Three-Dimensional Measurement Of hepatocellular carcinoma Ablation Zones And Margins For Predicting Local Tumor Progression

Arnaud Hocquelet, Hervé Trillaud, Nora Frulio, Panteleimon Papadopoulos, P. Balageas, C. Salut, Marie Meyer, Jean-Frédéric Blanc, Michel Montaudon, Baudouin Denis de Senneville

CLINICAL STUDY

Three-Dimensional Measurement of Hepatocellular Carcinoma Ablation Zones and Margins for Predicting Local Tumor Progression

Arnaud Hocquelet, MD, Hervé Trillaud, MD, PhD, Nora Frulio, MD, Panteleimon Papadopoulos, MD, Pierre Balageas, MD, Cécile Salut, MD, Marie Meyer, MD, Jean-Frédéric Blanc, MD, PhD, Michel Montaudon, MD, PhD, and Baudouin Denis de Senneville, PhD

ABSTRACT

Purpose: To propose a postprocessing technique that measures tumor surface with insufficient ablative margins ($\leq$ 5 mm) on magnetic resonance (MR) imaging to predict local tumor progression (LTP) following radiofrequency (RF) ablation.

Materials and Methods: A diagnostic method is proposed based on measurement of tumor surface with a margin $\leq$ 5 mm on MR imaging. The postprocessing technique includes fully automatic registration of pre- and post–RF ablation MR imaging, a semiautomatic segmentation of pre–RF ablation tumor and post–RF ablation volume, and a subsequent calculation of the three-dimensional exposed tumor surface area. The ability to use this surface margin $\leq$ 5 mm to predict local recurrence at 2 years was then tested on 16 patients with cirrhosis who were treated by RF ablation with a margin $\leq$ 5 mm in 2012: eight with LTP matched according to tumor size and number and α-fetoprotein level versus eight without local recurrence.

Results: The error of estimated tumor surface with a margin $\leq$ 5 mm was less than 12%. Results of a log-rank test showed that patients with a tumor surface area $>$ 425 mm^2 had a 2-year LTP rate of 77.5%, compared with 25% for patients with a tumor surface area $\leq$ 425 mm^2 ($P = .018$).

Conclusions: This proof-of-concept study proposes an accurate and reliable postprocessing technique to estimate tumor surface with insufficient ablative margins, and underscores the potential usefulness of tumor surface with a margin $\leq$ 5 mm to stratify patients with HCC treated by RF ablation according to their risk of LTP.

ABBREVIATIONS

AFP = α-fetoprotein, DSC = DICE similarity coefficient, HCC = hepatocellular carcinoma, LTP = local tumor progression, RF = radiofrequency, ROI = region of interest, 3D = three-dimensional, THRIVE = T1 high resolution isotropic volume excitation

Percutaneous thermal ablation, especially radiofrequency (RF) ablation, is a safe and effective treatment for hepatocellular carcinoma (HCC) (1,2), providing similar outcomes to surgical resection (2). However, despite progress in RF devices (3–6), the prognosis of cirrhotic patients with HCC treated by RF ablation as first-line is still impaired by local tumor progression (LTP) occurring within 2 years after RF ablation (7).

LTP occurs in 10%–21% of patients and is associated with poor prognosis (2,6). The likelihood of local progression increases with tumor size as a result of satellite nodules (8), which is why a 5-mm minimum ablation margin is recommended by several studies (9,10). However, some of these studies (9,11,12) used a rigid registration that is not optimal for the liver because of significant deformations by respiratory movements

From the Departments of Diagnostic and Interventional Imaging (A.H., H.T., N.F., P.P., P.B., C.S., M.Me.) and HepatoGastroenterology and Digestive Oncology (J.-F.B.), Hôpital Saint-André, Centre Hospitalier Universitaire de Bordeaux, Bordeaux, France; Department of Diagnostic and Interventional Imaging (M.Mo.), Hôpital Haut-Lévêque, Centre Hospitalier Universitaire de Bordeaux, Pessac, France; and Institut de Mathématiques de Bordeaux (B.D.d.S.), Unité Mixte de Recherche 5251, Centre National de Recherche Scientifique/Universitaire de Bordeaux, Talence, France. Received October 21, 2015; final revision received February 25, 2016. Address correspondence to A.H., Service d'Imagerie Médicale, Hopital Saint André, 1 rue Jean Burguet, 33000 Bordeaux, France; E-mail: arnaud.hocquelet@gmail.com

None of the authors have identified a conflict of interest.

Figure E1 and Appendices E1–E3 are available online at *www.jvir.org*.

and the ablation zone (13). Moreover, these studies assessed only two-dimensional minimal margins, failing to take into account the heterogeneous shape of the ablative volume around the tumor and that the minimal margin is not necessarily indicative of the margin sizes all around the tumor. This could explain the fact that LTP occurred in "only" as many as 30%–50% of patients with an ablation margin ≤ 5 mm (14). As illustrated in Figure 1, two patients with an identical minimal ablation margin can have different zones at risk of LTP. A measure of the tumor surface with insufficient ablative margins that represents the three-dimensional (3D) ablation margin therefore may be a better way to assess RF ablation accuracy during the procedure and follow-up and to improve the prognosis.

The aim of the present study is to propose a full post-processing technique designed to measure the tumor surface with insufficient ablation margin ≤ 5 mm and to assess its potential to predict LTP at 2 years after RF ablation.

MATERIALS AND METHODS

Patients

Institutional review board approval was obtained, and consent was waived for this retrospective study. From our hospital database, we retrospectively included patients with cirrhosis who underwent RF ablation in 2012 for a single HCC with minimum ablative margins ≤ 5 mm and experienced LTP (n = 8). Based on propensity score (15), according to tumor size and number and α-fetoprotein (AFP) level, we matched a control group of eight patients treated in our institution in 2012 with no LTP and a minimum treatment margin ≤ 5 mm. Overall, 16 patients with cirrhosis were included to test our postprocessing technique, eight in each group (with and without LTP). All tumors met the Milan criteria (16), so the maximum tumor size was 5 cm.

