

research article

Safety margin assessment after microwave ablation of liver tumors: inter- and intrareader variability

Jan Schaible, Benedikt Pregler, Wolf Bäuml, Ingo Einspieler, Ernst-Michael Jung, Christian Stroszczynski, Lukas Philipp Beyer

Department of Radiology, University Medical Center Regensburg, Germany

Radiol Oncol 2020; 54(1): 57-61.

Received 16 September 2019
Accepted 10 January 2020

Correspondence to: Jan Schaible, Department of Radiology, University Hospital Regensburg, 93053 Regensburg, Germany.
Phone: +49 941 944 7401; E-mail: jan.schaible@ukr.de

Disclosure: No potential conflicts of interest were disclosed.

Background. The aim of the study was to evaluate the inter- and intrareader variability of the safety margin assessment after microwave ablation of liver tumors using post-procedure computed tomography (CT) images as well as to determine the sensitivity and specificity of identification remnant tumor tissue.

Patients and methods. A retrospective analysis of 58 patients who underwent microwave ablation (MWA) of primary or secondary liver malignancies (46 hepatocellular carcinoma, 9 metastases of a colorectal cancer and 3 metastases of pancreatic cancer) between September 2017 and June 2019 was conducted. Three readers estimated the minimal safety margin in millimeters using side-by-side comparison of the 1-day pre-ablation CT and 1-day post-ablation CT and judged whether ablation was complete or incomplete. One reader estimated the safety margin again after 6 weeks. Magnetic resonance imaging (MRI) after 6 weeks was the gold standard.

Results. The intraclass correlation coefficient (ICC) for estimation of the minimal safety margin of all three readers was 0.357 (95%-confidence interval 0.194–0.522). The ICC for repeated assessment (reader 1) was 0.774 (95%-confidence interval 0.645–0.860). Sensitivity and specificity of the detection of complete tumor ablation, defined as no remnant tumor tissue in 6 weeks follow-up MRI, were 93%/82%/82% and 33%/17%/83%, respectively.

Conclusions. In clinical practice, the safety margin after liver tumor ablation is often assessed using side-by-side comparison of CT images. In the study, we were able to show, that this technique has a poor reliability (ICC 0.357). From our point of view, this proves the necessity of new technical procedures for the assessment of the safety distance.

Key words: radiology, interventional; liver neoplasms; safety margin; interindividual variability; intraindividual variability; tumor ablation

Introduction

Thermal ablation methods, such as microwave ablation (MWA), have established themselves in recent years as a suitable therapy for various malignancies.¹ The execution of a tumor ablation involves many challenges. One crucial factor for a successful ablation and the prevention of residual tumor tissue or the onset of a tumor recurrence is maintaining a sufficient safety margin. There is currently no clear consensus on what a sufficient

safety distance is.² Most authors recommend a safety margin between 5-10 mm.^{1,3-7} The precondition for the determination of a suitable safety distance, however, would initially be a proper measurement method. So far, this has been proved another major challenge. Several studies have investigated postinterventional methods and measurement techniques, such as computed tomography (CT) or magnetic resonance imaging (MRI), to be able to make a valid decision about a complete ablation.⁸⁻¹⁰ Other authors try to improve the intraprocedural

tumor detection and the assessment of the ablation margin as a recently published study used a FDG PET/CT guided ablation for the intraprocedural determination of the safety distance and achieved good results.¹¹ Many authors favour the MRI for ablation margin control.^{10,12} However, in most cases tumor ablation is performed under CT guidance and the safety margin is assessed in the CT images, at least in the periinterventional setting. For best treatment, the decision whether ablation is complete or not should be made as immediately as possible. Therefore, in most cases, native or contrast-enhanced CT scans are performed, and the extent of the ablation is decided by side-by-side comparison of the pre- and post-interventional images or by simple and fast measuring techniques during the intervention like measurement with a simple distance measurement tool. Unfortunately, we do not have reliable data on the consistency and reproducibility of these subjective estimations of a sufficient safety distance in a real-world setting. For this reason, in this study we investigated the inter- and intrareader variability of the safety margin assessment after microwave ablation of liver malignancies.

