Can accurate evaluation of the treatment success after radiofrequency ablation of liver tumors be achieved by visual inspection alone? Results of a blinded assessment with 38 interventional oncologists

Gregor Laimer , Peter Schullian , Daniel Putzer , Gernot Eberle , S. Nahum Goldberg & Reto Bale

Published online: 11 Dec 2020.

Submit your article to this journal ☑

Article views: 32

View related articles ☑

View Crossmark data ☑

Taylor & Francis
Taylor & Francis Group

Can accurate evaluation of the treatment success after radiofrequency ablation of liver tumors be achieved by visual inspection alone? Results of a blinded assessment with 38 interventional oncologists

Gregor Laimer[a], Peter Schullian[a], Daniel Putzer[a], Gernot Eberle[a], S. Nahum Goldberg[b,c] and Reto Bale[a]

[a]Interventional Oncology-Microinvasive Therapy (SIP), Department of Radiology, Medical University Innsbruck, Innsbruck, Austria; [b]Department of Radiology, Hadassah Hebrew University Medical Centre, Jerusalem, Israel; [c]Department of Radiology, Beth Israel Deaconess Medical Center, Boston, MA, USA

ABSTRACT

Objectives: To assess the difficulties in the immediate judgment of treatment success after radiofrequency ablation (RFA) of liver tumors by visual inspection alone and to evaluate whether radiologist's expertise affects the resultant judgment.

Methods: Peri-interventional CT-scans of nine patients with nine hepatocellular carcinomas with known outcomes after RFA were presented to 38 participants from 14 different countries. In a total of 342 reads, all interventional oncologists assessed the pre- and immediate post-interventional CT-scans through conventional side-by-side juxtapositioning of images and judged whether complete ablation (i.e., technical success and technique efficacy) was achieved. Results were compared regarding expertise in percutaneous tumor ablation (>50 interventions performed). An 'overcall' was defined as insufficient ablation that was misjudged as sufficient, and an 'undercall' as an erroneous assessment of complete ablation.

Results: Overall 3.97 ± 1.27 out of 9 (44.1%) cases per radiologist were misjudged. The mean number of overcalls and undercalls per radiologist were 0.74 ± 0.50 out of 2 (37.0%), and 3.24 ± 1.28 out of 7 (46.3%), respectively. 18/38 (47.4%) participants had considerable experience in percutaneous tumor ablation, with such expertise having no significant influence on the results (overall: $p = 0.70$; overcalls: $p = 0.87$; undercalls: $p = 0.75$).

Conclusions: Conventional side-by-side evaluation of treatment success after RFA of liver tumors by the juxtaposition of pre- and post-interventional CT-scans is very difficult for experienced radiologists. The implementation of advanced processing techniques such as rigid/non-rigid image fusion with the assessment of the periablational margin is thus likely needed in order to decrease errors and objectively evaluate technical success and predict technique efficacy of liver RFA.

Introduction

Radiofrequency ablation (RFA) is a minimally invasive, potentially curative local treatment option of liver malignancies and has emerged as a first-line approach for patients with small hepatocellular carcinoma (HCC) [1–3]. A crucial point in curative treatments of HCC is the evaluation of treatment success. Increasing evidence has pointed toward the periablational margin as a significant determinant of RFA success, associating a periablational margin of less than 5 mm with higher rates of local tumor progression (LTP) [4–9].

A widely used technique to assess the technical success of RFA is the comparison of pre- and post-interventional CT-scans in side-by-side juxtaposition. However, this method contains many possible sources of error and can be very challenging even for experienced radiologists. Indeed, the exact extent of variability and accuracy of this conventional method is currently unknown. Thus, the aim of this analysis

was to assess the difficulties in the immediate judgment of technical success and prediction of technique efficacy after RFA without image fusion and evaluation of the periablational margin, independent of the radiologist's expertise in percutaneous tumor ablation.

