

Published in final edited form as:

*Acad Radiol.* 2013 April ; 20(4): 446–452. doi:10.1016/j.acra.2012.11.009.

## Semi-automatic volumetric tumor segmentation for hepatocellular carcinoma: Comparison between C-arm Cone Beam Computed Tomography and MRI

Vania Tacher, MD<sup>1</sup>, MingDe Lin, PhD<sup>2</sup>, Michael Chao, BS<sup>1</sup>, Lars Gjestebj, PhD<sup>1</sup>, Nikhil Bhagat, MD<sup>1</sup>, Abdelkader Mahammedi, MD<sup>1</sup>, Roberto Ardon, PhD<sup>3</sup>, Benoit Mory, PhD<sup>3</sup>, and Jean-François Geschwind, MD<sup>1</sup>

(<sup>1</sup>)Russell H. Morgan Department of Radiology and Radiological Science, Division of Vascular and Interventional Radiology, Johns Hopkins Hospital, Baltimore, MD, USA

(<sup>2</sup>)Clinical Informatics, Interventional, and Translational Solutions (CIITS), Philips Research North America, Briarcliff Manor, NY, USA

(<sup>3</sup>)Philips Healthcare, Suresnes, France

### Abstract

**Rationale and Objectives**—To evaluate the precision and reproducibility of a semi-automatic tumor segmentation software in measuring tumor volume of hepatocellular-carcinoma-(HCC) before the first trans-arterial chemo-embolization-(TACE) on contrast-enhancement magnetic-resonance-imaging-(CE-MRI) and intra-procedural dual-phase C-arm cone-beam computed-tomography-(DP-CBCT) images.

**Materials and Methods**—19HCCs were targeted in 19patients(one per patient) who underwent baseline diagnostic CE-MRI and an intra-procedural DP-CBCT. The images were obtained from CE-MRI-(arterial-phase of an intra-venous contrast medium injection) and DP-CBCT-(delayed-phase of an intra-arterial contrast medium injection) before the actual embolization. Three readers measured tumor volumes using a semi-automatic 3D-volumetric segmentation software which used a region-growing method employing non-Euclidean radial basis functions. Segmentation time and spatial position were recorded. The tumor volume measurements between images sets were compared using linear-regression and Student t-test, and evaluated with Intraclass-Correlation analysis-(ICC). The inter-rater Dice Similarity Coefficient-(DSC) assessed the segmentation spatial localization.

**Results**—All 19 HCCs were analyzed. On CE-MRI and DP-CBCT examinations respectively, A) the mean segmented tumor volumes was  $87\pm 8\text{cm}^3$ [2–873] and  $92\pm 10\text{cm}^3$ [1–954], with no statistical difference of segmented volumes by readers of each tumor between the two imaging modalities and the mean time required for segmentation was  $66\pm 45\text{seconds}$  [21–173] and  $85\pm 34\text{seconds}$ [17–214] ( $p=0.19$ ), B),the ICCs were 0.99 and 0.974, showing a strong correlation

© 2012 The Association of University Radiologists. Published by Elsevier Inc. All rights reserved.

**Corresponding author:** Jean-François Geschwind MD. Professor of Radiology, Surgery and Oncology Director, Vascular and Interventional Radiology Director, Interventional Radiology Center Johns Hopkins University School of Medicine Interventional Radiology Center Sheikh Zayed Tower, Suite 7203 The Johns Hopkins Hospital 1800 Orleans Street Baltimore, MD 21287 Phone: (410) 614-6597 Fax: (410) 955-0233 jfg@jhmi.edu.

**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.

among readers, and C) the inter-rater DSCs showed a good to excellent inter-user agreement on the spatial localization of the tumor segmentation—(0.70±0.07 and 0.74±0.05, p=0.07).

**Conclusion**—This study shows a strong correlation, precision and reproducibility of semi-automatic tumor segmentation software in measuring tumor volume on CE-MRI and DP-CBCT images. The use of the segmentation software on DP-CBCT and CE-MRI can be a valuable and highly accurate tool to measure the volume of hepatic tumors.

### Keywords

Tumor segmentation software; Dice Similarity Coefficient; C-arm cone-beam CT; MRI; Hepatocellular carcinoma; TACE