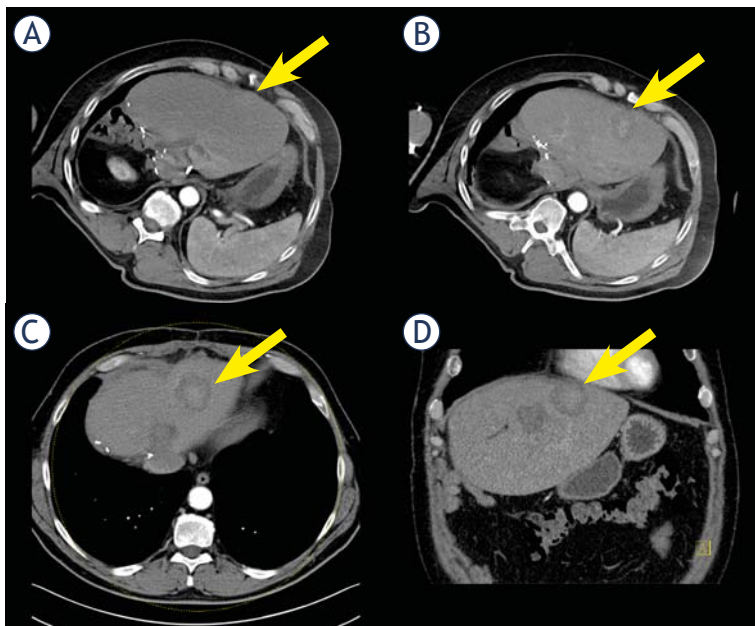


FIGURE 1. (A) Pre-interventional arterial phase in which the tumor is almost invisible. This not only complicates ablation but also post-interventional detection of residual tumor tissue. (B) The result immediately postinterventionally with a corresponding ablation defect. (C) Axial and (D) coronal show the situation one day postinterventionally. Due to the different breathing position, the tumor was more peripheral the day before, whereas on the following day healthy liver tissue around the ablation defect is visible. Measuring the safety distance is particularly difficult in these cases.

Patients and methods

Study design and participant selection

The local ethics committee approved this retrospective study. Written informed consent was obtained from all patients. A total of 58 patients were included in this study, who were treated with microwave ablation between September 2017 and June 2019. Tumor entities were hepatocellular carcinoma (HCC) and metastases of colorectal and pancreatic cancer. Exclusion criteria were the patient's refusal to participate in the study and other tumors than those mentioned above. All patients received a CT-scan one day before ablation and on the first postinterventional day (Figure 1). Subsequently, all patients were independently assessed by three interventional radiologists regarding the safety margin between tumor and healthy liver tissue using side-by-side measurement. No special evaluation software was used to simulate the procedure in everyday practice as accurately as possible. The orientation was based on reference points that could be reproduced exactly, *e.g.* prominent vessel outlets, foreign material or calcifications. Of course, the different breathing positions had to be taken into consideration as well. One of the three readers re-evaluated the patients after six weeks to detect possible intraindividual variability. The six weeks period of time between the two readings should ensure, that the reader could not remember the patients and avoid a bias. The minimum safety distance, the maximum safety distance and whether the ablation was considered complete or incomplete, *i.e.* technical efficacy, were estimated. The 6 weeks follow-up MRI was regarded as the gold standard for technical efficacy.

Statistics

Intraclass correlation (ICC) estimates and their 95% confident intervals were calculated using R irr statistical package version 3.5.1 based on a mean-rating ($k = 3$), absolute-agreement, 2-way mixed-effects model. The intraclass correlation coefficient (ICC) was calculated for the estimation of the minimal safety margin. ICC values less than 0.5 are considered indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability. Bland-Altman analyses were used to assess agreement in the side-by-side measurements between the two blinded readings (minimal safety margin) by the same radiologist and between the

TABLE 1. Baseline characteristics

Number of patients	N = 58
Age	
mean (years)	62.84 (10.85)
range (years)	36–83
Gender	
male (%)	53 (91)
Ablation method (%)	
microwave ablation	58 (100)
Liver segments	
I	3 (5)
II	4 (7)
III	6 (10)
IVa	7 (12)
IVb	3 (5)
V	7 (12)
VI	4 (7)
VII	12 (21)
VIII	9 (16)
Tumor entity	
Hepatocellular carcinoma	46 (79)
Metastasis colorectal cancer	9 (16)
Metastasis pancreatic cancer	3 (5)

readings (minimal safety margin) by the three independent radiologists.

