.C., L.D.W., M.L., R.D., Z.W., L.J.S., I.R.K.; clinical studies, J.C., M.L., R.D., Z.W., V.T., L.J.S., I.R.K., J.F.G.; experimental studies, J.C., L.D.W., M.L., R.D., T.C., D.L., Z.W., L.J.S., I.K.; statistical analysis, J.C., T.C., V.C., Z.W., L.J.S.; and manuscript editing, J.C., L.D.W., M.L., R.D., D.L., V.C., R.S., Z.W., V.T., L.J.S., I.K., J.F.G.

Funding:

This research was supported by the National Institutes of Health (grants R01 CA160771, P30 CA006973, and NCRRL1 RR 025005).

Conflicts of interest are listed at the end of this article.

3D quantitative assessment techniques, which we will herein refer to as quantitative EASL (qEASL) and quantitative apparent diffusion coefficient (qADC).

Materials and Methods

Philips provided the software graphical user interface for the in-house 3D therapy assessment software that was developed at Johns Hopkins. M.L. and D.L. are employees of Philips. The control of the data and the information in the article were maintained by the remaining authors. No other conflicts of interest exist.

Study Cohort

This retrospective single-institution study was conducted in compliance with the Health Insurance Portability and Accountability Act and approved by the institutional review board. The design of the study was in agreement with the Standards for Reporting of Diagnostic Accuracy guidelines (20). Patients treated with TACE between December 2005 and December 2012 were considered eligible for inclusion if the diagnosis of HCC was confirmed by means of biopsy or accepted radiologic findings, as described in the EASL guidelines (4). Only patients who underwent TACE (conventional TACE, TACE with drug-eluting beads) as the sole bridging therapy before orthotopic liver transplantation or surgical resection were further considered ($n = 82$). This group of patients was subjected to conservative exclusion criteria as summarized in Figure 1. Specifically, we excluded patients who received systemic or local-regional liver-targeted therapies other than TACE ($n = 12$) and those who did not undergo follow-up MR imaging within 90 days before surgical treatment ($n = 9$). In addition, patients with severe motion artifacts at MR imaging that affected the target tumor regions were further excluded ($n = 5$). Last, patients with large, pathologically nonassessable tumors (diameter >7.5 cm) as well as those with ambiguously marked and incomplete pathologic samples were also excluded ($n = 39$). The main premise of the applied exclusion

criteria was to achieve a narrow selection of patients specifically treated with TACE before surgical treatment. Thus, crossover patients were avoided to exclusively evaluate TACE-related tumor necrosis. An additional emphasis was on including patients with entirely preserved and well-assessable pathologic tumor samples. All included patients had undergone dynamic contrast material-enhanced and diffusion-weighted MR imaging before the initial TACE procedure.

Evaluation and Staging

All included patients underwent baseline assessment that included laboratory tests to evaluate liver function (serum albumin level, prothrombin time, total bilirubin level, aspartate transaminase level, alanine aminotransferase level). Eastern Cooperative Oncology Group performance status was recorded, and the stage of disease was assessed in all patients by using Child-Pugh and Barcelona Clinic Liver Cancer classification systems (4). Table 1 summarizes the baseline characteristics of all patients included in this study. The mean patient age was 55.8 years (range, 45–76 years). Hemochromatosis was diagnosed in one patient (6%).

TACE Protocol

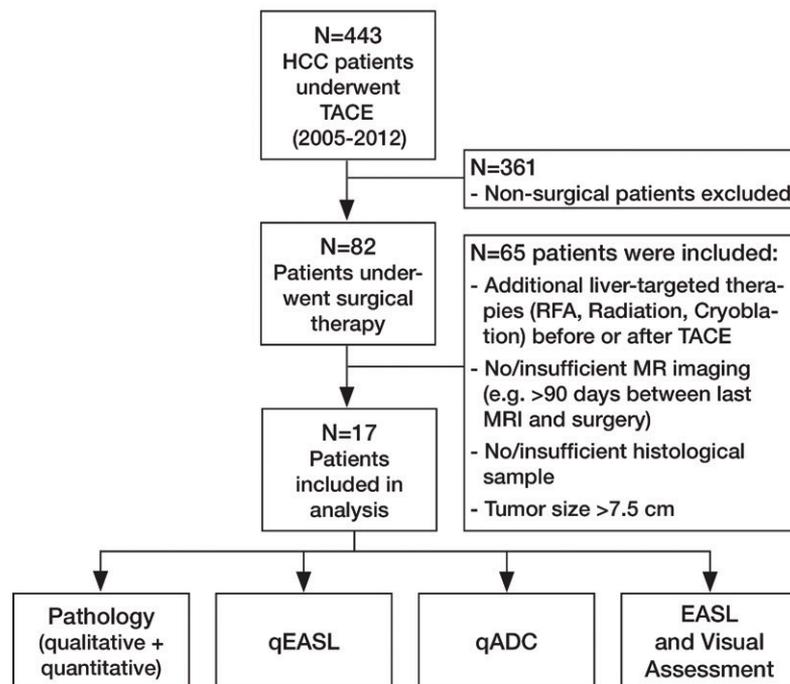
All TACE procedures were performed by one experienced interventional radiologist (J.F.G., with 16 years of experience in hepatic interventions). As described below, a consistent approach was used according to our standard institutional protocol. Initially, all patients underwent multiple angiographic steps to define the hepatic arterial anatomy and determine portal venous patency. For conventional TACE, patients were treated with selective (lobar or segmental) and superselective injections. A solution containing 50 mg of doxorubicin and 10 mg of mitomycin-C in a 1:1 mixture with iodized oil (Lipiodol; Guerbet, Aulnay-sous-Bois, France) was infused and followed by administration of 100–300- μ m-diameter microspheres (Embospheres; Merit Medical, South Jordan, Utah). Substantial arterial flow reduction to the tumor was

defined as the technical end point. For TACE with drug-eluting beads, patients were treated with selective to superselective injections. LC Beads (2 mL; BTG, Surrey, England) with a diameter of 100–300 μ m were loaded with 100 mg of doxorubicin hydrochloride (25 mg/mL) and mixed with an equal volume of nonionic contrast material. Up to 4 mL of drug-eluting beads was administered; however, complete occlusion was avoided to maintain the arterial pathway for repeat treatment. Additional details of the TACE protocol used in our institution have been previously described (21).