Material and methods

As part of the ESIR Course 'Reliability in Percutaneous Ablation', hosted at the Department of Radiology of the Medical University of Innsbruck (12–13th December 2019), a substantial population of interventional oncologists were asked to evaluate multiple RFA cases on a voluntary basis in order to assess their visual acuity and their ability to determine ablation success. Peri-interventional CT-scans of nine patients with nine HCCs treated with RFA at our institution were shown to the participants in small groups. Through the

---

CONTACT Reto Bale reto.bale@i-med.ac.at Interventional Oncology-Microinvasive Therapy, Department of Radiology, Medical University Innsbruck, Anichstraße 35, Innsbruck 6020, Austria

conventional side-by-side juxtaposition of pre- and post-interventional CT-scans without evaluation of the periablational margin, the participants had to decide whether the RFA treatment could be considered successful or not. Successful ablation was defined by the combination of technical success and technique efficacy.

Stereotactic radiofrequency ablation

All cases included in this study were treated with stereotactic radiofrequency ablation (SRFA) at a single institution (Interventional Oncology-Microinvasive Therapy (SIP), Department of Radiology, Medical University Innsbruck, Austria). SRFA includes three-dimensional planning of multiple overlapping ablation zones, precise stereotactic placement of multiple coaxial needles, and intraoperative assessment of the resultant ablation by means of image fusion. Full procedural details of SRFA have been described briefly [10] and in detail [11] elsewhere. With SRFA, the spectrum of locally curable liver lesions can be dramatically increased due to the creation of overlapping ablation zones using multiple ablation applicators [10,11]. Indeed, a recent study in explanted livers [12] reported a complete histopathological response in 183 of 188 nodules (97.3%) and 50 of 52 nodules >3 cm (96.2%).

Ablation assessment quiz

At first, a rigid (i.e., Syngo.via, Siemens Healthineers, Erlangen, Germany) and a non-rigid registration software (i.e.,

Ablation-fit, R.A.W. Srl, Milano, Italy) for the evaluation of the periablational margin were presented to the participants as part of the course. Both software showed promising results in recently published studies [9,13]. Thereafter, the participants attended on a voluntary basis the so-called 'ablation assessment quiz'. All attendees were asked to individually complete a form with the following parameters: country of their institution, number of years actively practicing interventional radiology, the estimated number of liver RFA/microwave ablations (MWA) (less than 10; more than 10; more than 50), image guidance used at their institution (US; CT; Cone beam-CT; MRI), and whether image fusion was used in daily clinical practice. Subsequently, using a diagnostic image viewing program (i.e., IMPAX Agfa HealthCare, Mortsel, Belgium), pre- and post-interventional CT-scans in side-by-side juxtaposition of nine patients with nine target tumors and known oncological outcome treated at our institution with SRFA were consecutively presented to the participants. The pre- and post-interventional CT-scans shown to the participants were obtained under general anesthesia as part of the SRFA procedure using the planning and control CT-scans of each session. They consisted of a dual-phase contrast-enhanced CT-scan (Siemens SOMATOM Sensation Open, sliding gantry with 82 cm diameter, Siemens AG, Erlangen, Germany) with 3 mm slice; 35–40 s and 70–80 s after initiation of contrast material injection (100–150 ml of Iopromide [Ultravist 370; Schering AG, Berlin, Germany]), representing the late arterial and late portal venous phases.

Table 1 shows patient/tumor characteristics and an exemplary case is shown in Figure 1. The participants had to

Table 1. Patient/tumor characteristics of nine HCCs presented to the participants of the 'ablation assessment quiz'.

No.	Age (years)	Gender	Tumor size (mm)	Tumor location	LTP	Time to LTP (months)	Follow-Up (months)
1	63	Male	34	Extrahepatic organ	Yes	19.4	
2	75	Male	29	None	No		19.1
3	59	Male	40	Subcapsular/extrahepatic organ/major vessel	Yes	6.1	
4	77	Male	25	Subcapsular/extrahepatic organ	No		45.6
5	49	Male	20	Major vessel	No		46.5
6	72	Male	16	Major vessel	No		20.3
7	74	Male	40	None	No		39.7
8	89	Male	60	None	No		24.9
9	79	Male	15	None	No		27.4