## Introduction

Evaluation of treatment response is a key aspect in cancer therapy. On imaging, tumor change after locoregional therapy (i.e. trans-arterial chemo-embolization (TACE)), is used to determine therapeutic success (1, 2). Indeed, the three accepted methods to assess response to TACE are measuring changes in tumor size (Response Evaluation Solid Tumor [RECIST]), enhancement size (European Association for the Study of the Liver [EASL]), and tumor enhancement size (modified Response Evaluation Criteria in Solid Tumors [mRECIST]) on MRI imaging (3-5). Indeed, RECIST, described in 2000, is based on the measurement of the longest diameter of a given target lesion, or the sum of the longest diameters for a set of target lesions. This uni-dimensional assessment is made on a single axial slice on cross sectional imaging (either computed tomography (CT) or magnetic resonance imaging (MRI)). But these criteria are limited for the following reasons: (1) it is a one-dimensional measurement of a tumor volume, (2) it disregards the extent of tumor viability/necrosis after therapy, and (3) it is susceptible to high inter-observer variability (6-8). While RECIST was appropriate at the time of its introduction, the simplicity of RECIST now underutilizes image processing functionalities available in modern imaging units. With the advent of Multi-Detector Computed Tomography (MDCT) and MRI, the ability to assess tumor volume using three-dimensional metrics appear more accurate and precise, as recognized in RECIST version 1.1. In addition, with the increased use of C-arm Cone Beam Computed Tomography (CBCT) for interventional oncology, tumor size and enhancement changes can be assessed during the intervention for planning or for treatment success evaluation (9). A new criteria, a volumetric RECIST, was described as a feasible method that can be performed in a realistic time frame to evaluate the tumor volume change after therapy (10). However, before volume-based metrics can supplant RECIST, these methods must be shown to be accurate and reproducible. A semi-automatic software for tumor volume segmentation has already been evaluated in a pre-clinical study using the VX2 rabbit hepatic tumor model. In this previous work, the segmentation results from CBCT and diagnostic CT were confirmed by histology (11). Because the segmentation software can be used in clinic, and especially in interventional oncology, we decided to confirm the similarity between the liver tumor volume assessments on intra-procedural imaging (DP-CBCT) to the diagnostic imaging gold standard (CE-MRI). The purpose of this study was to evaluate the correlation, the precision and the reproducibility of a semi-automatic tumor segmentation software used to measure the tumor volume of hepatocellular carcinoma (HCC) before trans-arterial chemo-embolization (TACE) by comparing contrast enhanced MRI (CE-MRI), the gold standard in liver imaging, to intra-procedural Dual-Phase CBCT (DP-CBCT) imaging.

## Material and Methods

### Patient study selection

This was a single institution prospective study (HIPPA compliant and IRB approved) but the data analysis was done retrospectively. All patients were provided with informed consent before inclusion in the study. The study group included all patients 1) with hepatocellular carcinoma (HCC), 2) scheduled to undergo their first TACE (conventional lipiodol based or drug eluting beads), 3) without any prior systemic or local therapy, 4) with dynamic contrast-enhanced MR imaging (CE-MRI) within four weeks before TACE, and 5) with pre-embolization intra-procedural DP-CBCT. From May 2, 2011 to April 11, 2012, the liver tumor board discussed the care of 303 patients who underwent TACE. Eligibility criteria for performing TACE were identical to those already published (9). Of the 303 patients, 249 patients were treated in an angiography suite equipped with DP-CBCT (12). Of those, 147 patients had HCC, with 60 undergoing their first treatment. 19 of these patients had intra-procedural DP-CBCT before chemoembolization and were included in this study. DP-CBCT were performed in all patients unless the CBCT option was not available (in the angiography room) or there was patient related factors that precluded the use of CBCT (i.e. inability to hold breath).

### Pre-Procedure MR Imaging Technique

All patients underwent baseline CE-MRI imaging using a 1.5-T MR unit (CV/I, GE Medical Systems, Milwaukee, WI, USA) and a phased-array torso coil within four weeks before the TACE. The imaging protocol included: 1) axial T2-weighted fast spin-echo images (TR/TE, 5000/100 msec; matrix size, 256 × 256; slice thickness, 8-mm; interslice gap, 2-mm; receiver bandwidth, 32-kHz), 2) axial single-shot breath-hold gradient-echo diffusion-weighted echo-planar images (TR/TE, 5000-6500/110 msec; matrix size, 128 × 128; slice thickness, 8-mm; interslice gap, 2-mm; *b* value, 500 sec/mm<sup>2</sup>; receiver bandwidth, 64-kHz), and 3) axial breath-hold unenhanced and contrast-enhanced (0.1 mmol/kg IV of gadodiamide, Omniscan, General Electric, Princeton, NJ) T1-weighted 3D fat-suppressed spoiled gradient-echo images (TR/TE, 5.1/1.2 msec; field of view, 320-400 mm<sup>2</sup>; matrix size, 192 × 160; slice thickness, 4-6-mm; receiver bandwidth, 64-kHz; flip angle, 15°) in the arterial and portal venous phases (20 and 70 seconds after intravenous contrast administration, respectively). The arterial phase of the CE-MRI imaging was used for the study (13).