Results

Patient and tumor characteristics

58 patients were included and evaluated. The mean age was 62.84 (10.85) years. 53 patients (91%) were male. All 58 patients were treated with MWA. Most tumors ($n = 12$) were located in liver segment VII, followed by segments VIII ($n = 9$) and IV a and V with $n = 7$ each. The minority of tumors were found in segments I and IV b with $n = 3$ each. The baseline data are shown in Table 1.

Inter- and intrareader variability

The intraclass correlation coefficient (ICC) for estimation of the interindividual variability of the assessment of the minimal safety margin for all three readers was 0.357 (95%-confidence interval 0.194–0.522), indicating a poor reliability. The ICC for estimation of the variability of two repeated estimations of reader 1 was 0.774 (95%-confidence interval 0.645–0.860), indicating a good reliability for repeated measurements.

Bland–Altman plots were calculated to show intra- and interindividual variability (Figure 2). A systematic error was not detectable. The standard deviation in the intrareader-result was smaller compared to the interindividual evaluations. Nevertheless, deviations of more than 5 mm can be detected in

TABLE 2. Contingency table of all the three independent readings compared with the six weeks follow-up MRI as gold standard

	Incomplete (6 weeks MRI)	Complete (6 weeks MRI)
Reader 1		
Incomplete	2	3
Complete	4	41
Reader 2		
Incomplete	1	8
Complete	5	36
Reader 3		
Incomplete	5	8
Complete	1	36

TABLE 3. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of the three independent readings (R 1, 2, 3) compared with the six weeks follow-up MRI as gold standard

	R 1	R 2	R 3
Sensitivity	93%	82 %	82 %
Specificity	33%	17 %	83 %
PPV	91 %	88 %	97 %
NPV	40 %	10 %	39%

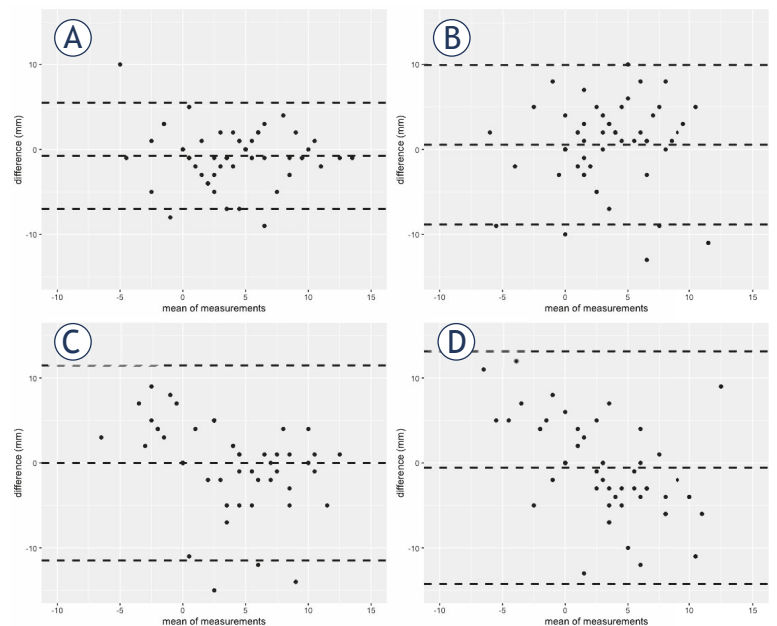


FIGURE 2. Bland–Altman plots: intra- (A) and inter-reader (B) = reader 1 vs. reader 2, (C) = reader 2 vs. reader 3, (D) = reader 1 vs. 3) agreement for minimum safety margin measurements. The middle line shows the mean percentage difference in measurements and the dashed lines above and below show the 95% reference range. Measurements within the 95% reference range can be considered as intrinsic measurement errors (or variations) that are associated with the given measurement tools and imaging techniques. Therefore, a narrower reference range indicates a lower measurement error/variation.