LTP: local tumor progression.
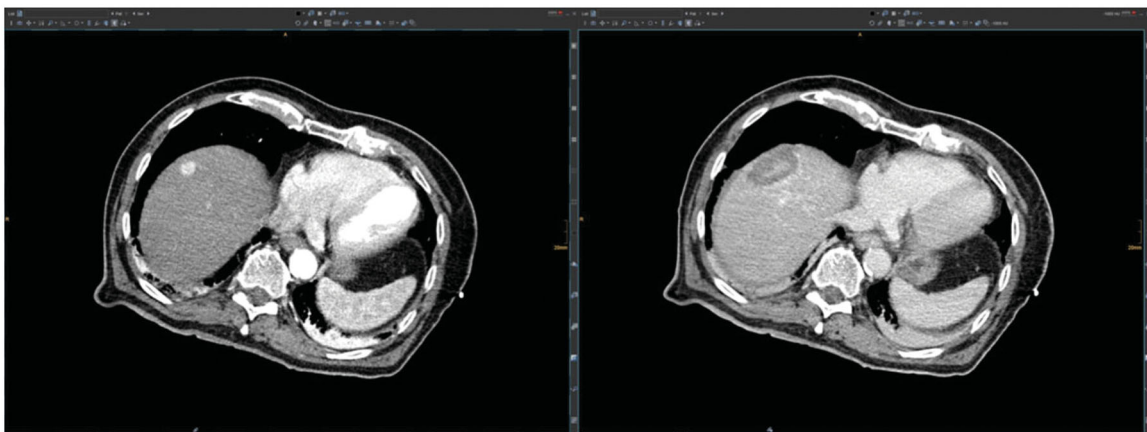
Figure 1. Evaluation of the technical success and technique efficacy after stereotactic radiofrequency ablation (SRFA) through conventional side-by-side juxtaposition of pre- and post-interventional CT-scans as shown to the participants of the so-called 'ablation assessment quiz'.
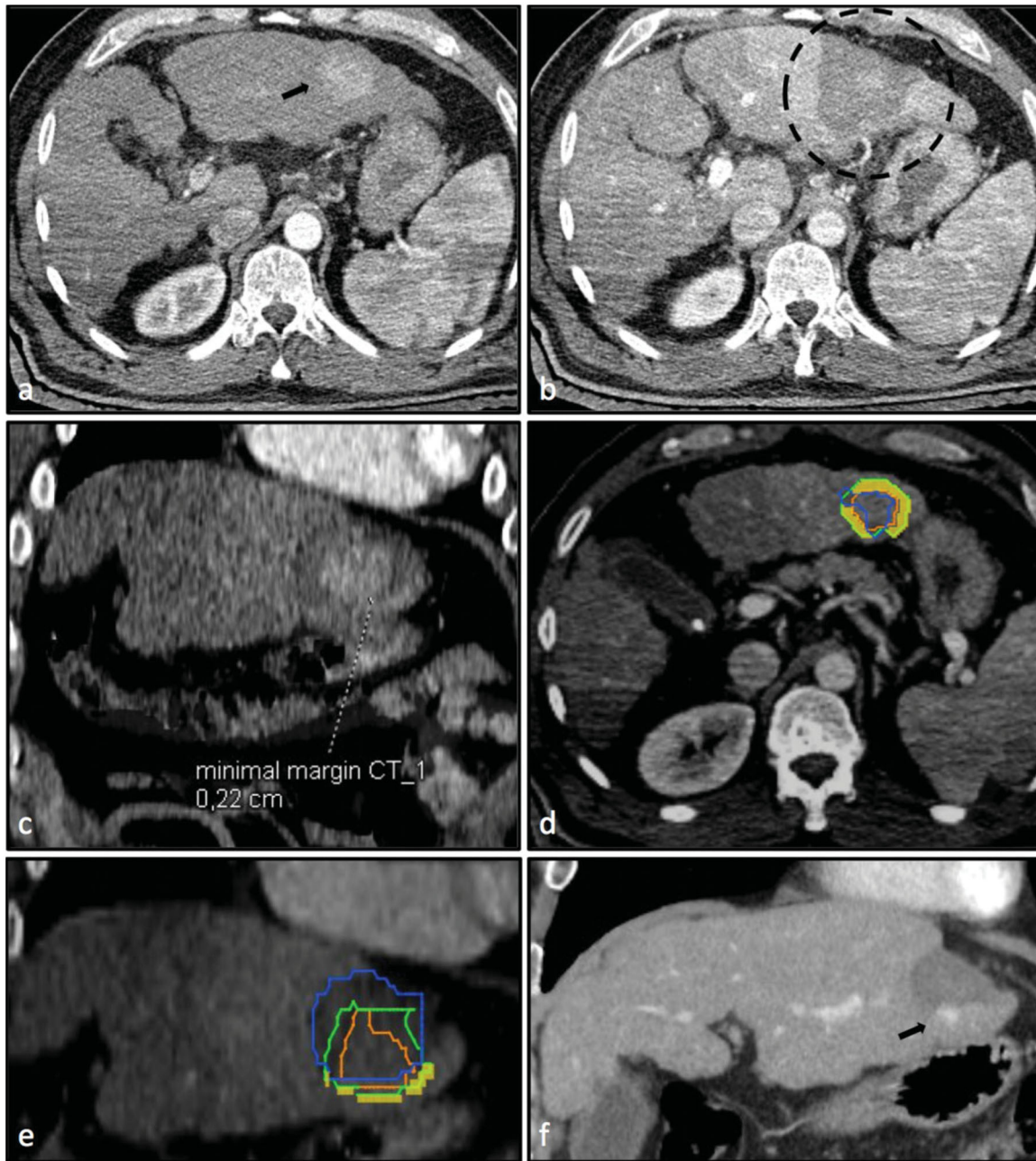
Figure 2. Case 1 of the so-called 'ablation assessment quiz': 63-year old male with a 34 mm HCC. Axial plane of pre-interventional CT-scan (a) with HCC in liver segment III (black arrow) and post-interventional CT-scan (b) with large necrosis zone (black dashed circle). (c) Axial plane after image fusion of pre- and post-interventional CT-scans with a rigid registration tool (Syngo.via, Siemens Healthineers, Erlangen, Germany) revealing a MAM of 2 mm. Axial (d) and coronal (e) plane after image fusion of pre- and post-interventional CT-scans with a non-rigid registration tool (Ablation-fit, R.A.W. Srl, Milano, Italy) revealing a lack of sufficient MAM in caudal position (orange line = HCC; green line = HCC with 5 mm margin; blue = necrosis zone; yellow area = unablated safety margin). (f) LTP with 10 mm in diameter caudally to necrosis zone 19 months after the intervention (black arrow).