MR Imaging Technique

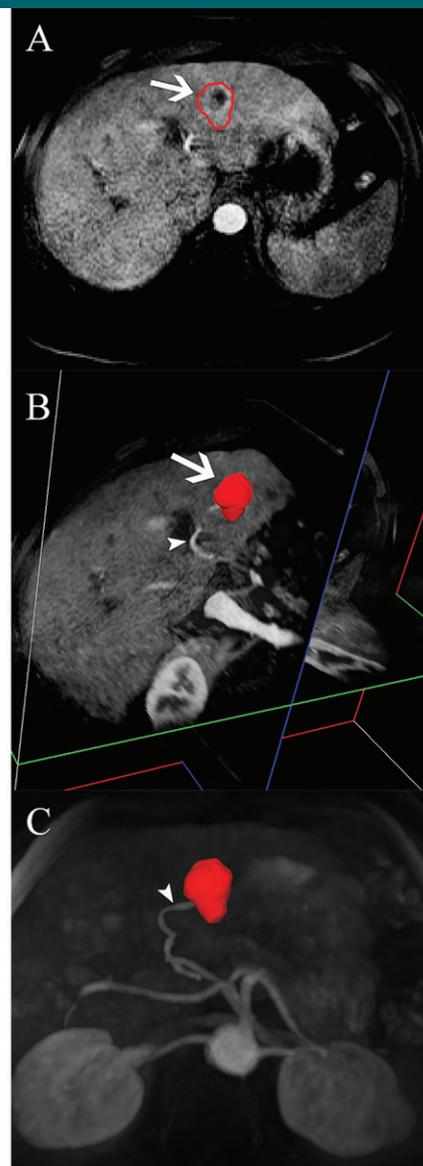
All patients included in this study underwent a standardized MR imaging protocol within 1 month before initial TACE. Follow-up MR imaging was performed 3–4 weeks after every procedure and within 90 days before surgical therapy. MR imaging was performed with a 1.5-T MR unit (Magnetom Avanto; Siemens, Erlangen, Germany) by using a phased-array torso coil. The protocol included breath-hold unenhanced and contrast-enhanced (0.1 mmol/kg intravenous gadopentetate dimeglumine [Magnevist; Bayer, Wayne, NJ]) T1-weighted 3D fat-suppressed spoiled gradient-echo imaging in the hepatic arterial phase (20 seconds after contrast material administration), portal venous phase (70 seconds after contrast material administration), and delayed phase (3 minutes after contrast material administration), as well as breath-hold diffusion-weighted echo-planar imaging. Reconstruction of ADC maps was performed by using a monoexponential fit between the two b values of 0 and 750 sec/mm² and the following equation: $S_b/S_0 = \exp(-b \times \text{ADC})$, where S_b is the MR signal strength in the presence of diffusion weighting and S_0 the MR signal strength in the absence of diffusion weighting (22,23). The high reproducibility of ADC measurements has been previously reported on and was not studied further (24). The detailed acquisition parameters are shown in Table E1 (online).

Figure 1



a.

Figure 1: (a) Flowchart of study design. Chart summarizes patient selection, exclusion criteria, and assessment techniques. Patients who received systemic or local-regional liver-targeted therapies other than TACE ($n = 12$) were excluded to avoid any non-TACE-related tumor response and/or necrosis. Patients who did not undergo follow-up MR imaging within 90 days before surgical treatment ($n = 9$) were excluded to avoid any potential discordance between imaging and tumor pathology owing to disease progression. Patients with severe motion artifacts on MR images affecting the target tumor regions were further excluded ($n = 5$). Only patients with well-defined index lesions smaller than 7.5 cm in largest diameter at baseline imaging were selected for further analysis. We excluded larger tumors ($n = 24$) because of the potential for bias in submission of representative pathologic sections of large tumors in this retrospective analysis. We also excluded patients in whom all pathologic slices of tumor were not available for review ($n = 15$). RFA = radiofrequency ablation. (b) Segmentation technique. A, Representative contrast-enhanced T1-weighted MR image demonstrates semiautomated tumor segmentation. Arrow = tumor rim. B, Three-dimensional model of upper abdomen shows volume rendering for segmented tumor. Arrow = tumor, arrowhead = branch of right hepatic artery that feeds tumor. C, Maximum intensity projection demonstrates contrast-enhanced blood vessels. Arrowhead = tumor-feeding artery.



b.

Imaging Data Evaluation

Subjective visual image analysis was performed by two radiologists (R.D. and V.T., with 7 and 4 years of experience in abdominal MR imaging, respectively) who were not involved in the TACE procedures, and a third radiologist (Z.W., with 10 years of experience in abdominal MR imaging) evaluated the lesions according to EASL guidelines (described in detail below). All evaluations were performed independently, and any remaining ambiguity

was resolved by consensus. In addition, a junior radiology resident (J.C.) performed the 3D image analysis using a software prototype (described below). He had 1 year of experience with the software.

Three-dimensional quantitative image analysis was performed with a software prototype (Medisys; Philips Research, Suresnes, France) (18). A target lesion was defined as the dominant lesion with the largest diameter treated during the first TACE session.

In accordance with previously mentioned inclusion criteria, one targeted lesion per patient was selected for analysis. The software used semiautomatic 3D tumor segmentation on the last arterial phase contrast-enhanced MR image obtained before orthotopic liver transplantation or surgical resection (Fig 1b; Appendix E1 [online]). The overall tumor volume was directly calculated on the basis of this segmentation (Fig 1b; Appendix E1 [online]). The resulting 3D segmentation

mask was then used for both arterial phase enhancement-based and diffusion-weighted quantitative analysis to achieve high intermethod agreement (8). The precision and reader-independent reproducibility of the utilized segmentation software has been previously described (25). Additional methodologic specifications for the segmentation technique and application are itemized in Appendix E1 (online). The qEASL calculation was based on image subtraction and followed the previously described algorithm (18,26). In brief, after creation of the 3D segmentation mask, the mask was then transferred onto the subtraction image and a region of interest (ROI) placed into extratumoral liver parenchyma as a reference to calculate the relative enhancement values within the tumor (Fig 2, A) (23). The patient-specific average signal intensity within the ROI was then defined as a threshold to estimate tumor necrosis within the 3D mask. Subsequently, nonenhancing regions were expressed as a percentage of the previously calculated overall tumor volume and visualized by using a color map overlay on the arterial phase MR image (Fig 2, A) (18). The qADC calculation used an analogous method. Additional methodologic specifications for the qADC and qEASL techniques as well as the ROI specifications are itemized within Appendix E1 (online).