Diagnosis and Staging of HCC and RF Ablation

Noninvasive criteria of the European Association for the Study of the Liver were used to diagnose HCC (16). All RF ablation procedures were performed percutaneously under general anesthesia. Real-time ultrasound (US) with a 4-MHz probe was chosen as the first-line guidance modality for all patients. A senior interventional radiologist (at least 5 y of experience) performed RF ablation with use of monopolar expandable LeVeen needles (RF 3000; Boston Scientific, Marlborough, Massachusetts). Thermal ablation was performed according to the manufacturer's instructions. RF ablation was performed under US surveillance to determine sufficient overlap of the hyperechoic ablation zone on the HCC lesion. Follow-up after ablation was performed with MR imaging at 1 month.

LTP (17) describes the appearance of tumor foci at the edge of the ablation zone after at least one contrast-enhanced follow-up study has documented adequate ablation and an absence of viable tissue in the target tumor and surrounding ablation margin based on imaging criteria. This term applies regardless of whether tumor foci were discovered early or late in the course of imaging follow-up.

Experimental Setup

MR imaging acquisition. Pre- and post–RF ablation MR imaging was performed on a 1.5-T clinical unit (Tesla Archiva 1.5; Philips, Best, The Netherlands). MR protocol consisted of axial T2-weighted sequences with and without fat saturation, a T1-weighted in- and out-of-phase sequence, and a 3D T1-weighted sequence with fat

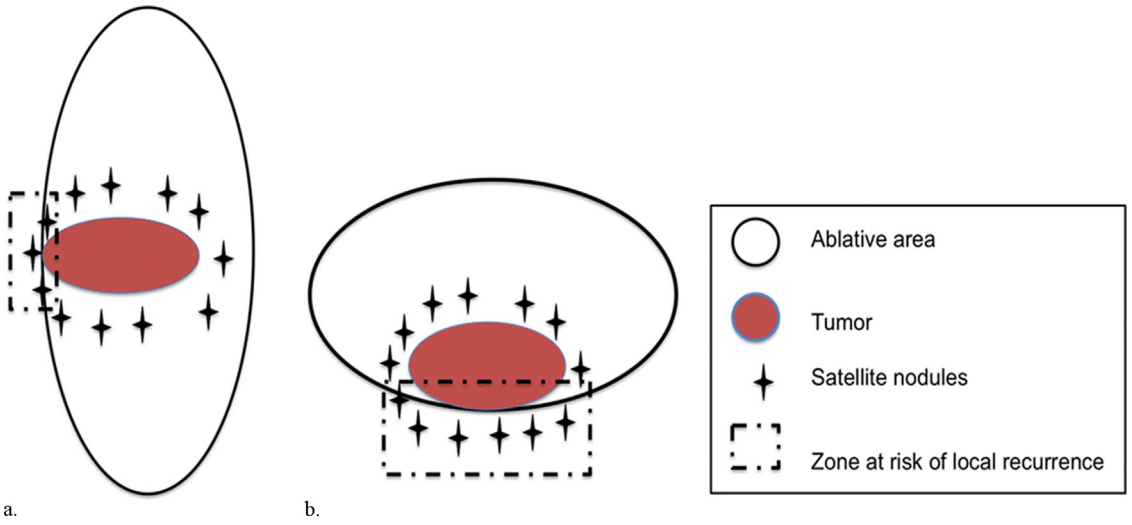

Figure 1. Schematic illustration of two typical ablative surfaces depicts similar minimal margins but different tumor areas exposed to insufficient margins. Minimal margins equal to 0 mm are shown with a small **(a)** and large **(b)** tumor area exposed to insufficient margins. Note that the probability of untreated satellite nodules is higher in the second case **(b)**.

saturation before and after intravenous contrast agent injection (gadoterate meglumine; Dotarem; Guerbet, Villepinte, France) in arterial, portal, and late phases. Breath-hold transverse 3D T1 high-resolution isotropic volume excitation (THRIVE) sequences were obtained with the following parameters: repetition/echo times, 3.9/1.82 ms; flip angle, 10°; field of view, 346 mm (it could be adapted to cover all liver parenchyma); voxel size, 2.11 × 2.07 × 2.2 mm^3; and reconstruction matrix, 164 × 142. Post–RF ablation MR imaging studies were performed 1 month after ablation and then every 3 months. THRIVE was the only sequence used to perform segmentation and analysis.

Estimation of 3D liver deformation/registration. We employed a 3D image registration algorithm to compensate for the liver deformation and inherent positioning errors arising between the two MR imaging sessions. The registration was designed to provide voxel-wise 3D motion estimates to reflect complex 3D deformations such as the ones that the liver may undergo. We employed the modality-independent neighborhood descriptor initially proposed by Heinrich et al (18) and Østergaard et al (19) for multimodal deformable registration. We registered the post–RF ablation portal-phase THRIVE MR imaging sequence on the pre–RF ablation arterial-phase THRIVE MR imaging sequence.

Tumor and ablation area segmentation. An anisotropic spatial diffusion filter was applied to reduce noise while preserving image anatomic features (20). Additional details are provided in **Appendix E1** (available online at *www.jvir.org*). The tumor was then segmented on THRIVE sequences by using the acquisition phase providing the best contrast between the region of interest (ROI) and the remaining liver, ie, arterial phase for HCC and portal phase for ablation area segmentation. Tumor and ablation area segmentation was performed with ITK-SNAP freeware (www.itksnap.org) (21), which performs fast and accurate semiautomatic segmentation by using region-competition snakes. We are aware of no previous study that has used tumor and ablation area segmentation to perform automatic calculation of treatment margins. Manual segmentation or measurement of margins could produce highly variable results and are time-consuming. A radiologist with 20 years of experience validated all segmentations.

Calculation of tumor area exposed to insufficient margins. At this point, our aim was to calculate the 3D tumor perimeter exposed to post–RF ablation margins ≤ 5 mm based on two digital data sets. Unfortunately, an exact recovery of the perimeter is not possible because of the loss of information arising from the digitization process (**Fig E1** [available online at *www.jvir.org*]). Further information about voxel-based surface area estimation methods can be found in previous studies (22–24) and in **Appendix E2** (available online at *www.jvir.org*) (25–35). To calculate the tumor area exposed to a specific postablation margin, we used an extended application of the surface area estimator of Mullikin and Verbeek (36). The estimator is local and operates directly on voxels, and is therefore a fast and an easy-to-implement method that obtains very reasonable accuracy. The pre- and registered postablation MR images were first "regridded" by using a trilinear interpolation to achieve an isotropic 1 × 1 × 1 mm^3 voxel aspect. Note that, for the remainder of the section, this aspect ratio was maintained for all data of interest. Although this task was not computationally mandatory, this rendered coding and debugging easier. The following tasks were then performed sequentially.