### Intra-Procedural Dual-Phase Cone-Beam Computed Tomography Technique

The imaging was performed using a commercially available angiographic system (Allura Xper FD20, Philips Healthcare, Best, The Netherlands). This system was equipped with the XperCT option, enabling C-arm cone-beam CT acquisition and volumetric image reconstruction (Feldkamp back projection) (14). For each CBCT scan, the area of interest was positioned in the system isocenter, and over approximately 5 seconds, 312 projection images (60 frames per second) were acquired with the motorized C-arm covering a 200° clockwise arc at 40° per sec rotation speed (matrix size, 256 × 256 × 198; field of view, 25 × 25 × 19 cm). As the images were being acquired, the projections were transferred to the reconstruction computer to produce volumetric data. One-millimeter isotropic images were reconstructed from DP-CBCT scans. The dual-phase prototype feature allowed for the acquisition of two sequential multi-phasic CBCT scans using only one contrast injection (9, 12, 15). The same contrast injection protocol was applied to all cases: the contrast injection was done through the catheter placed into the proper hepatic artery, and the scan was triggered after contrast was injected for three seconds (amount, 18ml; rate 2ml/s; Oxilan 300mgI/ml, Guerbet LLC, Bloomington, IN). The patients were instructed to be at end-expiration apnea during each of the CBCT scans with free breathing between the early and delayed arterial

phase scans. If needed, oxygen was administered to patients during the acquisition to minimize the discomfort of breath holding. In this work, the second phase was used because the tumors were best visualized (16).

### Imaging Data Evaluation Overview

The CE-MRI and DP-CBCT images were retrospectively evaluated offline and used for tumor segmentation in two modality separate blind and random sessions. The evaluation was done by three readers (one with 8-years of interventional radiology experience (\_\_\_), one radiology resident (\_\_\_), and one biomedical engineering student (\_\_\_)). For each patient, the same target tumor was segmented by consensus from the CE-MRI and the DP-CBCT. The target tumor was defined as the tumor with the largest diameter that was treated. The time required to segment the tumor was recorded. The segmentations were compared within each imaging modality and between readers using a spatial overlap index, the inter-rater Dice similarity coefficient (DSC) (17-19).

### Semi-automatic Tumor Segmentation

A semi-automatic 3D volume tumor segmentation software (Medisys, Philips Research, Suresnes, France) was used by the three readers on the DP-CBCT and CE-MRI to segment the tumors in the same way for each image set. The segmentation was done on the DP-CBCT and CE-MRI in two sessions separated by few days (3-21 days). This prototype software was shown in previous pre-clinical work to accurately segment tumors with minimal user interaction (11). In practice, the images sets were opened as a digital imaging and communication in medicine (DICOM) images format in the segmentation software. The user first identified the target tumor and placed an initial control point on the DP-CBCT images or on the CE-MRI images, and from there, the user could interactively expand or contract a 3D adaptive region of interest (Figure 1). Additional segmentations could be included by placing more control points. Corrections were made in the same volumetric way. This software was inspired by non-Euclidean geometry and the theory of radial basis functions (mathematically, a function whose value depends only on the distance from the origin). This subset was spanned by a novel class of non-Euclidean Radial Basis Functions, built from image-dependant metrics using local images features, such as intensity distribution or edge information (20). This method allowed for segmentations that follow 3D image features including straight edges and corners. This method was used because it could accurately segment in 3D with minimal user interaction (11). The tumor volume was directly calculated from the 3D tumor segmentation based on the voxel's size.

### Dice Similarity Coefficient (DSC)

The DSC is commonly used as a standardized metric to evaluate the performance of spatial overlap/localization of two sets of segmentations and evaluates inter-user agreement (19). DSC is defined by the following equation:

$$DSC = \frac{2 \times |Region A \cap Region B|}{|Region A| + |Region B|}$$

where region A and region B of the equation represent the tumor volume segmented by two readers. The DSC varies between zero and one, where 0 implies no overlap and 1 represents identical regions with perfect overlap. Landis and Koch described different ranges of values: values greater than 0.75 represent an excellent agreement, below 0.4 a poor agreement, and between 0.75 and 0.40 a good agreement (21). The inter-user agreement value was calculated as the average of the three DSC values of pairwise comparisons between the readers on each imaging modality. The DSC calculation and visualization was done using

MATLAB (MathWorks, Natick, MA) and ImageJ (National Institutes of Health, Bethesda, MD) (22).

### Statistical analysis

All statistical analyses were performed using Minitab Statistical Software (Minitab, Inc., State College, PA). To evaluate the inter-user agreement, the analysis was performed based on measurements of the segmented tumor volume and their spatial location. Intraclass correlation coefficient (ICC) was used to measure the strength of agreement between observers on each imaging modality (23). The readers' average tumor volumes measured on CE-MRI and DP-CBCT were compared using linear regression, including the R-squared value of correlation, to evaluate consistency values between the two imaging modalities (24). In addition, the least squares method was applied to this data to visualize a linear fit function and to calculate a 95% confidence interval for volume measurements between the two imaging modalities. Each tumor volume segmented by each reader on each imaging modality was compared using Student's t-test. All inter-rater DSC values within each imaging modality were compared using Student's t-test. The time (in seconds) taken for each reader to perform one tumor segmentation was recorded and compared between imaging modalities using a Student's t-test. Two-sided p-values  $\leq 0.05$  were considered statistically significant.