In addition to the 3D quantitative techniques, subjective visual response assessment of target lesions was performed. Two readers (R.D. and V.T.) visually assessed the percentage of tumor enhancement on the last preoperative pre- and postcontrast MR images (arterial phase) in all included patients. The readers were blinded to the results of the 3D quantitative assessment. Areas of tumor enhancement were considered indicative of viable tumor tissue, and unenhanced areas were considered necrotic. The percentage of enhancement was recorded subjectively in 5% increments, ranging from no enhancement to 100% enhancement (27). Hence, the reciprocal percentage values were recorded to represent the amount of tumor

Table 1

Baseline Patient Characteristics	
Parameter	No. of Patients (n = 17)
Demographics	
Age	
<65 y	15 (88)
≥65 y	2 (12)
Sex	
M	15 (88)
F	2 (12)
Ethnicity	
Caucasian	10 (59)
African American	6 (35)
Asian	1 (6)
Etiology	
HBV	2 (12)
HCV	11 (65)
HBV + HCV	3 (18)
Alcohol	7 (41)
Cryptogenic	1 (6)
Hemochromatosis	1 (6)
Cirrhosis	
Present	15 (88)
Absent	2 (12)
Methods of diagnosis	
Biopsy	16 (94)
Imaging (MR imaging)	1 (6)
ECOG performance status of 0	17 (100)
Staging system	
BCLC class	
A	11 (65)
B	6 (35)
Child-Pugh class	
A	12 (71)
B	5 (29)
Tumor characteristics	
Distribution	
Unilobar	14 (82)
Bilobar	3 (18)
Multiplicity	
Unifocal	10 (59)
Bifocal	2 (12)
Multifocal	5 (29)
Size	
<3 cm	5 (29)
3–5 cm	11 (65)
>5 cm	1 (6)

Note.—Numbers in parentheses are percentages. BCLC = Barcelona Clinic Liver Cancer, ECOG = Eastern Cooperative Oncology Group, HBV = hepatitis B virus, HCV = hepatitis C virus.

Figure 2

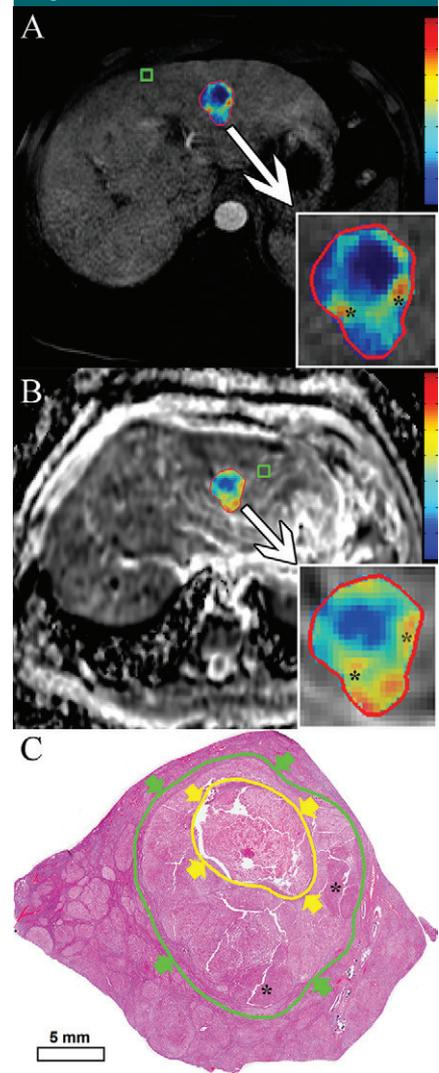


Figure 2: Quantitative assessment techniques. *A*, qEASL color map of tumor (red represents maximum enhancement and blue represents no enhancement and/or necrotic tumor tissue, normalized by the ROI). *B*, qADC color map of same lesion (with color coding as in *A*). *C*, Digital scan (hematoxylin-eosin stain) obtained with histosegmentation technique. Yellow arrows and circle indicate necrosis, green arrows and circle indicate entire tumor. Corresponding highly viable areas of tumor are shown on all images (*).

necrosis for each lesion. Only lesions treated with TACE and measured with the 3D quantitative method were included in the final analysis.

Finally, target lesions were assessed by another radiologist (Z.W.) with use of EASL guidelines. Multiple previously published studies identified the EASL guidelines to be strongly predictive of histopathologic tumor necrosis, showing higher sensitivity and specificity when compared with other response criteria (15,28). These findings prompted us to include EASL in our study with the purpose of comparing these two-dimensional criteria with our 3D quantitative technique. Response according to EASL criteria was defined in a conservative fashion, as follows: complete response, absence of any enhancing tissue in the target lesion; partial response, at least 50% decrease in the amount of enhancing tissue in the target lesion; stable disease, less than 50% decrease in the amount of enhancing tissue in the target lesion; and progressive disease, any increase in the amount of enhancing tissue in the target lesion that would lead to additional intraarterial therapy (13,14,27).

Pathologic Evaluation

The radiologically selected lesions were matched with the pathologic samples according to descriptions in the pathology reports by an experienced liver pathologist (L.D.W.) in the presence of the MR image readers to ensure that identical lesions were assessed at histopathologic examination in cases with multiple lesions ($n = 7$). The pathologist was blinded to all MR imaging findings. We obtained 0.5–1.0-cm slices of the treated index lesions according to our institutional protocol of serial liver sectioning. The orientation of sample sectioning was extracted from the available gross pathology reports and matched with the quantitative MR image analysis. Standard hematoxylin-eosin staining was used for all slides. Only fully or largely preserved HCC lesions were considered for further analysis. To our knowledge, most pre-existing radiologic-pathologic studies

used a tabulated approach for classifying histopathologic necrosis (8,13–15,26,28,29) and relied on a Gestalt-based (30) subjective evaluation rather than quantitative measurements. The rationale for this method lies in the nature of one- and two-dimensional quantitative treatment response criteria that also follow a tabulated approach according to the World Health Organization (4,11,14,31). To achieve a linear comparison with the specific numeric qEASL and/or qADC output, we followed a dual approach. First, the pathologist visually assessed and recorded the percentage of necrosis of the treated lesions in 10% increments, ranging from no necrosis (0%) to complete pathologic necrosis (100%). Then, all slides were scanned and digitalized at 20 \times magnification by using a high-spatial-resolution system (Aperio; Leica Biosystems Vista, Calif). The digitalized slides were then assessed with software (Aperio ImageScope). A manual slide-by-slide segmentation of necrotic as well as total tumor areas was performed by the same pathologist (Fig 2, C). The overall necrosis-to-tumor ratio was expressed as a calculated percentage on the basis of the areas of the identified necrotic and viable regions. Furthermore, tumor necrosis was tabulated in four pathologic classes: less than 25%, 25%–49%, 50%–74%, and 75%–100%. The tabulation was used for reasons of clarity when presenting the overall distribution of necrosis in our sample and furthermore facilitated the correlation with the EASL technique. No tabulation was used for the correlation with the 3D quantitative technique to avoid a data reduction.