First, the algorithm began by detecting all "surface voxels," ie, tumor voxels in the segmented pre–RF ablation images that were 6-connected to background voxels. Second, for each surface voxel, the minimal Euclidean distance toward the external edge in the post–RF ablation image was computed. Surface voxels in the outer pre–RF ablation tumor contour exposed to a postablation margin less than 5 mm were considered as exposed surface voxels. Third, similar to the method of Mullikin and Verbeek (36), surface area weights were assigned to the exposed surface voxels to remove local length estimation errors. (More information on the calculation of the surface area weights is provided in previous studies (37,38). The total surface area could then be estimated by summing the area contributed by all exposed surface voxels.

Quality assessment of postprocessing tasks. Quality assessment of the registration was analyzed by qualitative and quantitative strategies. For the qualitative analysis, the registration was first analyzed by computing the DICE similarity coefficient (DSC) between the ROI encompassing the liver in the registered volume and its counterpart in the reference frame (39) (**Appendix E3** [available online at *www.jvir.org*]) (40,41). A DSC of 1 indicates an ideal registration (ie, perfect ROI matching). For the quantitative analysis, the validity of the estimated motion was then tested by manually tracking three blood vessel bifurcations in the liver (**Appendix E2** [available online at *www.jvir.org*]). Assessment of the tumor and ablative area segmentation process was also performed by using the DSC (**Appendix E2** [available online at *www.jvir.org*]).

Assessment of tumor surface area exposed to insufficient postablation margins ≤ 5 mm. The estimation of the tumor surface area was assessed by using a typical case scenario initially defined in the continuous domain, which includes a preablation and a postablation scenario. Subsequent quantification of the

continuous scenes was performed to produce pre- and post–RF ablation digital data sets; in this way, the exact analytic tumor exposition surface area could be compared versus the one derived from the digital images, the latter being obtained by using the method described in the previous section. The typical case scenario is described in details in **Appendix E3** (available online at *www.jvir.org*). Briefly, the tumor in the pre–RF ablation images was simulated using a sphere, which was embedded in a larger sphere representing the postablation margins. The percentage of variation between the estimated and analytic exposed surface areas was calculated for studied spherical tumor diameters of 20, 30, and 40 mm.

Statistical Analysis

Data are expressed as medians and interquartile ranges and were compared by using the Mann–Whitney *U* test. Percentages were compared by χ^2 test. All patients were followed up over a period of 24 months or until death. Survival without local recurrence was defined as the time from RF ablation treatment to the time of local recurrence or 24 months. Progression-free survival was defined from the time of RF ablation treatment to status at 24 months or time of recurrence (of any type). Optimal cutoffs to predict LTP were selected by using the "Survival ROC" package for R software (R Foundation for Statistical Computing, Vienna, Austria). Survival curves were computed by Kaplan–Meier method and compared by log-rank test. Cox regressions were performed to assess the correlation between tumor surface with margins $\leq$ 5 mm and LTP. A P value $<$ .05 was considered statistically significant. Statistical analyses were performed with Stata (version 13; StataCorp, College Station, Texas) and R software.

RESULTS

Patients

The mean follow-up time was 2.2 years (95% confidence interval, 1.76–2.63 y). As shown in **Table 1**, there were no differences between the two groups (with and without LTP) in terms of tumor size or AFP level. However, albumin levels were higher in the LTP group.

Accuracy of Registration

The computation time required for the registration process was less than 7 minutes for each patient tested on our platform. The mean DSC of the errors increased from 0.7 ± 0.17 to 0.94 ± 0.02 with the use of the implemented registration algorithm. More than 75% of errors were reduced to less than 2 mm (ie, one voxel) when the implemented motion compensation strategy was applied. In that case, the average positioning error also decreased from 17.3 mm ± 8 to 2 mm ± 0.9 (**Fig 2**). However, the precision of the manually tracked displacements was limited by the discrete nature of the acquired images. The intrarater variation in positioning landmarks was 2 mm (ie, voxel size). Therefore, tracking errors of as much as one voxel were to be expected in practice, as variations in manual landmark positioning of approximately 2 mm were measured.

Segmentation Repeatability

Semiautomatic segmentations of the preablation tumor and the postablation margins were achieved quickly with ITK-SNAP (< 1 min) in all cases. The average DSC was calculated between two repeated semiautomatic segmentations of the same volume (ie, tumor and ablation area). The DSC was constantly higher than 0.9. The mean DSC values on pre- and postablation scans were 0.98 ± 0.03 and 0.96 ± 0.03, respectively.

Table 1. Demographics and Clinical Characteristics of Patients who Received RF Ablation for HCC (N = 16)

Characteristic	Local Recurrence (n = 8)	No Local Recurrence (n = 8)	*P* Value
Age (y)			.752
Median	62.1	64.5	
IQR	58–74	61–70	
Male sex	8 (100)	8 (100)	.302
ALT (UI/L)			.874
Median	49.5	47	
IQR	25.5–193	27.5–77.5	
Platelet count ($\times 10^9$/L)			.431
Median	92.5	91	
IQR	66.5–151	66–108	
Albumin (g/dL)			.030
Median	37.5	33.4	
IQR	35.5–41.1	30.5–36	
Bilirubin (mg/dL)			.900
Median	29	25.5	
IQR	13.5–37.5	16–38	
INR			.896
Median	1.1	1.1	
IQR	1.04–1.12	1.07–1.13	
Child class B disease	1 (12.5)	5 (62.5)	.039
Etiology			.097
Viral	5	2	
Alcohol	3	6	
AFP (ng/mL)			.831
Median	10	8.45	
IQR	4–138	4.4–15.1	
Tumor size (mm)			.598
Median	29	25	
IQR	18–35	21.5–27	

AFP = α-fetoprotein; ALT = alanine aminotransferase; HCC = hepatocellular carcinoma; INR = International Normalized Ratio; IQR = interquartile range; RF = radiofrequency.