### Results

All patient characteristics are described in Table 1. The presence of a target tumor (HCC) before TACE was observed in all 19 patients on CE-MRI and DP-CBCT (one tumor per patient). All CE-MRI and DP-CBCT images were acquired before embolization for the included patient group and were used for segmentation. The semi-automatic tumor segmentation was performed retrospectively as shown in Figure 1. The mean tumor volumes for all patients measured by all readers were: on CE-MRI  $87 \pm 8 \text{ cm}^3$  [2 – 873] and on DP-CBCT was  $92 \pm 10 \text{ cm}^3$  [1 – 954],  $p=0.30$ . The mean time for all readers to segment each of the 19 CE-MRI examinations was  $66 \pm 45$  seconds [21 – 173] and the 19 DP-CBCT examinations was  $85 \pm 34$  seconds [17 – 214]. There was no statistically significant difference between DP-CBCT and CE-MRI ( $p=0.19$ ). The specific tumor volumes and time needed for segmentation for each of the 19 patients on CE-MRI and DP-CBCT are described in Table 2. No statistical difference was found between readers segmentations on the two imaging modalities for each tumor, according to the Student t-test except for one tumor which was probably due to the difference of enhancement between imaging sets (Table 2).

The ICC for inter-observer reliability of the mean volume measurements for each tumor on the two imaging sets were 0.997 and 0.990 on CE-MRI and DP-CBCT, respectively, showing a strong correlation among readers. Additionally, the high correlation between readers for the mean acquired volume within each of the imaging modality was evidenced by the linear regression analysis shown in Figure 2. An R-squared value of 0.99 indicated strong agreement between the mean tumor volumes acquired from CE-MRI and DP-CBCT, showing consistent measurements across both imaging modalities.

The inter-rater DSC allowed for evaluation of inter-user agreement. The spatial segmentation of a tumor is illustrated on a representative case in Figure 3. No significant difference ( $p=0.07$ ) was found between the inter-rater DSC according to the imaging modality per lesion, with CE-MRI DSC  $0.70 \pm 0.07$  [0.44 – 0.88] and DP-CBCT DSC  $0.74 \pm 0.05$  [0.49 – 0.92]. Table 3 reports specific inter-rater DSC for each patient. However, the inter-rater DSC of each reader when comparing between imaging modalities showed a significantly higher value of inter-rater DSC on DP-CBCT vs. CE-MRI ( $p=0.006$ , Table 4).

The different values of inter-rater DSC depending on the imaging modality showed similar variation, as described in Figure 4. For CE-MRI, 6 tumor segmentations had excellent agreement (inter-rater DSC  $\geq 0.75$ ), 13 had good agreement ( $0.4 \leq$  inter-rater DSC  $< 0.75$ ), and 0 had poor agreement (inter-rater DSC  $< 0.4$ ). For DP-CBCT, 11 segmentations had an excellent agreement, 8 had a good agreement, and 0 had poor agreement.

## Discussion

Our study confirms the precision and reproducibility of a semi-automatic tumor segmentation software across two imaging modalities (CE-MRI and DP-CBCT) obtained by different readers with different levels of experience. We chose three readers with very different levels of expertise to highlight the strength and the reproducibility of the tumor segmentation software. We found a strong correlation between the two imaging sets regarding hepatic tumor volume and spatial localization according to all readers. The ICC for inter-observer reliability showed a strong correlation and consistent measurements across both modalities between readers. Thus, the results reinforce that tumor volumes can be determined with high consistency on CE-MRI and DP-CBCT, both for each individual reader and across the group of readers. The trend of the increased mean tumor volume found on DP-CBCT compared to CE-MRI can be explained by the combination of two hypotheses: the four week time span between the pre-procedural CE-MRI imaging and the DP-CBCT with a potential progression of the tumor and the increased conspicuity with the intra-arterial injection.

This semi-automatic tumor segmentation software system was also efficient, as the mean time required to segment using either modality was less than two minutes. The time required by readers to perform segmentation on CE-MRI was not statistically different from that of DP-CBCT. The tumor segmentation spatial overlap agreement between readers, as measured by inter-rater DSC, was good to excellent on both CE-MRI and DP-CBCT, with a higher overlap agreement on DP-CBCT. This could be explained by the higher contrast between tumor and hepatic parenchyma seen on DP-CBCT images because of the arterial blood supply to HCC and the fact that contrast was directly injected intra-arterially.

Change in tumor size is an important indicator for determining the success of locoregional therapies. Currently, one or two-dimensional measurements are being used to estimate tumor size for HCC staging and assessment of treatment response. In the latest RECIST version, the authors recognized the need to study volumetric assessment in greater detail before uni-dimensional assessment could be improved upon because of insufficient standardization (4). A part of the standardization effort is the development of tumor segmentation software that can accurately segment tumors in realistic time frames. However, studies have found that 3D tumor volumetric may be a more effective standard than 1D and 2D measurements (25-27). New tumor assessments as volumetric RECIST and quantitative EASL have recently been introduced to evaluate tumor response (10). As a result, research efforts from medical imaging companies have increased to provide tumor volume segmentation algorithms that allow for accurate and reproducible results across different imaging modalities (11). In response to this need, our research team developed a semi-automated tumor segmentation software that is reproducible and precise. This study is the first clinical step towards a 3D intra-procedural volumetric criteria evaluation of HCC before chemoembolization.