Statistical Analysis

All statistical computations were performed with commercial software (Prism, version 6; GraphPad, San Diego, Calif). The summary of data was performed by using descriptive statistics. Count and frequency were used for categorical variables. Mean and range were used for continuous variables. Linear regression analysis was performed to investigate the correlation of results measured with

radiologic and pathologic techniques. The Pearson correlation coefficient (R) was calculated, and the coefficient of determination (R^2) was used to report the correlation of the methods. In addition, residual plots were used to assess drift, variance, and deviation in each radiologic technique when correlated with the standard of reference of tumor pathology. The residual standard error (RSE) was calculated to measure the discrepancy between the linear model predictions and the observed data (pathologic findings). Sensitivity, specificity, overall accuracy (32), positive predictive value, and negative predictive value were calculated for all radiologic methods in reference to the pathologic analysis for the tabular classes of tumor necrosis. For this purpose, responders (complete and partial response) and nonresponders (stable disease and progressive disease) were tabulated in two groups.

Results

Baseline Imaging, Treatment, and Follow-up

All TACE procedures were technically successful, and no major toxicity was reported. Table 2 provides information about types of intraarterial and surgical therapy and the frequency of the TACE procedures. The mean interval between the last TACE session and the last preoperative MR examination was 67 days (range, 20–181 days). The mean interval between the last TACE session and surgical therapy was 98 days (range, 23–204 days). The mean interval between the latest MR images and surgical therapy was 34 days (range, 2–83 days). The effects of time between treatment and surgery on tumor necrosis have been previously studied and were not followed up further (13). The median time to follow-up was 44 months (range, 12–79 months).

Pathologic Findings

Table 3 reports the extent of pathologic tumor necrosis assessed by using the quantitative histosegmentation

Table 2

Intraarterial and Surgical Therapy

Parameter	Value
Intraarterial therapy	
Treatment type	
Conventional TACE	9 (53)
TACE with drug-eluting beads	7 (41)
Both	1 (6)
Treatment frequency	
Overall no. of sessions	22
Mean no. of sessions per patient	1.29
Total no. of sessions	
Conventional TACE	11
TACE with drug-eluting beads	11
No. of patients treated with multiple sessions	
Conventional TACE	2 (12)
TACE with drug-eluting beads	3 (18)
Surgical therapy	
Orthotopic liver transplantation	13 (76)
Tumor resection	4 (24)

Note.—Except where indicated, data are numbers of patients. Numbers in parentheses are percentages.

Table 3

Histopathologic Necrosis according to Tumor Size at Baseline MR Imaging

Pathologic Class	<3 cm (n = 5)	3–5 cm (n = 11)	>5 cm (n = 1)
75%–100%	3	6	0
50%–74%	1	3	0
25%–49%	0	1	1
<25%	1	1	0

technique and stratified according to tumor size (largest diameter) on baseline MR images. Most analyzed lesions (65%) were 3–5 cm in size. Tumor necrosis was observed in all evaluated lesions and appeared to be independent of tumor size or TACE technique. Complete pathologic necrosis without identifiable viable cell clusters was found in seven lesions (41%). Both pathologic assessment techniques were compared. The mean percentage tumor necrosis with the histosegmentation technique was 70% (range, 10%–100%). When assessed visually, mean tumor necrosis was 71% (range, 10%–100%). Both methods showed a strong correlation with each other ($R^2 = 0.9786$) (Fig 3, A). No drift (as a marker for over- or underestimation of tumor necrosis) was observed

for the histosegmentation technique when compared with the reference of visual assessment. The variance for this technique was low and constant, showing a maximum deviation of 12%, with two cases deviating by more than 10% (RSE = 4.7) (Fig 3, B). The mean variation between the methods was 3.4% (range, 0%–12%).

Correlation of Radiologic and Pathologic Findings

Three-dimensional quantitative imaging.—Measurements of tumor necrosis calculated with 3D quantitative techniques were correlated with the results from histopathologic examination. The quantitative histosegmentation technique was used as a reference. The qEASL method showed a strong correlation with

histopathologic examination ($R^2 = 0.9657$) (Fig 4, A). Here, the residual plot showed no drift and a constantly low variance. The qEASL results deviated no more than 12% from actual histopathologic tumor necrosis, and only three cases (18%) exhibited a deviation of more than 10% (RSE = 6.38) (Fig 4, B). In the determination of necrosis according to four tabular classes, qEASL showed 100% overall accuracy. Figure 2, C, shows a representative, largely viable HCC lesion with central tumor necrosis (10% necrotic according to histopathologic examination), which had been detected with qEASL and measured as 14% necrotic (Fig 2, A). The qADC method showed an equally strong correlation with histopathologic findings ($R^2 = 0.9662$) (Fig 2, C). According to the residual plot, the maximum deviation of the qADC results was 13% and only two lesions (12%) showed a deviation of more than 10% (RSE = 6.33) (Fig 4, D). Similar to qEASL, overall accuracy was 100% in all tabular classes of tumor necrosis. Figure 2, B, demonstrates the qADC color map of the same representative lesion. According to qADC, this tumor showed 11% tumor necrosis. Because both 3D quantitative methods are based on different image acquisition techniques and calculated by using different ROIs, a lesion-by-lesion comparison of results was deemed necessary. According to a linear regression model, the correlation between qEASL and qADC was strong ($R^2 = 0.9685$) (Fig 4, E). Only three values (18%) deviated more than 10%, with a maximum of 13% (RSE = 5.99) (Fig 4, F). In two of these cases, the qEASL value was closer to the pathology reference and qADC provided a higher accuracy in one remaining case.

Subjective assessment.—Subjective measurements of 17 lesions by two readers resulted in 34 recorded values. Both readers were able to predict histopathologic findings, showing similar correlation patterns (reader 1: $R^2 = 0.8751$; reader 2: $R^2 = 0.8926$) (Fig 5, A and C). As for the analysis of residuals, no drift was observed.

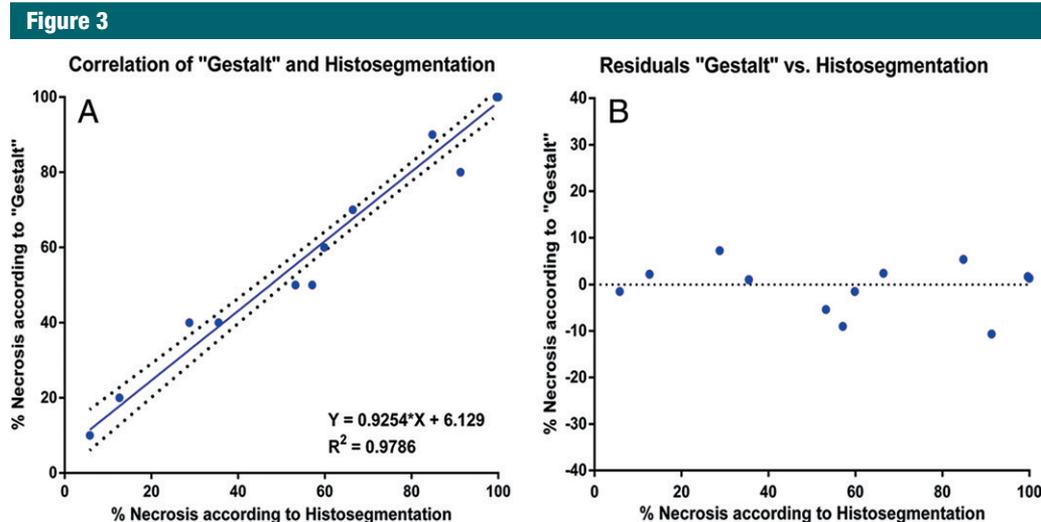


Figure 3: Comparison of pathologic assessment techniques. *A*, Correlation and, *B*, residual plots of Gestalt-based assessment with histosegmentation according to a linear regression model.