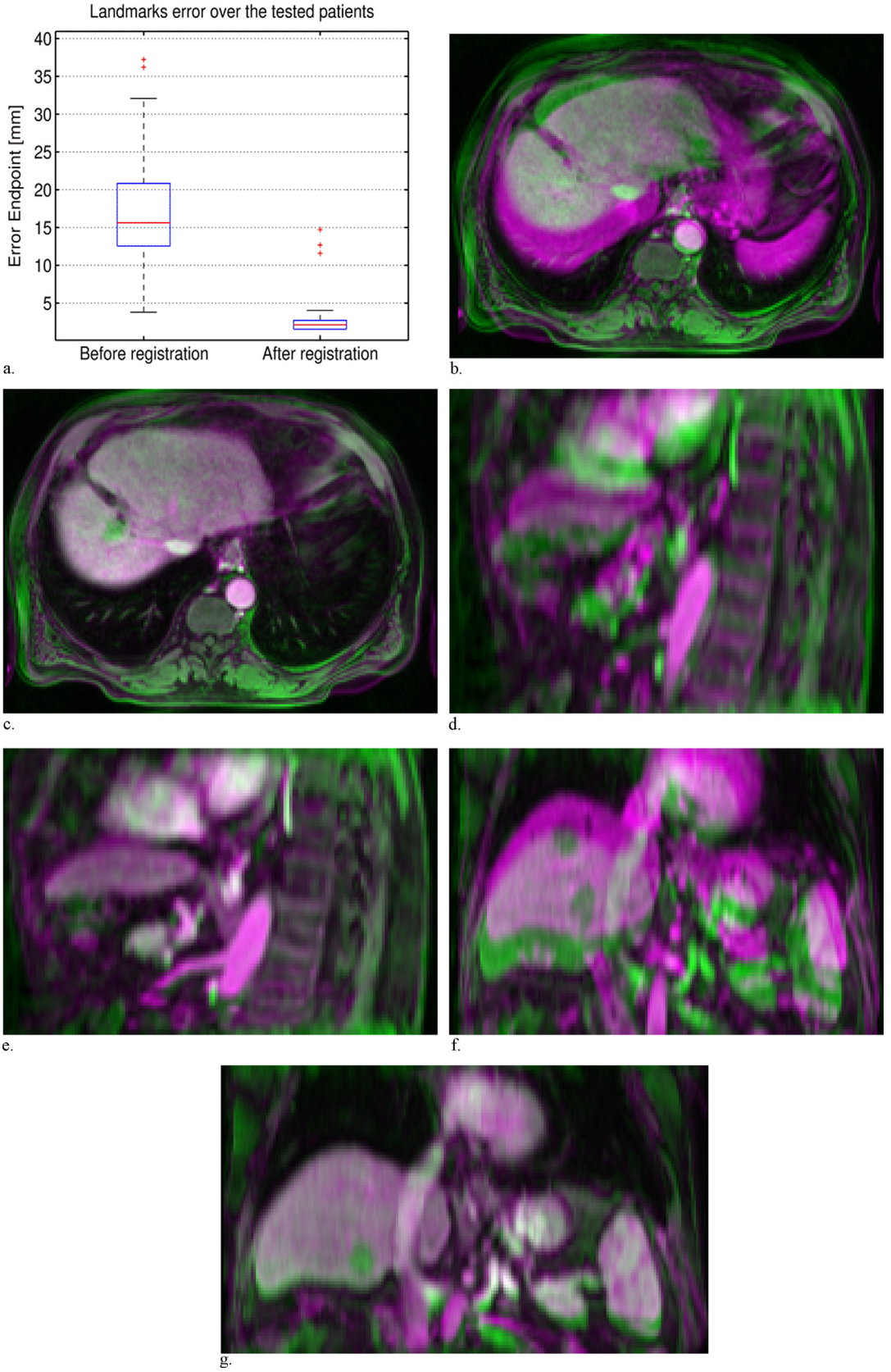

Figure 2. (**a**) Box plots of positional errors that would affect the analysis of tumor margin in the absence (left plot) and presence (right plot) of a motion-compensation strategy. The values on the ordinate indicate the estimated signed displacements in millimeters. **(b–g)** Images detail typical registration results obtained in one patient: superimposed pre- (green layer) and post–RF ablation (purple layer) anatomic images without (**b**, **d**, and **f**) and with (**c, e,** and **g**) the use of the automatic registration process implemented in the study. Transverse (**b** and **c)**, sagittal (**d** and **e**), and coronal planes (**f** and **g**) are shown.

Estimation of Exposed Tumor Surface Area

Figure 3 shows an analysis of the precision of the implemented algorithm for the estimation of exposed tumor surface area under various experimental conditions. For each tested tumor diameter, a scatter plot is shown. The error remained well below 12% for voxel sizes lower than $2 \times 2 \times 2$ mm^3 in all simulated case scenarios.

Tumor Area Exposed to Margins ≤ 5 mm

The mean tumor surface with margins ≤ 5 mm was 648 mm^2 ± 844 (standard deviation), and the median was 388 mm^2 (interquartile range, 105–960 mm^2). In the group with LTP, mean tumor surface with margins ≤ 5 mm was 976 mm^2 ± 1,064 (median, 748 mm^2; range, 329–1,045 mm^2), compared with 319 mm^2 ± 385 (median, 157 mm^2; range, 64–456 mm^2) in the group without LTP ($P = .046$). The area under the curve of tumor surface with margins ≤ 5 mm to predict recurrence was 0.830. The best cutoff value according to the Youden index was 425 mm^2.

Prediction of Local Tumor Recurrence

Based on a cutoff value of 425 mm^2, two groups with significantly different 2-year LTP cumulative incidences were identified (Fig 4). The cumulative 1- and 2-year LTP incidences in the group with tumor surface with insufficient margins ≤ 425 mm^2 were 12.5% and 25%, respectively, versus 62.5% and 77.5%, respectively, in the group with tumor surface with insufficient margins > 425 mm^2 ($P = .018$). The univariate hazard ratio of tumor surface with insufficient margins > 425 mm^2 to predict 2-year LTP was 6.71 (95% confidence interval, 1.27–35.5; $P = .014$).