There were some limitations to this study. First, complex appearances of tumor (heterogeneous, infiltrative, necrotic, core, and irregular shape), their adjacent structures, and imaging artifacts could negatively impact the tumor segmentation reading. However, the inter-rater DSC among readers was between good to excellent, indicating that these factors may not play a large role using this segmentation software. Secondly, the CBCT (25 cm) has

a limited field of view, where a peripheral tumor could be excluded. Thirdly, the total number of patients (and tumors) was relatively small, and our results should be further validated in a larger cohort of patients. Finally, this study was limited to HCC, and the performance of the segmentation software should also be investigated for different liver tumor histologies (i.e. neuroendocrine, colorectal, etc.) and for tumors occurring in different organs.

## Conclusion

Our study demonstrated the precision and reproducibility of a work-flow efficient, semi-automatic tumor volume segmentation software on clinical data acquired with two imaging modalities. The volume and the spatial localization of the segmented tumors by multiple readers showed strong consistency, correlation and a high level of agreement. Thus, the use of this segmentation software on DP-CBCT and CE-MRI can be a useful tool for volumetric hepatic tumor assessment.

## Acknowledgments

### Funding and Financial Support:

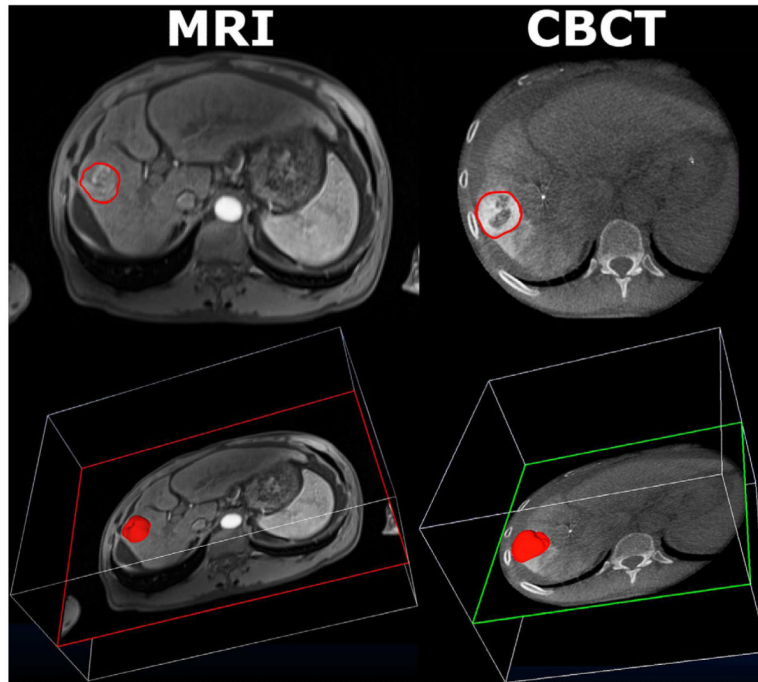
This study was funded by NIH/NCI R01 CA160771 and NCRN UL1 RR 025005.

## References

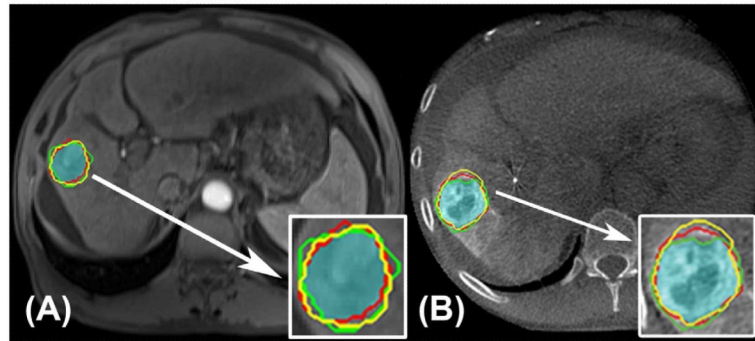
1. Bruix J, Sherman M. Management of hepatocellular carcinoma. *Hepatology* (Baltimore, Md. 2005; 42(5):1208–36.
2. Vossen JA, Buijs M, Kamel IR. Assessment of tumor response on MR imaging after locoregional therapy. *Techniques in vascular and interventional radiology*. 2006; 9(3):125–32. [PubMed: 17561215]
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. *Journal of the National Cancer Institute*. 2000; 92(3):205–16. [PubMed: 10655437]
4. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009; 45(2):228–47. [PubMed: 19097774]
5. Kamel IR, Liapi E, Reyes DK, Zahurak M, Bluemke DA, Geschwind JF. Unresectable hepatocellular carcinoma: serial early vascular and cellular changes after transarterial chemoembolization as detected with MR imaging. *Radiology*. 2009; 250(2):466–73. [PubMed: 19188315]
6. Forner A, Ayuso C, Varela M, et al. Evaluation of tumor response after locoregional therapies in hepatocellular carcinoma: are response evaluation criteria in solid tumors reliable? *Cancer*. 2009; 115(3):616–23. [PubMed: 19117042]
7. Lim HK, Han JK. Hepatocellular carcinoma: evaluation of therapeutic response to interventional procedures. *Abdominal imaging*. 2002; 27(2):168–79. [PubMed: 11847576]
8. Suzuki C, Torkzad MR, Jacobsson H, et al. Interobserver and intraobserver variability in the response evaluation of cancer therapy according to RECIST and WHO-criteria. *Acta oncologica* (Stockholm, Sweden). 2010; 49(4):509–14.
9. Loffroy R, Lin M, Yenokyan G, et al. Intra-Procedural C-Arm Dual-Phase Cone Beam CT Imaging to predict Response to Drug-Eluting Beads Transcatheter Arterial Chemoembolization in Patients with Hepatocellular Carcinoma. *Radiology*. 2012 In Press.
10. Lin M, Pellerin O, Bhagat N, et al. Quantitative and Volumetric EASL and RECIST: Feasibility of a Semi-automated Software Method to Assess Tumor Response after Transcatheter Arterial Chemoembolization (TACE). *Journal of Vascular and Interventional Radiology*. 2012 In Press.