However, variance was high and a total of 15 deviations greater than 10% for a total of 34 readings (44%) were noted (reader 1: RSE = 12.18; reader 2: RSE = 11.28). Furthermore, a total of five deviations greater than 15% with a maximum of 37% were recorded (Fig 5, *B* and *D*). Table 4 summarizes sensitivity, specificity, positive predictive value, and negative predictive value in identifying the extent of tumor necrosis for both readers. When plotted against each other, both readings showed a correlation of the results ($R^2 = 0.8321$) (Fig 5, *E*). In addition, residual analysis showed a maximum deviation of more than 30% between the readers (deviating $>15\%$ in four cases; RSE = 12.14) (Fig 5, *F*).

EASL.—The degree of pathologic necrosis in treated lesions was stratified according to tumor response by using EASL criteria. Complete response, partial response, and stable disease were found to have complete pathologic necrosis in 100%, 20%, and 0% of the lesions, respectively. No lesion was classified as progressive disease. The sensitivity and specificity of EASL in the classification of patients as responders versus nonresponders were 100% (11 of 11 patients; 95% confidence interval [CI]: 73.33%, 100%) and 67% (four of six patients; 95%

CI: 22.68%, 94.67%), respectively. The positive predictive value was 85% (11 of 13 patients; 95% CI: 54.54%, 97.63%), the negative predictive value was 100% (four of four patients; 95% CI: 40.23%, 100%), and the overall accuracy was 88% (15 of 17 patients; 95% CI: 58.66%, 95.70%). Table 5 provides more details about the radiologic-pathologic agreement for the EASL method. However, because of the tabulated nature of the EASL guidelines, which are based on four classes rather than continuous numeric values for each patient, no linear regression model was considered for the comparison.

Complete response.—Seven lesions (41%) showed complete pathologic necrosis according to both histologic assessment techniques. Both 3D quantitative radiologic methods failed to identify any of these lesions as 100% necrotic. However, qEASL identified 100% (seven of seven lesions) as at least 95% necrotic and 57% (four of seven lesions) as at least 99% necrotic. With the qADC method, 86% of these lesions (six of seven lesions) were classified as at least 95% necrotic and 14% (one of seven lesions) were classified as at least 99% necrotic. Subjective readings identified complete pathologic necrosis in 57% (four of seven lesions) and 42%

(three of seven lesions) for each reader, respectively. When assessed according to EASL criteria (comparing baseline and follow-up MR imaging), 85% (six of seven) of the completely necrotic lesions were identified correctly as showing complete response.

Discussion

The main finding of this study is that both 3D quantitative MR image analysis techniques demonstrated a strong correlation with pathologically measured tumor necrosis. In addition, 3D quantitative assessment showed a higher sensitivity, higher specificity, lower discrepancy, and a lower variance when estimating the amount of pathologic tumor necrosis when compared with subjective assessment or measurements according to EASL guidelines.

Patient survival continues to be the focus of most clinical trials in cancer research, and tumor response at cross-sectional imaging is widely considered to be a weak outcome predictor (7). However, most clinical trials rely on imaging-based surrogate markers (eg, progression-free survival, time to progression, and disease-free survival), in part because stronger surrogate markers of survival are lacking (7). This radiologic-pathologic correlation