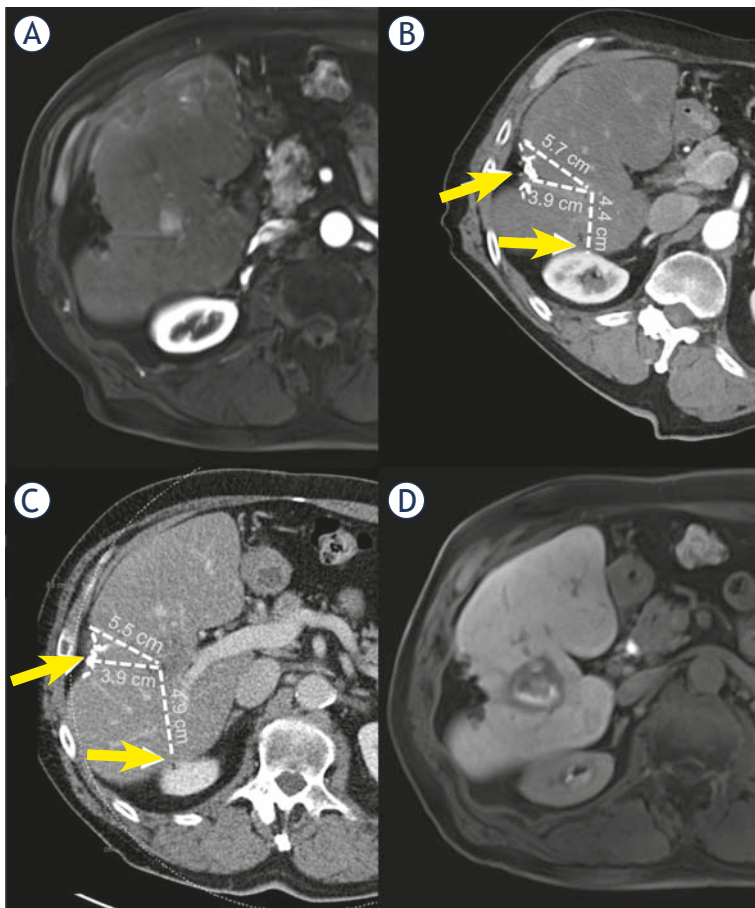


FIGURE 3: The HCC in segment VIII (A) was pre-interventionally localized using landmarks such as metal clips (arrow) and anatomical landmarks like the kidney (arrow) (B). Post-interventionally, the same landmarks are used and the target area is localized by distance measurements (dashed lines) from different angles (C). The MRI follow-up after 6 weeks confirmed complete ablation (D).

some measurements. The differences of the safety margins measured by the two readers are clearly larger in comparison to the deviations between both measurements performed by one reader.

Assessment of complete ablation

The readers achieved a sensitivity and specificity of 93%/82%/82% and 33%/17%/83%, respectively. The positive predictive value (PPV) was 91%/88%/97%. The negative predictive value (NPV) was 40%/10%/39%. The results are shown in Table 2 and 3.

Discussion

There is agreement that a safety distance is necessary after ablation of a liver tumor to prevent local