11. Pellerin O, Lin M, Bhagat N, Ardon R, Mory B, Geschwind JF. Comparison of Semi-automatic Volumetric VX2 Hepatic Tumor Segmentation from Cone Beam CT and Multi-detector CT with Histology in Rabbit Models. *Acad Radiol*. 2012
12. Lin M, Loffroy R, Noordhoek N, et al. Evaluating tumors in transcatheter arterial chemoembolization (TACE) using dual-phase cone-beam CT. *Minimally invasive therapy & allied technologies : MITAT : official journal of the Society for Minimally Invasive Therapy*. 2011; 20(5):276–81.
13. Mori K, Yoshioka H, Takahashi N, et al. Triple arterial phase dynamic MRI with sensitivity encoding for hypervascular hepatocellular carcinoma: comparison of the diagnostic accuracy among the early, middle, late, and whole triple arterial phase imaging. *AJR American journal of roentgenology*. 2005; 184(1):63–9. [PubMed: 15615952]
14. Feldkamp L, Davis L, Kress J. Practical cone-beam algorithm. *JOpt Soc Am*. 1984; 1(6):A1, 612–9.
15. Miyayama S, Yamashiro M, Okuda M, et al. Detection of corona enhancement of hypervascular hepatocellular carcinoma by C-arm dual-phase cone-beam CT during hepatic arteriography. *Cardiovascular and interventional radiology*. 2011; 34(1):81–6. [PubMed: 20333382]
16. Loffroy R, Lin M, Rao P, et al. Comparing the detectability of hepatocellular carcinoma by C-arm dual-phase cone-beam computed tomography during hepatic arteriography with conventional contrast-enhanced magnetic resonance imaging. *Cardiovascular and interventional radiology*. 35(1):97–104. [PubMed: 21328023]
17. Dice LR. Measures of the Amount of Ecologic Association Between Species. *Ecological Society of America*. 1945; 26(3):297–302.
18. Qazi AA, Pekar V, Kim J, Xie J, Breen SL, Jaffray DA. Auto-segmentation of normal and target structures in head and neck CT images: A feature-driven model-based approach. *Medical Physics*. 2011; 38(11):6160–70. [PubMed: 22047381]
19. Zou KH, Warfield SK, Bharatha A, et al. Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. *Academic Radiology*. 2004; 11(2):178–89. [PubMed: 14974593]
20. Mory B, Ardon R, Yezzi AJ, Thiran JP. Non-Euclidean image-adaptive Radial Basis Functions for 3D interactive segmentation. *Book Non-Euclidean image-adaptive Radial Basis Functions for 3D interactive segmentation*. 2009:787–94.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33(1):159–74. [PubMed: 843571]
22. Heimann T, van Ginneken B, Styner MA, et al. Comparison and evaluation of methods for liver segmentation from CT datasets. *IEEE transactions on medical imaging*. 2009; 28(8):1251–65. [PubMed: 19211338]
23. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin*. 1979; 86(2):420–8. [PubMed: 18839484]
24. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 1(8476):307–10. [PubMed: 2868172]
25. Dachman AH, MacEneaney PM, Adedipe A, Carlin M, Schumm LP. Tumor size on computed tomography scans: is one measurement enough? *Cancer*. 2001; 91(3):555–60. [PubMed: 11169938]
26. Hopper KD, Kasales CJ, Egli KD, et al. The impact of 2D versus 3D quantitation of tumor bulk determination on current methods of assessing response to treatment. *Journal of computer assisted tomography*. 1996; 20(6):930–7. [PubMed: 8933793]
27. 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):416–9. [PubMed: 12409574]