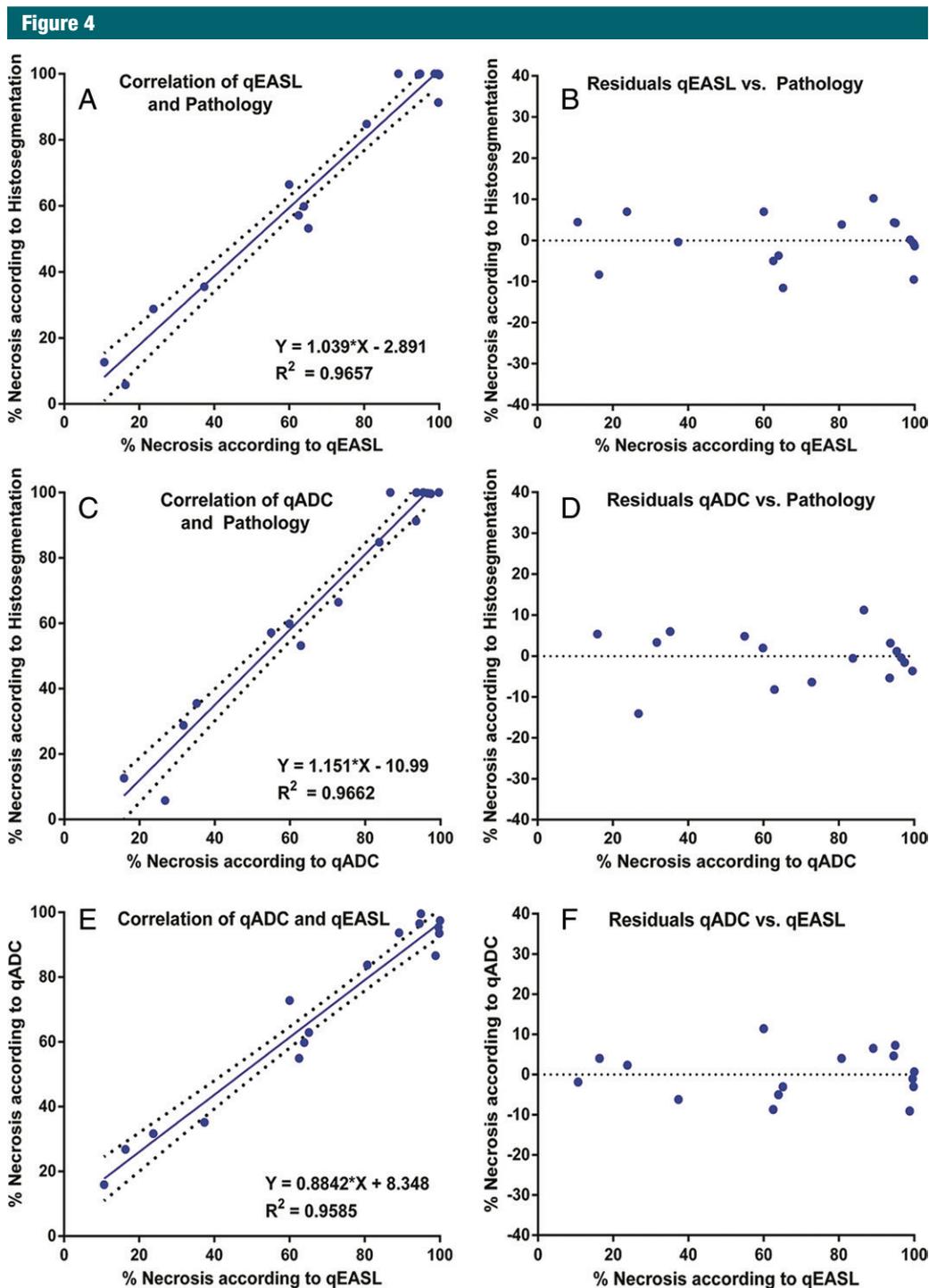


Figure 4: Correlation of 3D quantitative techniques with pathologic examination. *A*, Correlation and, *B*, residual plots demonstrate results for qEASL. *C*, Correlation and, *D*, residual plots demonstrate results for qADC. *E*, Correlation and, *F*, residual plots show intermethod agreement between qEASL and qADC.

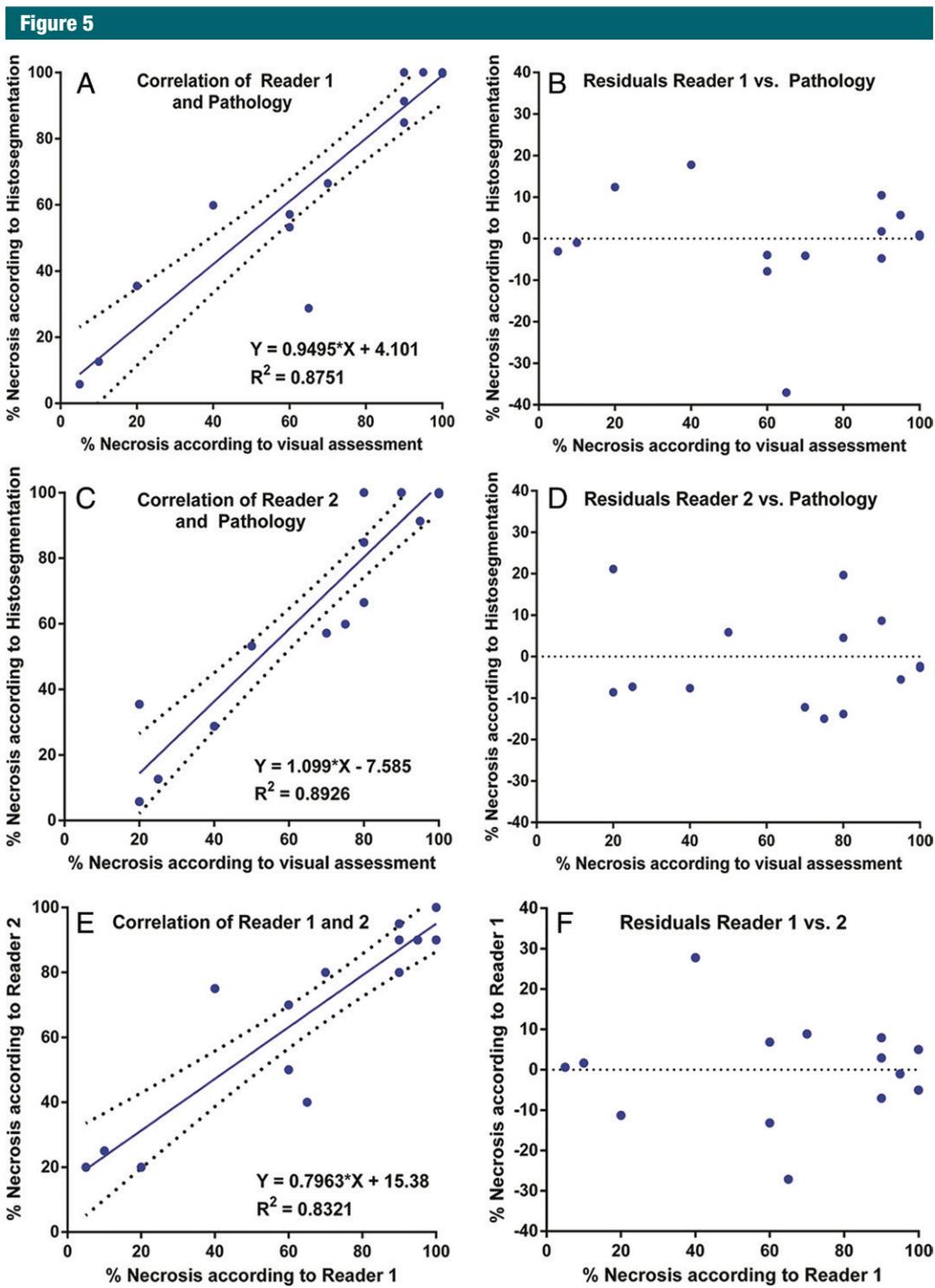


Figure 5: Correlation of subjective radiologic readings with pathologic examination. *A*, Correlation and, *B*, residual plots show results for reader 1. *C*, Correlation and, *D*, residual plots show results for reader 2. *E*, Correlation and, *F*, residual plots demonstrate agreement between readers.

Table 4

Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value of Subjective Readings

Reader and Pathologic Class	Necrosis according to MR Reading			
	75%–100%	50%–74%	25%–49%	<25%
Reader 1				
75%–100%	9	3	0	0
50%–74%	0	0	1	0
25%–49%	0	1	0	1
<25%	0	0	0	2
Reader 2				
75%–100%	9	0	0	0
50%–74%	2	2	0	0
25%–49%	0	0	1	1
<25%	0	0	1	1

Note.—For reader 1, sensitivity = 92% (12 of 13 patients; 95% CI: 63.90%, 98.72%), specificity = 75% (three of four patients; 95% CI: 20.34%, 95.88%), positive predictive value = 92% (12 of 13 patients; 95% CI: 63.90%, 98.72%), negative predictive value = 75% (three of four patients; 95% CI: 20.34%, 95.88%), and overall accuracy = 88% (15 of 17 patients; 95% CI: 50.10%, 93.04%). For reader 2, sensitivity = 100% (13 of 13 patients; 95% CI: 75.12%, 100.00%), specificity = 100% (four of four patients; 95% CI: 40.23%, 100.00%), positive predictive value = 100% (13 of 13 patients; 95% CI: 75.12%, 100.00%), negative predictive value = 100% (four of four patients; 96% CI: 40.23%, 100.00%), and overall accuracy = 100% (17 of 17 patients; 95% CI: 75.12%, 100.00%).