Overall Survival and Disease Recurrence (Any Pattern)

The 1- and 2-year cumulative incidences of disease recurrence in the group with insufficient margins ≤ 425 mm^2 were 12.5% and 50%, respectively, versus 75% and 87.5%, respectively, in the group with insufficient margins > 425 mm^2 ($P = .033$). One- and 2-year overall survival rates were higher for patients with insufficient margins ≤ 425 mm^2 than for patients with insufficient margins > 425 mm^2, at 100% and 100%, respectively, versus 87.5% and 37.5%, respectively ($P = .008$).

DISCUSSION

The postprocessing technique developed for the present study requires only two 3D MR imaging volumes of the pathologic region as input, one being obtained before RF ablation and the other afterward. Careful attention was paid to the ability of each task of the technique to assess the specific endpoint, ie, the 3D tumor perimeter exposed to post–RF ablation margins ≤ 5 mm. The results demonstrate that, for patients considered at risk for local recurrence (ie, with minimal ablative margins ≤ 5 mm), calculating the tumor surface exposed to this insufficient treatment margin enabled the stratification of patients into two groups with significantly different incidences of 2-year LTP.

First, an accurate assessment of ablative margins requires an accurate registration to compensate for inherent liver displacements occurring between the pre- and post–RF ablation MR images. It must be emphasized that several studies performed MR/MR or CT/CT scan registration to assess minimal treatment margins (11,42,43) but mainly used rigid registration. Except for the study of Kim et al (42), none of these studies performed a quantitative evaluation of registration accuracy

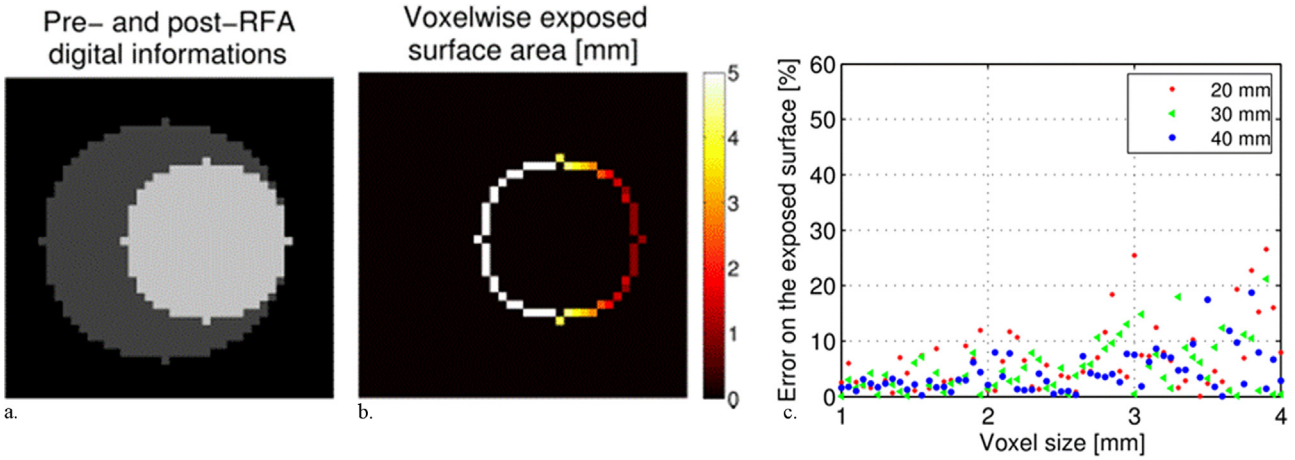

Figure 3. Assessment of the algorithm employed for the estimation of tumor surface area from pre- and post–RF ablation 3D digital data sets. **(a)** Typical cross-section of a simulated 3D digital data (x–y plane) that depicts the tumor (light gray) superimposed on the postablation volume (dark gray). Here, the tumor diameter is 20 mm and the simulated voxel size is $1 \times 1 \times 1$ mm^3. **(b)** Corresponding minimal Euclidean distance map from each surface voxel (ie voxels located in the outer preablation tumor contour) toward the external edge of the postablation margins. **(c)** Error of estimated exposed surface obtained for various simulated voxel and tumor sizes.

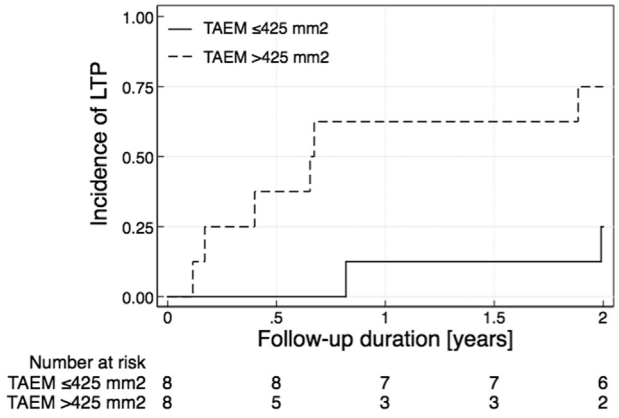

Figure 4. Graph of LTP according to insufficient margin surfaces $\leq$ 425 mm^2 (solid line) and $>$ 425 mm^2 (dashed line).

even though it is of great interest when measuring millimeter margins. In addition, in the present study, we employed an image registration algorithm based on the modality-independent neighborhood descriptor (16,17). The benefit of this algorithm lies in its robustness against possible gray-level intensity variations, which are likely to occur between the two MR imaging sessions as a result of signal intensity changes corresponding to the phase of acquisition (ie, not attributed to motion). The implemented algorithm offered accurate voxel-wise estimates of 3D liver deformation: a 2-mm difference in distance was measured between landmarks after registration that could easily be explained by the intrarater variation of approximately 2 mm in landmark placement (considering that the landmarks are manually located with a precision equal to the voxel size). Finally, tumor surface with insufficient margins could be measured from the segmentations of the tumor and ablation area. Tumor surface error fluctuations of 5%–10% were observed, even for similar voxel sizes, as a result of image aliasing considerations (ie, amplitude of scatter plots in **Fig 3c**). By using the proposed case scenarios for validation, we found a potential error of $<$ 12% with the use of a 2 $\times$ 2 $\times$ 2 mm^3 voxel size for all tested tumor diameters, which is considered sufficient for our diagnostic endpoint.