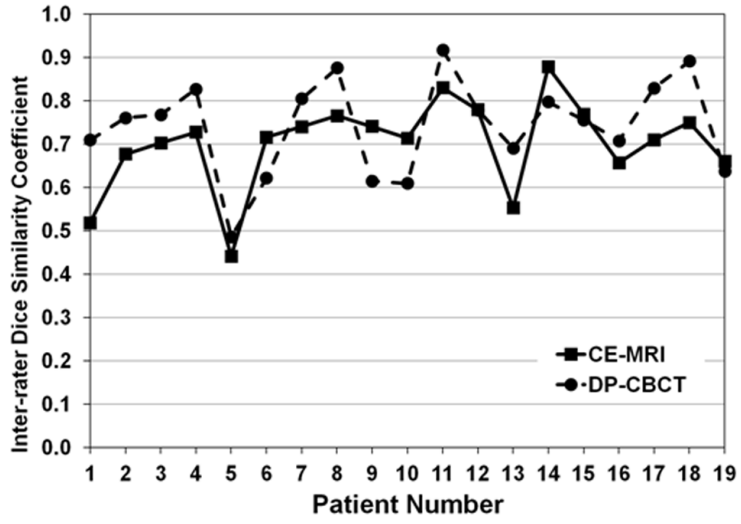




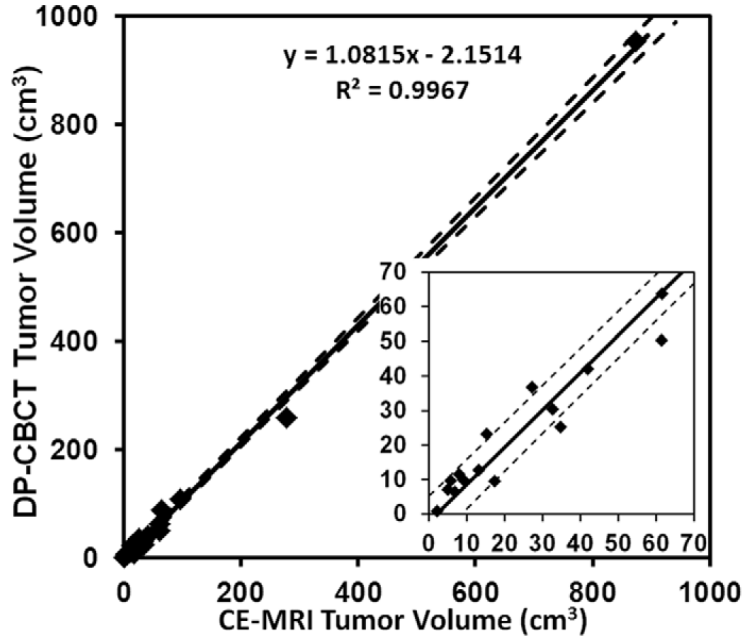
**Figure 1.** Representative CE-MRI (left column) and DP-CBCT (right column) images after tumor segmentation by one reader (□). The first row shows a 2D axial slice of the tumor segmented by the “elastic band.” The second row shows the 3D projected tumor volume over the same axial slice above.



**Figure 2.** Least squares linear regression curve for mean tumor volume measured from CE-MRI and DP-CBCT. The solid line is modeled by the equation shown on the figure and the dotted lines represent the 95% confidence interval. The inset is a zoom to show the smaller tumor volumes.



**Figure 3.** CE-MRI (A) and DP-CBCT (B) images after tumor segmentation was done by the three readers. The red, green, and yellow lines indicate where each reader segmented and the highlighted blue-green region represents the common area segmented by all readers. The tumor segmentation is magnified and represented in the white box. The inter-rater DSC was 0.75 for CE-MRI and 0.89 for DP-CBCT.



**Figure 4.** Graphical representation of inter-rater DSC values for each patient on each imaging modality, CE-MRI (solid line) and DP-CBCT (dashed lined).

**Table 1**

## Patient's characteristics and HCC staging

	Value (n or mean $\pm$ SD)
<b>Baseline Characteristics</b>	
Number of patients/Number of tumors	19/19
Age (year)	61 $\pm$ 13
Sex (Male/Female)	13/6
Etiology (Alcohol/HCV/HBV/NASH/Cryptogenic)	0/10/2/2/5
ECOG performance analysis (0/1/2/3/4)	19/0/0/0/0
Type (Unifocal/Bifocal/Multiple/Diffuse)	6/4/8/1
Portal vein thrombosis	2
<b>HCC staging</b>	
Child-Pugh class (A/B/C)	8/11/0

**Table 2**

Mean tumor volume and segmentation time measured on CE-MRI and DP-CBCT by three readers for 19 patients. Volume values are in cubic centimeters ( $\text{cm}^3$ ) and time values are in seconds (s).