judge whether the ablation was successful or not, knowing that a periablational margin of >5 mm was sought to achieve in all ablations. Successful ablations were defined by the combination of technical success (i.e., ablation zone completely overlaps target tumor; no residual vital tumor tissue at first follow-up CT-scan) and technique efficacy (i.e., future absence of local tumor progression in subsequent follow-up CT-scans) [14]. A retrospective review by three experienced interventional radiologists documenting the absence of local tumor progression at a minimum of 18 months' follow-up was established as the reference standard ('ground truth') in consensus. All nine selected cases included target tumors

that were technically successfully treated as they showed no residual tumor tissue at the first follow-up CT-scan approximately one month after the intervention. Two target tumors (Case 1 and Case 3) developed local tumor progression as documented by subsequent follow-up CT-scans, resulting in an insufficient ablation (i.e., insufficient technique efficacy). Case 1 with periablational margin assessment is shown as an example in Figure 2.

If an insufficient ablation was judged by the participant as sufficient to achieve complete ablation, it was defined as an 'overcall'. On the other hand, if a sufficient ablation was judged to be insufficient, it was defined as an 'undercall'. At

the end of the ESIR Course, quiz results together with the evaluated periablational margin as determined by image fusion using both software [9,13] were presented.

Of the 40 participants, 38 gave written permission to a further analysis of their results.

Statistical analysis

All statistical analyses were performed using SPSS Version 22 (SPSS Inc., Chicago, Illinois).

Data are expressed as total numbers, mean, and range. The distribution (homogeneous/non-homogeneous) of all variables was assessed using histograms. Differences between categorical variables were evaluated with the χ^2 test, while differences between independent continuous variables were evaluated with the independent Student's t-test. A p-value <0.05 was considered statistically significant.