The results demonstrate significantly higher local and global tumor progression rates in the group with tumor surface with insufficient margin $>$ 425 mm^2, which supports the proposal that tumors with large areas with margins $\leq$ 5 mm have higher risk of recurrence arising from untreated satellite nodules. The present results highlight the utility of this concept based on tumor surface exposed to margins $\leq$ 5 mm, and not only on two-dimensional minimal margins. The probability of LTP for the population with margins $\leq$ 5 mm is heterogeneous, and, as evidenced here, tumor surface with insufficient margin can be used to stratify these patients according to their LTP risk. The stratification according to local recurrence risk could be useful to choose which patients need to be treated repeatedly. Indeed, the decision to perform a new RF ablation is problematic in patients with complete response according to modified Response Evaluation Criteria In Solid Tumors (44) and margins $\leq$ 5 mm, considering that 50%–70% of these patients will not experience LTP (14,30). In addition, RF ablation is mainly performed under general anesthesia, so patients who are to undergo repeat treatment should be carefully chosen in view of the risk of anesthesia complications. As our technique is mainly automated and takes fewer than 10 minutes for all the processes, it is compatible for use in a clinical environment.

In the present study, we used our postprocessing technique to predict LTP, but this postprocessing technique can obviously be used during RF ablation procedures (under CT or cone-beam CT or guidance) to provide fast semiautomatic 3D margin computation. Hence, RF probes could be moved to insufficiently treated areas to complete the ablation during the same session, which should drastically reduce LTP arising from insufficient margins and multiple RF ablation sessions to treat the same nodule. This stratification could also be used to modulate follow-up. Indeed, in cases of patients at high risk with large tumor surface at risk, a shorter follow-up interval could be proposed to detect recurrence at an earlier stage, allowing curative treatment and avoiding worsening prognosis as a result of LTP (45).

Several limitations should be emphasized in this retrospective proof-of-concept study, including its small sample size, even though we used propensity-score matching to avoid selection bias (2,15). Although encouraging results were obtained, a per–ablation procedure validation is needed. The low number of patients does not allow multivariate analysis, which would be of great interest for the validation of the method. In addition, although only patients with ablative margins $\leq$ 5 mm were studied, an analysis including patients with larger margins should also be performed.

In conclusion, this proof-of-concept study describes a unique approach to predict LTP based on the measurement of tumor surface exposed to insufficient margins. This easy-to-perform measurement can be used to stratify patients generally considered at risk of LTP (ie, with ablative margin $\leq$ 5 mm) following RF ablation for HCC according to whether they are at low or high risk. This postprocessing technique could also be used for 3D margin computation during RF ablation procedures.

ACKNOWLEDGMENT

The authors thank Pippa McKelvie-Sebileau for medical editorial services.

REFERENCES

1. N'Kontchou G, Mahamoudi A, Aout M, et al. Radiofrequency ablation of hepatocellular carcinoma: long-term results and prognostic factors in 235 Western patients with cirrhosis. Hepatol Baltim Md 2009; 50:1475–1483.