Patient Number	Volume ( $\text{cm}^3$ ) $\pm$ SD			Time (s) $\pm$ SD		
	CE-MRI	DP-CBCT	p-value	CE-MRI	DP-CBCT	p-value
1	15 $\pm$ 5	23 $\pm$ 4	0.26	89 $\pm$ 69	120 $\pm$ 46	0.65
2	62 $\pm$ 23	50 $\pm$ 14	0.41	94 $\pm$ 76	161 $\pm$ 41	0.48
3	8 $\pm$ 1	12 $\pm$ 3	0.24	60 $\pm$ 72	40 $\pm$ 12	0.61
4	13 $\pm$ 5	13 $\pm$ 1	0.87	34 $\pm$ 31	89 $\pm$ 26	0.10
5	17 $\pm$ 10	10 $\pm$ 5	0.19	57 $\pm$ 52	77 $\pm$ 42	0.14
6	33 $\pm$ 2	30 $\pm$ 14	0.82	57 $\pm$ 23	76 $\pm$ 7	0.30
7	35 $\pm$ 3	25 $\pm$ 3	0.02	51 $\pm$ 35	156 $\pm$ 36	0.15
8	6 $\pm$ 2	10 $\pm$ 1	0.17	35 $\pm$ 44	30 $\pm$ 4	0.85
9	42 $\pm$ 5	42 $\pm$ 7	1.00	58 $\pm$ 38	67 $\pm$ 54	0.83
10	7 $\pm$ 1	6 $\pm$ 2	0.64	21 $\pm$ 9	53 $\pm$ 26	0.18
11	5 $\pm$ 1	7 $\pm$ 0.3	0.12	64 $\pm$ 67	17 $\pm$ 4	0.34
12	873 $\pm$ 17	954 $\pm$ 86	0.29	173 $\pm$ 62	214 $\pm$ 199	0.80
13	9 $\pm$ 1	10 $\pm$ 1	0.64	66 $\pm$ 49	51 $\pm$ 29	0.22
14	278 $\pm$ 7	259 $\pm$ 9	0.09	84 $\pm$ 41	90 $\pm$ 30	0.46
15	96 $\pm$ 20	109 $\pm$ 22	0.59	90 $\pm$ 62	72 $\pm$ 37	0.73
16	65 $\pm$ 18	88 $\pm$ 15	0.16	60 $\pm$ 9	109 $\pm$ 20	0.26
17	62 $\pm$ 19	64 $\pm$ 5	0.87	51 $\pm$ 38	96 $\pm$ 14	0.09
18	27 $\pm$ 8	37 $\pm$ 2	0.19	69 $\pm$ 58	77 $\pm$ 19	0.73
19	2 $\pm$ 1	1 $\pm$ 0.2	0.10	43 $\pm$ 27	17 $\pm$ 2	0.20
<b>Mean <math>\pm</math> SD:</b>	<b>87 <math>\pm</math> 8</b>	<b>92 <math>\pm</math> 10</b>	<b>-</b>	<b>66 <math>\pm</math> 45</b>	<b>85 <math>\pm</math> 34</b>	<b>0.19</b>

**Table 3**

Dice Similarity Coefficient (DSC) for each patient according to imaging modality: CE-MRI and DP-CBCT.

Patient number	DSC Inter-rater	
	CE-MRI	DP-CBCT
1	0.52 ± 0.19	0.71 ± 0.02
2	0.68 ± 0.08	0.76 ± 0.05
3	0.70 ± 0.05	0.77 ± 0.03
4	0.73 ± 0.05	0.83 ± 0.01
5	0.44 ± 0.16	0.49 ± 0.10
6	0.72 ± 0.09	0.62 ± 0.11
7	0.74 ± 0.02	0.80 ± 0.04
8	0.77 ± 0.06	0.88 ± 0.2
9	0.74 ± 0.10	0.61 ± 0.07
10	0.71 ± 0.05	0.61 ± 0.08
11	0.83 ± 0.02	0.92 ± 0.02
12	0.78 ± 0.02	0.78 ± 0.02
13	0.55 ± 0.08	0.69 ± 0.15
14	0.88 ± 0.00	0.80 ± 0.04
15	0.77 ± 0.02	0.76 ± 0.03
16	0.66 ± 0.07	0.71 ± 0.04
17	0.71 ± 0.06	0.83 ± 0.04
18	0.75 ± 0.06	0.89 ± 0.03
19	0.66 ± 0.15	0.64 ± 0.08
<b>Average</b>	<b>0.70 ± 0.07</b>	<b>0.74 ± 0.05</b>

**Table 4**

The mean Dice Similarity Coefficient (DSC) for each reader according to imaging modality: CE-MRI and DP-CBCT.

Reader	Intra-reader DSC	
	CE-MRI	DP-CBCT
( )	0.70 ± 0.11	0.75 ± 0.12
( )	0.69 ± 0.13	0.73 ± 0.13
( )	0.71 ± 0.13	0.75 ± 0.11