Results

Participants were from 14 different countries (Table 2). 18/38 (47.4%) showed considerable experience in percutaneous tumor ablation with more than 50 interventions performed at their institution. The mean number of years in interventional radiology of all participants was 8.1 (0–37), with the most experienced attendee having 37 years of experience in interventional radiology. Almost all attendees had experience with more than one image guidance, CT (31 [81.6%]) and Ultrasound (29 [76.3%]) being the most prevalent. Nine (23.7%) participants had been using image fusion for the evaluation of the technical success and prediction of technique efficacy at the time of the quiz. Table 3 summarizes the descriptive statistics of the attendees.

A total of 151 out of 342 case reads were misjudged with the mean number of misjudged cases being 3.97 (±1.27) out of 9 (44.1%) per radiologist. None of the participants assessed all cases correctly. Independent Student's t-test

revealed no significant influence of the expertise in percutaneous tumor ablation on the number of misjudged cases at the ablation assessment quiz (p = 0.70; Table 4). Radiologists with more than 50 liver RFA/MWAs performed (mean = 3.89; SD = 1.37; n = 18) showed nearly identical results as radiologists with less than 50 liver RFA/MWAs performed (mean = 4.05; SD = 1.19; n = 20). A further subdivision of the cases based on overcalling vs. undercalling also revealed no significant differences between the groups (overcalling: p = 0.87; undercalling: p = 0.75). Overall, the mean number of overcalls in insufficient ablations was 0.74 ± 0.50 out of 2 per radiologist and the mean number of undercalls in sufficient ablations was 3.24 ± 1.28 out of 7 per radiologist as illustrated in Table 4.

Subdivided per case, no significant differences between the two groups regarding the results were observed (Table 5).

Discussion

Our analysis clearly shows that accurate evaluation of treatment success after RFA through the conventional side-by-side juxtaposition of pre- and post-interventional CT-scans is very challenging. The mean number of misjudged cases in this 'ablation assessment quiz' was 3.97 (±1.27) out of 9, independent of the radiologist's expertise (p = 0.70). This high number of misjudged cases (in our study >40%) leads to a substantial number of overcalls with the result of residual vital tumor tissue or LTP. Also, a non-inconsequential number of undercalls should not be underestimated. In order to guarantee a sufficient ablation, an undercall can lead to additional, unnecessary secondary needle placement. This, in turn, increases the risk of treatment-associated complications.

The high number of misjudged cases in this study reinforces the assumption that an assessment of the periablational margin with solutions such as image fusion of pre- and post-interventional CT-scans is likely needed to overcome difficulties encountered with conventional side-by-side juxtaposition.

Several studies already confirmed that an ablation with a circumscribed periablational margin > 5 mm can be considered successful due to an extremely low probability of local tumor progression (LTP) [4,9,14]. A recently published study evaluating technique efficacy in HCC patients referred to SRFA showed a relative LTP risk reduction of 30% for each millimeter increase in the periablational margin [9].

Table 2. Country of the institution of 38 participants.

Country	DE	NL	UK	ES	US	BR	HU	PL	UA	IT	SE	MX	DK	PT
No. of participants	9	6	6	3	3	2	2	1	1	1	1	1	1	1

Table 3. Characteristics of 38 participants of the 'ablation assessment quiz' performed at the ESIR course 'reliability in percutaneous ablation' hosted at the Department of Radiology of the Medical University of Innsbruck (12–13th December 2019).

Participant characteristics	
Estimated number of liver RFA/MWAs performed, n (%)	
Less than 10	13 (34.2)
More than 10	7 (18.4)
More than 50	18 (47.4)
Image guidance used, n (%)	
CT	31 (81.6)
Ultrasound	29 (76.3)
Cone beam CT	6 (15.8)
MRI	3 (7.9)
Image fusion used at institution, n (%)	9 (23.7)
No. of years in interventional oncology, mean (range)	8.1 (0–37)

RFA: radiofrequency ablation; MWA: microwave ablation; CT: computed tomography; MRI: magnetic resonance imaging; No.: number.

Table 4. Misjudged cases per participant in the 'ablation assessment quiz': experienced (more than 50 liver RFA/MWAs) vs. unexperienced interventional radiologists (less than 50 liver RFA/MWAs).