2. Hocquelet A, Balageas P, Laurent C, et al. Radiofrequency ablation versus surgical resection for hepatocellular carcinoma within Milan criteria: a study of 281 western patients. Int J Hyperthermia 2015; 31: 749–757.
3. Hocquelet A, Montaudon M, Balageas P, et al. Comparison of surgical resection and radiofrequency ablation for hepatocellular carcinoma: take care not to neglect radiofrequency technic and device. Ann Surg June 15, 2015. http://journals.lww.com/annalsofsurgery/Pages/default.aspx; published online. http://dx.doi.org/10.1097/SLA.0000000000001293.
4. Seror O. Percutaneous hepatic ablation: what needs to be known in 2014. Diagn Interv Imaging 2014; 95:665–675.
5. Seror O. Ablative therapies: advantages and disadvantages of radiofrequency, cryotherapy, microwave and electroporation methods, or how to choose the right method for an individual patient? Diagn Interv Imaging 2015; 96:617–624.
6. Cartier V, Boursier J, Lebigot J, et al. Radiofrequency ablation of hepatocellular carcinoma: mono or multipolar? J Gastroenterol Hepatol 2016; 31:654–660.
7. Ahmed M, Solbiati L, Brace CL, et al; International Working Group on Image-Guided Tumor Ablation; Interventional Oncology Sans Frontières Expert Panel; Technology Assessment Committee of the Society of Interventional Radiology; Standard of Practice Committee of the Cardiovascular and Interventional Radiological Society of Europe. Image-guided tumor ablation: standardization of terminology and reporting criteria–a 10-year update. J Vasc Interv Radiol 2014; 25:1691–1705.e4.
8. Sasaki A, Kai S, Iwashita Y, et al. Microsatellite distribution and indication for locoregional therapy in small hepatocellular carcinoma. Cancer 2005; 103:299–306.
9. Kim Y-S, Lee WJ, Rhim H, et al. The minimal ablative margin of radiofrequency ablation of hepatocellular carcinoma (> 2 and < 5 cm) needed to prevent local tumor progression: 3D quantitative assessment using CT image fusion. AJR Am J Roentgenol 2010; 195:758–765.
10. Liu C-H, Arellano RS, Uppot RN, et al. Radiofrequency ablation of hepatic tumours: effect of post-ablation margin on local tumour progression. Eur Radiol 2010; 20:877–885.
11. Wang X-L, Li K, Su Z-Z, et al. Assessment of radiofrequency ablation margin by MRI-MRI image fusion in hepatocellular carcinoma. World J Gastroenterol 2015; 21:5345–5351.
12. Sakakibara M, Ohkawa K, Katayama K, et al. Three-dimensional registration of images obtained before and after radiofrequency ablation of hepatocellular carcinoma to assess treatment adequacy. AJR Am J Roentgenol 2014; 202:W487–W495.
13. Korin HW, Ehman RL, Riederer SJ, et al. Respiratory kinematics of the upper abdominal organs: a quantitative study. Magn Reson Med 1992; 23:172–178.
14. Nakazawa T, Kokubu S, Shibuya A, et al. Radiofrequency ablation of hepatocellular carcinoma: correlation between local tumor progression after ablation and ablative margin. AJR Am J Roentgenol 2007; 188:480–488.
15. Austin PC. An Introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivar Behav Res 2011; 46:399–424.
16. 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; 56:908–943.
17. Ahmed M, Solbiati L, Brace CL, et al. Image-guided tumor ablation: standardization of terminology and reporting criteria–a 10-year update. Radiology 2014; 273:241–260.
18. Heinrich MP, Jenkinson M, Bhushan M, et al. MIND: modality independent neighbourhood descriptor for multi-modal deformable registration. Med Image Anal 2012; 16:1423–1435.
19. Østergaard Noe K, De Senneville BD, Elstrøm UV, et al. Acceleration and validation of optical flow based deformable registration for image-guided radiotherapy. Acta Oncol Stockh Swed 2008; 47:1286–1293.
20. Mirebeau J-M, Fehrenbach J, Risser L, Tobji S. Anisotropic diffusion in ITK. ArXiv:1503.00992. 2015. Available at: http://arxiv.org/abs/1503.00992. Accessed April 18, 2016.
21. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. NeuroImage 2006; 31:1116–1128.
22. Windreich G, Kiryati N, Lohmann G. Voxel-based surface area estimation: from theory to practice. Pattern Recognit 2003; 36:2531–2541.
23. Ziegel J, Kiderlen M. Estimation of surface area and surface area measure of three-dimensional sets from digitizations. Image Vis Comput 2010; 28:64–77.
24. Lindblad J. Surface area estimation of digitized 3D objects using weighted local configurations. Image Vis Comput 2005; 23:111–122.
25. Dorst L, Smeulders AWM. Length estimators for digitized contours. Comput Vis Graph Image Process 1987; 40:311–333.
26. Koplowitz J, Bruckstein AM. Design of perimeter estimators for digitized planar shapes. IEEE Trans Pattern Anal Mach Intell 1989; 11: 611–622.
27. Kiryati N, Kübler O. Chain code probabilities and optimal length estimators for digitized three-dimensional curves. Pattern Recognit 1995; 28:361–372.
28. Kiryati N, Székely G. Estimating shortest paths and minimal distances on digitized three-dimensional surfaces. Pattern Recognit 1993; 26: 1623–1637.
29. Jonas A, Kiryati N. Length estimation in 3-D using cube quantization. J Math Imaging Vis 1998; 8:215–238.
30. Verwer BJH. Local distances for distance transformations in two and three dimensions. Pattern Recognit Lett 1991; 12:671–682.
31. Grossmann R, Kiryati N, Kimmel R. Computational surface flattening: a voxel-based approach. IEEE Trans Pattern Anal Mach Intell 2002; 24: 433–441.
32. Windreich G, Kiryati N, Lohmann G. Voxel-based surface area estimation: from theory to practice. Pattern Recognit 2003; 36:2531–2541.
33. Ziegel J, Kiderlen M. Estimation of surface area and surface area measure of three-dimensional sets from digitizations. Image Vis Comput 2010; 28:64–77.
34. Lindblad J. Surface area estimation of digitized 3D objects using weighted local configurations. Image Vis Comput 2005; 23:111–122.
35. Mullikin JC, Verbeek PW. Surface area estimation of digitized planes. Bioimaging 1993; 1:6–16.
36. Mullikin JC, Verbeek PW. Surface area estimation of digitized planes. Bioimaging 1993; 1:6–16.
37. Koplowitz J, Bruckstein AM. Design of perimeter estimators for digitized planar shapes. IEEE Trans Pattern Anal Mach Intell 1989; 11: 611–622.
38. Kiryati N, Kübler O. Chain code probabilities and optimal length estimators for digitized three-dimensional curves. Pattern Recognit 1995; 28:361–372.
39. Carillo V, Cozzarini C, Perna L, et al. Contouring variability of the penile bulb on CT images: quantitative assessment using a generalized concordance index. Int J Radiat Oncol Biol Phys 2012; 84:841–846.
40. Carillo V, Cozzarini C, Perna L, et al. Contouring variability of the penile bulb on CT images: quantitative assessment using a generalized concordance index. Int J Radiat Oncol Biol Phys 2012; 84:841–846.
41. Wolf I, Vetter M, Wegner I, et al. The medical imaging interaction toolkit. Med Image Anal 2005; 9:594–604.
42. Kim KW, Lee JM, Klotz E, et al. Safety margin assessment after radiofrequency ablation of the liver using registration of preprocedure and postprocedure CT images. AJR Am J Roentgenol 2011; 196: W565–W572.
43. Shin S, Lee JM, Kim KW, et al. Postablation assessment using follow-up registration of CT images before and after radiofrequency ablation (RFA): prospective evaluation of midterm therapeutic results of RFA for hepatocellular carcinoma. AJR Am J Roentgenol 2014; 203:70–77.
44. Lencioni R, Llovet J. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. Semin Liver Dis 2010; 30:52–60.
45. Lee DH, Lee JM, Lee JY, et al. Radiofrequency ablation of hepatocellular carcinoma as first-line treatment: long-term results and prognostic factors in 162 patients with cirrhosis. Radiology 2014; 270: 900–909.

APPENDIX E1
IMAGE PREFILTERING FOR TUMOR AND ABLATIVE AREA SEGMENTATION

A prefiltering process was first employed to increase the robustness of the subsequent segmentation of the lesion and the post–radiofrequency ablation margin. The filter employed required two control parameters to be determined, namely the number of iterations and the diffusion sensitivity range. These parameters were determined based on the finding that an increment of each value improved the noise reduction performance but also subsequently resulted in a decrease in specificity of the image features. In the scope of this study, 15 iterations and a diffusion sensitivity range of 0.3 were found to be a good compromise and were commonly employed.