Table 5

EASL Response according to Pathologic Class

Pathologic Class with EASL Criteria	Complete Response (n = 6)	Partial Response (n = 5)	Stable Disease (n = 6)	Progressive Disease (n = 0)
75%–100%	6	3	0	0
50%–74%	0	2	2	0
25%–49%	0	0	2	0
<25%	0	0	2	0

Note.—Sensitivity = 100% (11 of 11 patients; 95% CI: 71.33%, 100.00%), specificity = 67% (four of six patients; 95% CI: 22.68%, 94.67%), positive predictive value = 85% (11 of 13 patients; 95% CI: 54.54%, 97.63%), negative predictive value = 100% (four of four patients; 95% CI: 40.23%, 100.00%), overall accuracy = 82% (15 of 17 patients; 95% CI: 54.54%, 97.63%).

study provides evidence for the diagnostic value of 3D quantitative imaging with regard to tumor necrosis after TACE. Our methods are based on a semiautomatic, work flow-efficient, and reproducible tumor segmentation software (18,25), which addresses important concerns with regard to clinical practicability of 3D quantitative imaging (9,33). The presented data provide multilayer proof of the precision of 3D quantitative contrast-enhanced and diffusion-weighted MR imaging when correlated with pathologic examination as

a standard of reference. Both methods showed a high level of correlation with histopathologic findings and a strong intermethod agreement. Comparison of 3D quantitative and subjective techniques demonstrated the advantage of a semiautomated approach, showing a stronger correlation and a significantly lower deviation of the 3D measurements from actual pathologic tumor necrosis. The narrow margin of error shown for qEASL and qADC might very well help achieve consensus in defining a standardized approach to reporting tumor

response in three dimensions. An interesting observation was made for the assessment of complete response, showing clear advantages for EASL and for the subjective visual assessment. This can be easily explained by the nature of the 3D quantitative measurements, which use a voxel-by-voxel approach. This circumstance can be viewed as a double-edged sword regarding the accuracy of this method because a single voxel with a brightness value above the threshold of the ROI may influence the overall result. Mathematically, this may prevent the software from correctly identifying complete pathologic necrosis, thus misclassifying the entire tumor. Hence, it is safe to assume that lesions showing complete response according to a subjective assessment or EASL criteria will not further benefit from a 3D quantitative approach. This finding is in agreement with findings from other reports, underlining the challenge to reliably predict complete pathologic necrosis (8). However, qEASL and qADC demonstrated high overall accuracy in classifying complete pathologic necrosis as greater than 95% necrotic according to MR imaging (100% [seven of seven lesions] and 86% [six of seven lesions] overall accuracy, respectively). By implication, our results suggest that qEASL and qADC are more specific and highly sensitive in predicting tumor necrosis in all classes except for complete response, achieving a close estimate for the latter. The agreement between both techniques as well as the equally good performance of qADC was not expected. This result was rather in contradiction to previously published data that demonstrated the lower diagnostic performance of ADC when compared with pure diffusion coefficients and enhancement-based techniques (23). However, the analyzed HCC lesions in our sample largely lacked fibrotic tissue, a factor known for interfering with diffusion-weighted MR sequences and delaying arterial enhancement predominantly in non-HCC lesions (34).

This study has several limitations. First, the retrospective nature of this analysis prevented us from performing gross pathologic examination of tumor

explants. We attempted to counter this limitation with a rigorous selection of patients, excluding all lesions larger than 7.5 cm. Because the pathologic assessment in this study relied solely on histologic examination, it was further necessary to exclude all incompletely preserved lesions. Those circumstances contributed to the second limitation: a relatively small sample size. Our conservative exclusion criteria further reduced the sample size. Because TACE has a unique mechanism of inducing tumor necrosis, patients treated with other intraarterial therapies like yttrium 90 radioembolization and ablative techniques were excluded to preserve a homogeneous sample. When designing this study, the use of conservative exclusion criteria that included patient history, imaging, and pathology as well as the focus on a single therapeutic modality (TACE) were considered pivotal to reliably assess the diagnostic accuracy. Another limitation is the selection of patients according to MR image quality; however, only a small minority (<1%) of initially reviewed MR images demonstrated severe artifacts in the region of targeted lesions. In addition, we recognize that the results of ADC measurements can be highly heterogeneous even though no guideline has been made available until today. Although the technique presented herein is experimental and requires additional validation, the results of this study rely on pathologic examination as a standard of reference and encourage us to further investigate qADC. Despite the lack of consensus and the highly debated technology, we appreciate the value of ADC imaging for the assessment of tumor response after TACE and consider qADC as complementary to the more commonly used enhancement-based technique (35). As opposed to diffusion-weighted MR imaging, the enhancement-based qEASL technique helps quantify tumor vascularity by using arterial phase images and provides intuitive feedback for interventional radiologists who apply intraarterial therapies. In light of the advent of 3D quantitative tumor response assessment at enhancement-based intraprocedural cone-beam computed

tomography (36), the authors suggest the use of qEASL as a correspondingly enhancement-based and translatable instrument to measure tumor response to TACE at follow-up MR imaging.

In summary, the high precision of qEASL and qADC might have implications regarding clinical decisions as to whether a patient should undergo repeat treatment after an initial TACE session. It can be speculated that, according to the presented results, 3D quantitative techniques have the potential to enhance the diagnostic performance of follow-up imaging after TACE and to facilitate therapeutic decisions on the basis of an improved stratification between responders and nonresponders. However, the ramifications of this hypothesis must be clarified with further studies that will investigate patient survival on the basis of qEASL and qADC.

Acknowledgments: The authors acknowledge the continuous support and guidance provided by Rolf W. Günther, MD, Berlin, Germany. We also thank Nikhil Bhagat, MD, for the critical review of the presented data.