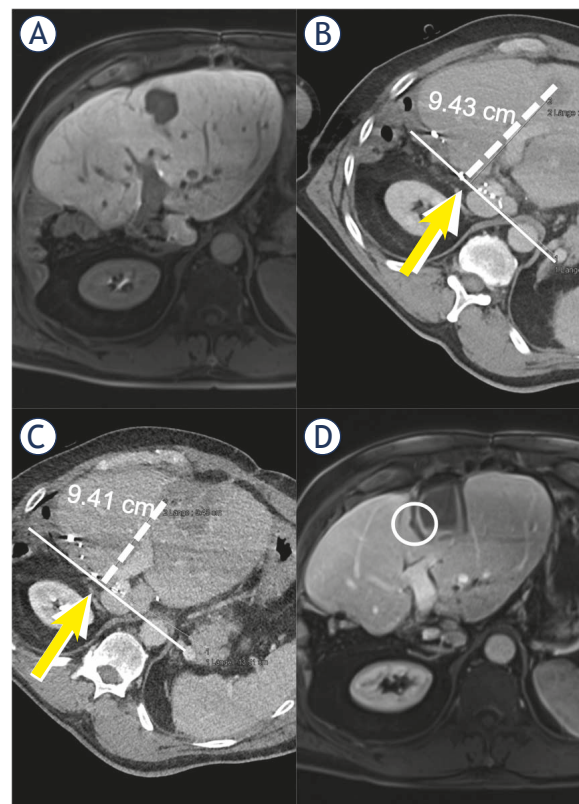


FIGURE 4: Metastasis was best seen in the portal venous phase (B). In this case, clips after hemihepatectomy serve as orientation. A line is drawn (solid line) and the distance (dashed line) is measured by means of a clip at an angle of 90 degrees. The same fixed points are used postinterventionally. This already shows only a small safety distance in the peripheral area. In the 6 weeks follow-up MRI residual tumor tissue (circle) was detected.

tumor recurrence. When defining the optimal safety distance, there are already different approaches and no general definition. Most authors favour a minimum distance of 5 mm (1,3–7).

We agree with this in principle. In our opinion, however, the measurement methods are rarely described or questioned. Therefore, our approach was to question the measurement of the safety distance in the daily routine (Figure 3 and 4).

This confirmed our impression that measurement with the standard tools provided by the CT software can lead to difficulties in measurement and thus to considerable intraindividual differences. Although the reading was performed by three experienced interventional radiologists, the inter-reader variability was poor.

One reason could be the localization of the tumor. Subcapsular tumors represent a special measuring challenge. The same applies for tumors in the

immediate vicinity of other organs or vessels that are also difficult to measure.

Another aspect that can lead to considerable differences in the evaluation of the distance is the choice of the reconstruction planes and the layer thickness. Zhao *et al.* claims to achieve best results with 1D or 3D 2.5 mm slices compared to 2D. From our point of view, an evaluation of the ablation zone in three planes is absolutely but also leads to a higher interreader variability.

A contentious aspect is always the experience of the interventionalist. Therefore, in our study the reading was carried out by an experienced radiologist (5 years experience), a specialist radiologist (7 years experience) and the head of the Centre for Interventional Oncological Radiology. The aim was to rule out diagnostic errors due to inexperience. Nevertheless, there were considerable differences between all three readers, which called the measuring method into question.

In our opinion, the fact that the intraindividual differences were smaller shows that there is no systematic error. The measurement results are interindividually different but not random. In our opinion, this indicates that our study results are reliable and meaningful.

New measurement methods or software for tumor segmentation are already being investigated in some studies.^{5,8,9,12-16} The results were promising and improved the assessment of ablation success. Our study was able to show that conventional measurement methods are inaccurate and can lead to large interindividual differences. We therefore support the development of new measurement methods to achieve more reliable measurement results.