	Overall $n = 38$	Experienced IO $n = 18$	Unexperienced IO $n = 20$	p-Value
No. of cases misjudged	3.97 ± 1.27 (44.1%)	3.89 ± 1.37 (43.2%)	4.05 ± 1.19 (45.0%)	0.700
No. of 'overcalls'	0.74 ± 0.50 (8.2%)	0.72 ± 0.46 (8.0%)	0.75 ± 0.55 (8.3%)	0.868
No. of 'undercalls'	3.24 ± 1.28 (36.0%)	3.17 ± 1.29 (35.2%)	3.30 ± 1.30 (36.7%)	0.754

IO: interventional oncologist.

Table 5. Misjudgments of participants in the 'ablation assessment quiz' subdivided per case and group: experienced (more than 50 liver RFA/MWAs) vs. unexperienced interventional oncologists (less than 50 liver RFA/MWAs).

	Successful ablation	Overall $n = 38$	Experienced IO $n = 18$	Unexperienced IO $n = 20$	p-Value
Misjudgments per case, n (%)					
Case 1	No	22 (57.9)	9/18 (50)	13/20 (65)	0.350
Case 2	Yes	17 (44.7)	6/18 (33.3)	11/20 (55)	0.180
Case 3	No	6 (15.8)	4/18 (22.2)	2/20 (10)	0.302
Case 4	Yes	3 (7.9)	0/18 (0)	3/20 (15)	0.087
Case 5	Yes	18 (47.4)	10/18 (55.6)	8/20 (40)	0.338
Case 6	Yes	24 (63.2)	11/18 (61.1)	13/20 (65)	0.804
Case 7	Yes	35 (92.1)	17/18 (94.4)	18/20 (90)	0.612
Case 8	Yes	19 (50)	8/18 (44.4)	11/20 (55)	0.516
Case 9	Yes	7 (18.4)	5/18 (27.8)	2/20 (10)	0.158

IO: interventional oncologist.

Nevertheless, in many institutions, neither an intraprocedural nor an immediate evaluation of the periablational margin after RFA is performed in daily clinical practice. Through image fusion of pre- and post-interventional CT-scans and immediate assessment of the periablational margin, a statement on technical success and technique efficacy after RFA can be given independently of the performing radiologist and facilitate further ablation in the same session or very promptly, if needed.

In several diagnostic image-viewing programs, rigid-registration tools are already implemented and could therefore be used for image-fusion of pre-and post-interventional CT-scan for evaluation of the periablational margin. On the other hand, non-rigid registration tools, where liver parenchyma can be deformed section by section, require performance using fully automatically by a specialized software because of their complex registration algorithms. Such a fully automatic software platform has been used recently in a retrospective study and showed very promising results [13].

In our analysis, one case (i.e., Case 7) was misjudged by 35/38 (92.1%) participants. In this case, a periablational margin of only 2 mm could be achieved. Nevertheless, this tumor did not develop LTP on follow-up CT-scan to 39.7 months, which may be attributed to the strong encapsulation of the target tumor. These findings underline once again the difficulty in objectively judging of the technical success and technique efficacy of RFA through a conventional side-by-side juxtaposition – even for experienced radiologists.

It must be noted that a circumscribed periablational margin >5 mm cannot always be achieved for several practical reasons. Electrode placement in conventional RFA may be difficult when targeting particularly large or hard to reach targets; this can be overcome by stereotactic needle placement in SRFA [10,15]. Vessel proximity on the other hand remains a limiting factor regardless of the approach, since the presence of a vessel restricts the periablational margin, often below the critical 5 mm. This lack of a sufficiently large periablational margin is further compounded by the heat sink effect of the vessel. In such cases, LTP can be prevented by increased duration and power of the ablation and by probes positioned preferentially next to vessel sites. Yet, if possible, a circumscribed periablational margin of >5 mm, assessed through image fusion of pre- and post-interventional CT-scans, has to be achieved to consider an ablation successful at the time of the intervention [4–9].

It is further acknowledged that the conclusions of this study may be limited by the relatively small sample size (38 interventional radiologists), particularly the small number of cases, and a bias in the selection of the cases. Nevertheless, additional studies are likely not warranted to draw our main conclusion particularly based upon the very high level of misjudgment revealed in this study. Thus, although a larger sample size may slightly change the results even a change of ± 10 percentage points of undercalling or overcalling will still be deemed clinically unsatisfactory by virtually all clinicians. Another limitation of the study is reflected by the fact that the small number of cases did not allow for an analysis of the predictive factors for miscalls. This topic should be analyzed in future studies with a higher number of cases. Furthermore, the heterogeneity of the cases regarding tumor size may have impacted the results as it is likely that many interventional oncologists tend to be more skeptical regarding technical success with larger tumors.