Disclosures of Conflicts of Interest: J.C. disclosed no relevant relationships. L.D.W. disclosed no relevant relationships. M.L. Activities related to the present article: is an employee of Philips. Activities not related to the present article: disclosed no relevant relationships. Other relationships: disclosed no relevant relationships. R.D. disclosed no relevant relationships. T.C. disclosed no relevant relationships. D.L. Activities related to the present article: is an employee of Philips. Activities not related to the present article: disclosed no relevant relationships. Other relationships: disclosed no relevant relationships. V.C. disclosed no relevant relationships. R.S. disclosed no relevant relationships. Z.W. disclosed no relevant relationships. V.T. disclosed no relevant relationships. L.J.S. disclosed no relevant relationships. I.R.K. disclosed no relevant relationships. J.F.G. Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: received a grant from Biocompatibles/BTG, Bayer Healthcare, Philips Medical, Nordio/BTG, Threshold, Guerbet, and DOD; is a consultant for Biocompatibles/BTG, Bayer Healthcare, Guerbet, Nordio/BTG, Philips Healthcare, and Jennerex. Other relationships: disclosed no relevant relationships.

References

1. 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):421-430.

2. Lencioni R, Crocetti L. Local-regional treatment of hepatocellular carcinoma. *Radiology* 2012;262(1):43-58.
3. Forner A, Llovet JM, Bruix J. Hepatocellular carcinoma. *Lancet* 2012;379(9822):1245-1255.
4. European Association for the Study of the Liver; European Organisation for Research and Treatment of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* 2012; 56(4):908-943.
5. Chapiro J, Tacher V, Geschwind JF. Intra-arterial therapies for primary liver cancer: state of the art. *Expert Rev Anticancer Ther* 2013;13(10):1157-1167.
6. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. *Semin Liver Dis* 2010;30(1):52-60.
7. Llovet JM, Di Bisceglie AM, Bruix J, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. *J Natl Cancer Inst* 2008;100(10):698-711.
8. Vouche M, Kulik L, Atassi R, et al. Radiological-pathological analysis of WHO, RECIST, EASL, mRECIST and DWI: imaging analysis from a prospective randomized trial of Y90 ± sorafenib. *Hepatology* 2013;58(5):1655-1666.
9. Gonzalez-Guindalini FD, Botelho MP, Harmath CB, et al. Assessment of liver tumor response to therapy: role of quantitative imaging. *RadioGraphics* 2013;33(6):1781-1800.
10. Therasse P, Arbuuck 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):205-216.
11. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981;47(1):207-214.
12. Welsh JL, Bodeker K, Fallon E, Bhatia SK, Buatti JM, Cullen JJ. Comparison of response evaluation criteria in solid tumors with volumetric measurements for estimation of tumor burden in pancreatic adenocarcinoma and hepatocellular carcinoma. *Am J Surg* 2012;204(5):580-585.
13. Riaz A, Lewandowski RJ, Kulik L, et al. Radiologic-pathologic correlation of hepatocellular carcinoma treated with chemoembolization. *Cardiovasc Intervent Radiol* 2010;33(6):1143-1152.

14. 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.
15. Riaz A, Kulik L, Lewandowski RJ, et al. Radiologic-pathologic correlation of hepatocellular carcinoma treated with internal radiation using yttrium-90 microspheres. *Hepatology* 2009;49(4):1185–1193.
16. Lencioni R. New data supporting modified RECIST (mRECIST) for hepatocellular carcinoma. *Clin Cancer Res* 2013;19(6):1312–1314.
17. Mantatzis M, Kakolyris S, Amarantidis K, Karayiannakis A, Prassopoulos P. Treatment response classification of liver metastatic disease evaluated on imaging: are RECIST unidimensional measurements accurate? *Eur Radiol* 2009;19(7):1809–1816.
18. 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* 2012;23(12):1629–1637.
19. Bonekamp S, Halappa VG, Geschwind JF, et al. Unresectable hepatocellular carcinoma: MR imaging after intraarterial therapy. II. Response stratification using volumetric functional criteria after intraarterial therapy. *Radiology* 2013;268(2):431–439.
20. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Radiology* 2003;226(1):24–28.
21. 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.
22. Bonekamp S, Li Z, Geschwind JF, et al. Unresectable hepatocellular carcinoma: MR imaging after intraarterial therapy. I. Identification and validation of volumetric functional response criteria. *Radiology* 2013;268(2):420–430.
23. Wagner M, Doblas S, Daire JL, et al. Diffusion-weighted MR imaging for the regional characterization of liver tumors. *Radiology* 2012;264(2):464–472.
24. Kim SY, Lee SS, Byun JH, et al. Malignant hepatic tumors: short-term reproducibility of apparent diffusion coefficients with breath-hold and respiratory-triggered diffusion-weighted MR imaging. *Radiology* 2010;255(3):815–823.
25. 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):446–452.
26. Kim S, Mannelli L, Hajdu CH, et al. Hepatocellular carcinoma: assessment of response to transarterial chemoembolization with image subtraction. *J Magn Reson Imaging* 2010;31(2):348–355.
27. Loffroy R, Lin M, Yenokyan G, et al. Intra-procedural C-arm dual-phase cone-beam CT: can it be used to predict short-term response to TACE with drug-eluting beads in patients with hepatocellular carcinoma? *Radiology* 2013;266(2):636–648.
28. Shim JH, Han S, Shin YM, et al. Optimal measurement modality and method for evaluation of responses to transarterial chemoembolization of hepatocellular carcinoma based on enhancement criteria. *J Vasc Interv Radiol* 2013;24(3):316–325.
29. Kwan SW, Fidelman N, Ma E, Kerlan RK Jr, Yao FY. Imaging predictors of the response to transarterial chemoembolization in patients with hepatocellular carcinoma: a radiological-pathological correlation. *Liver Transpl* 2012;18(6):727–736.
30. Koontz NA, Gunderman RB. Gestalt theory: implications for radiology education. *AJR Am J Roentgenol* 2008;190(5):1156–1160.
31. World Health Organization. Handbook for Reporting Results of Cancer Treatment. WHO offset publication. Geneva, Switzerland: World Health Organization, 1979; 48.
32. Alberg AJ, Park JW, Hager BW, Brock MV, Diener-West M. The use of “overall accuracy” to evaluate the validity of screening or diagnostic tests. *J Gen Intern Med* 2004;19(5 Pt 1):460–465.
33. 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(2):416–419.
34. Rimola J, Forner A, Reig M, et al. Cholangiocarcinoma in cirrhosis: absence of contrast washout in delayed phases by magnetic resonance imaging avoids misdiagnosis of hepatocellular carcinoma. *Hepatology* 2009;50(3):791–798.
35. Taouli B, Vilgrain V, Dumont E, Daire JL, Fan B, Menu Y. Evaluation of liver diffusion isotropy and characterization of focal hepatic lesions with two single-shot echo-planar MR imaging sequences: prospective study in 66 patients. *Radiology* 2003;226(1):71–78.
36. Wang Z, Lin M, Lesage D, et al. Three-dimensional evaluation of lipiodol retention in HCC after chemoembolization: a quantitative comparison between CBCT and MDCT. *Acad Radiol* 2014;21(3):393–399.