APPENDIX E2
CALCULATION OF TUMOR AREA EXPOSED TO INSUFFICIENT MARGINS

Our aim was to calculate the three-dimensional (3D) tumor perimeter exposed to post–radiofrequency margins $\leq$ 5 mm based on two digital data sets. Unfortunately, an exact recovery of the perimeter is not possible because of the loss of information arising from the digitization process (**Fig E1**). However, estimation is possible provided the following assumptions are met: mainly, the curve is smooth and contains a small number of corners, and the curve can be considered nearly straight within small neighborhoods. A number of methods have been proposed in the past to address the two-dimensional perimeter estimation problem (25,26). Recently, the theory of two-dimensional perimeter estimation has been extended to 3D curves (27–30), and plays a large role in 3D shape analysis techniques (31). Further information about this method can be found in a pervious study (32), which provides a comprehensive overview of voxel-based surface area estimation methods. Two other publications (33,34) provide more detailed presentations of certain specific aspects of voxel-based surface area estimation methods. In the present study, we used the surface area estimator of Mullikin and Verbeek (36) as described in the main text.

APPENDIX E3
QUALITY ASSESSMENT OF POSTPROCESSING TASKS

Quality Assessment of Motion Estimation

Qualitative analysis. The registration was first analyzed by computing the DICE similarity coefficient (DSC) between the region of interest (ROI) encompassing the liver in the registered volume (ROI_{reg}) and its counterpart in the reference frame (ROI_{ref}) as follows (40):

$$DSC = \frac{2(ROI_{ref} \cap ROI_{reg})}{ROI_{ref} + ROI_{reg}}$$

Each ROI consisted of the whole liver and was obtained by automatic segmentation using Medical Imaging Interaction Toolkit organ segmentation (41).

Quantitative analysis. The validity of the estimated motion was then tested by manually tracking blood vessel bifurcations in the liver. For each volunteer, vessel bifurcations were tracked in three zones: one in the upper part of the liver (lobes VII and VIII), one in the middle part of the liver (lobes II and III), and one in the lower part of the liver (lobes V and VI). The landmarks were placed twice by a radiologist to assess intrarater variation of manual positioning. The manually measured displacements were used to quantify the

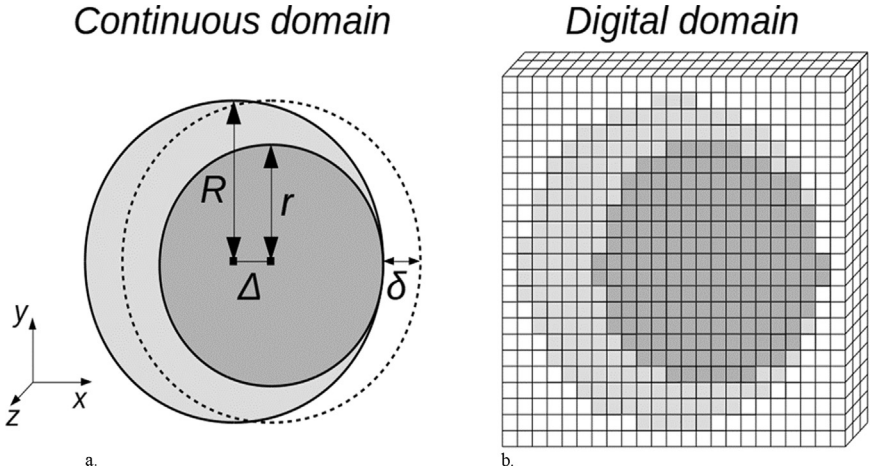

Figure E1. Assessment of the method employed for the determination of tumor surface area exposed to postablation margins $\leq$ 5 mm from a digital data set. **(a)** Proposed simulated case scenario in the continuous domain. **(b)** Corresponding digital data provided by pre- and postablation scans. The tumor present in the preablation scan is simulated by using the smaller sphere represented in dark gray. The postablation margin is simulated by a larger sphere represented in light gray.

positional variations caused by different patient positioning/imaging calibrations between the two scan sessions. In addition, to assess the performance of the motion estimation process, the manually tracked landmarks were compared versus the displacements provided by the employed image-registration algorithm. This was achieved by comparing the average amplitude of the displacement vectors obtained through manual tracking (corresponding to the positional errors we would have in the absence of the motion-compensation strategy) with the difference between these displacements and the displacements estimated by the employed image registration algorithm (which corresponds to the residual positional errors when the motion-compensation strategy was employed).

Assessment of Tumor and Ablative Area Segmentation Process

Segmentation of the tumor and ablative area requires a manual intervention by the radiologist to choose the ROI where the semiautomatic segmentation provided by the ITK-SNAP software will be performed. To assess the repeatability of the segmentation process, the average DSC values were calculated between two semiautomatic segmentations performed by two radiologists (A.H. and P.B.) with the same volume for each patient.

Assessment of Tumor Surface Area Exposed to Insufficient Postablation Margins $\leq$ 5 mm

The proposed case scenario is summarized in **Figure E1**. The tumor in the preablation image was simulated by using a sphere (radius r; light gray in **Fig E1a**) embedded in a larger sphere representing the postablation lesion (rayon R; dark gray in **Fig E1a**). The sphere representing the tumor was positioned so that the two spheres had only one single geometric point in common. We denote by Δ the Euclidean distance between the center of the two spheres (with $\Delta = R - r$) and by δ the maximal allowed tumor exposition margin ($\delta = 5$ mm for this study). That way, the right part of the sphere representing the tumor is exposed to the radiofrequency ablation margin equal to or less than δ (tumor margin in **Fig E1a**, delimited by the dashed circle, which exceeds the postablation lesion in the right part of the scheme). Assuming, for mathematical simplification, that the tumor exposition margin δ is equal to the difference between the two rayons/radiuses R and r (ie, $\delta = \Delta$), the analytic tumor exposition surface S can be obtained by using the following equation:

$$s = 2\,\pi\,r\left(r + \frac{\Delta\,r}{2\,R}\right)$$

To support this expression, one can observe that S corresponds to the surface of a spherical cap of height h with:

$$h = \left(r + \frac{\Delta\,r}{2\,R}\right)$$

The percentage of variation between the estimated and the analytic exposed surface area was calculated for spherical tumor diameters of 20, 30, and 40 mm. Various quantifications of the continuous scenarios were achieved (**Fig 1b**) to produce digital data sets with isotropic voxel sizes varying from 1 to 4 mm (at increments of 0.05 mm).