In conclusion, our findings have several important implications. First, we have established that the evaluation of technical success and technique efficacy after RFA through the conventional side-by-side juxtaposition of pre- and post-interventional CT-scans used in most institutions can be very difficult. Next, we note that even experienced radiologists with more than 50 percutaneous tumor ablations performed misjudge in many cases. Accordingly, the implementation of a rigid/non-rigid image fusion of pre- and post-interventional CT-scans with an assessment of the periablational margin can help to prevent any errors and objectively evaluate technical success after RFA and predict technique efficacy.

Disclosure statement

S. Nahum Goldberg performs unrelated consulting for Angiodynamics and Cosman Instruments. The other authors do not have any conflicts of interest to disclose.

ORCID

Reto Bale http://orcid.org/0000-0003-0021-5792

References

[1] Forner A, Llovet JM, Bruix J. Hepatocellular carcinoma. Lancet. 2012;379(9822):1245–1255.

[2] Facciorusso A, Di Maso M, Muscatiello N. Microwave ablation versus radiofrequency ablation for the treatment of hepatocellular carcinoma: a systematic review and meta-analysis. Int J Hyperthermia. 2016;32(3):339–344.

[3] EASL. EASL clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol. 2018;69:182–236.

[4] Nakazawa T, Kokubu S, Shibuya A, et al. Radiofrequency ablation of hepatocellular carcinoma: correlation between local tumor progression after ablation and ablative margin. Am J Roentgenol. 2007;188(2):480–488.

[5] Kaye EA, Cornelis FH, Petre EN, et al. Volumetric 3D assessment of ablation zones after thermal ablation of colorectal liver metastases to improve prediction of local tumor progression. Eur Radiol. 2019;29(5):2698–2705.

[6] Shady W, Petre EN, Do KG, et al. Percutaneous microwave versus radiofrequency ablation of colorectal liver metastases: ablation with clear margins (A0) provides the best local tumor control. J Vasc Interv Radiol. 2018;29(2):268–275.

[7] Wang X, Sofocleous CT, Erinjeri JP, et al. Margin size is an independent predictor of local tumor progression after ablation of colon cancer liver metastases. Cardiovasc Intervent Radiol. 2013; 36(1):166–175.

[8] Jiang C, Liu B, Chen S, et al. Safety margin after radiofrequency ablation of hepatocellular carcinoma: precise assessment with a three-dimensional reconstruction technique using CT imaging. Int J Hyperthermia. 2018;34(8):1135–1141.

[9] Laimer G, Schullian P, Jaschke N, et al. Minimal ablative margin (MAM) assessment with image fusion: an independent predictor for local tumor progression in hepatocellular carcinoma after stereotactic radiofrequency ablation. Eur Radiol. 2020;30(5): 2463–2472.

[10] Schullian P, Johnston EW, Putzer D, et al. Safety and efficacy of stereotactic radiofrequency ablation for very large (≥8 cm) primary and metastatic liver tumors. Sci Rep. 2020; 10(1):1618.

[11] Bale R, Widmann G, Haidu M. Stereotactic radiofrequency ablation. Cardiovasc Intervent Radiol. 2011;34(4):852–856.

[12] Bale R, Schullian P, Eberle G, et al. Stereotactic radiofrequency ablation of hepatocellular carcinoma: a histopathological study in explanted livers. Hepatology. 2019;70(3): 840–850.

[13] Solbiati M, Muglia R, Goldberg SN, et al. A novel software platform for volumetric assessment of ablation completeness. Int J Hyperthermia. 2019;36:337–343.

[14] Motoyama T, Ogasawara S, Chiba T, et al. Coronal reformatted CT images contribute to the precise evaluation of the radiofrequency ablative margin for hepatocellular carcinoma. Abdom Imaging. 2014;39(2):262–268.

[15] Schullian P, Johnston EW, Putzer D, et al. Stereotactic radiofrequency ablation of subcardiac hepatocellular carcinoma: a case-control study. Int J Hyperthermia. 2019;36(1):876–885.