References

- Vogl TJ, Nour-Eldin NA, Hammerstingl RM, Panahi B, Naguib NNN. Microwave ablation (MWA): basics, technique and results in primary and metastatic liver neoplasms – review article. *Rofo* 2017; **189**: 1055-66. doi: 10.1055/s-0043-117410
- Ahmed M, Technology Assessment Committee of the Society of Interventional Radiology. Image-guided tumor ablation: standardization of terminology and reporting criteria - a 10-year update: supplement to the consensus document. *J Vasc Interv Radiol* 2014; **25**: 1706-8. doi: 10.1016/j.jvir.2014.09.005
- Teng W, Liu KW, Lin CC, Jeng WJ, Chen WT, Sheen IS, et al. Insufficient ablative margin determined by early computed tomography may predict the recurrence of hepatocellular carcinoma after radiofrequency ablation. *Liver Cancer* 2015; **4**: 26-38. doi: 10.1159/000343877
- Ke S, Ding XM, Qian XJ, Zhou YM, Cao BX, Gao K, et al. Radiofrequency ablation of hepatocellular carcinoma sized > 3 and ≤ 5 cm: is ablative margin of more than 1 cm justified? *World J Gastroenterol* 2013; **19**: 7389-98. doi: 10.3748/wjg.v19.i42.7389
- Shin S, Lee JM, Kim KW, Joo I, Han JK, Choi BI, 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**: 70-7. doi: 10.2214/AJR.13.11709
- Wang X, Sofocleous CT, Erinjeri JP, Petre EN, Gonen M, Do KG, et al. Margin size is an independent predictor of local tumor progression after ablation of colon cancer liver metastases. *Cardiovasc Intervent Radiol* 2013; **36**: 166-75. doi: 10.1007/s00270-012-0377-1
- Kurilova I, Gonzalez-Aguirre A, Beets-Tan RG, Erinjeri J, Petre EN, Gonen M, et al. Microwave ablation in the management of colorectal cancer pulmonary metastases. *Cardiovasc Intervent Radiol* 2018; **41**: 1530-44. doi: 10.1007/s00270-018-2000-6
- Hocquet A, Trillaud H, Frulio N, Papadopoulos P, Balageas P, Salut C, et al. Three-dimensional measurement of hepatocellular carcinoma ablation zones and margins for predicting local tumor progression. *J Vasc Interv Radiol* 2016; **27**: 1038-45.e2. doi: 10.1016/j.jvir.2016.02.031
- Tani S, Tatli S, Hata N, Garcia-Rojas X, Olubiyi OI, Silverman SG, et al. Three-dimensional quantitative assessment of ablation margins based on registration of pre- and post-procedural MRI and distance map. *Int J Comput Assist Radiol Surg* 2016; **11**: 1133-42. doi: 10.1007/s11548-016-1398-z
- Kim SM, Shin SS, Lee BC, Kim JW, Heo SH, Lim HS, et al. Imaging evaluation of ablative margin and index tumor immediately after radiofrequency ablation for hepatocellular carcinoma: comparison between multidetector-row CT and MR imaging. *Abdom Radiol* 2017; **42**: 2527-37. doi: 10.1007/s00261-017-1146-z
- Shyn PB, Casadaban LC, Sainani NI, Sadow CA, Bunch PM, Levesque VM, et al. Intraprocedural ablation margin assessment by using ammonia perfusion PET during FDG PET/CT-guided liver tumor ablation: a pilot study. *Radiology* 2018; **288**: 138-45. doi: 10.1148/radiol.2018172108
- Koda M, Tokunaga S, Okamoto T, Hodozuka M, Miyoshi K, Kishina M, et al. Clinical usefulness of the ablative margin assessed by magnetic resonance imaging with Gd-EOB-DTPA for radiofrequency ablation of hepatocellular carcinoma. *J Hepatol* 2015; **63**: 1360-7. doi: 10.1016/j.jhep.2015.07.023
- Shyn PB, Casadaban LC, Sainani NI, Sadow CA, Bunch PM, Levesque VM, et al. Intraprocedural ablation margin assessment by using ammonia perfusion PET during FDG PET/CT-guided liver tumor ablation: a pilot study. *Radiology* 2018; **288**: 138-45. doi: 10.1148/radiol.2018172108
- Luu HM, Niessen W, van Walsum T, Klink C, Moelker A, Klink C, et al. An automatic registration method for pre- and post-interventional CT images for assessing treatment success in liver RFA treatment. *Med Phys* 2015; **42**: 5559-67. doi: 10.1118/1.4927790
- Solbiati M, Muglia R, Goldberg SN, Ierace T, Rotilio A, Passera KM, et al. A novel software platform for volumetric assessment of ablation completeness. *Int J Hyperthermia* 2019; **36**: 337-43. doi: 10.1080/02656736.2019.1569267
- Sakakibara M, Ohkawa K, Katayama K, Imanaka K, Ishihara A, Hasegawa N, 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**: W487-95. doi: 10.2214/AJR.13.11384