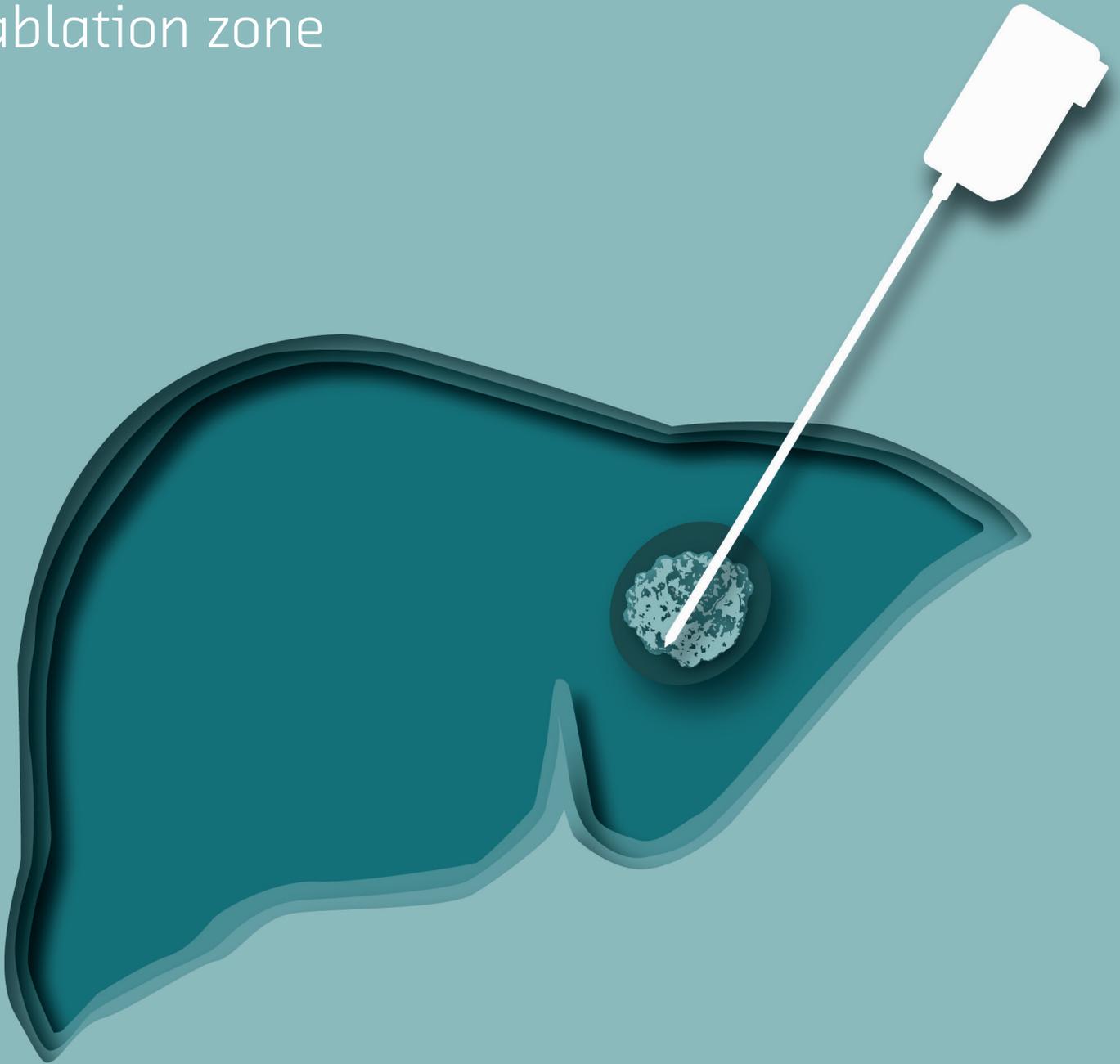


# Thermal ablation of primary liver tumors: Correspondence of the predicted ablation zone with the clinically obtained ablation zone



Gonnie van Erp - December 2022



# **THERMAL ABLATION OF PRIMARY LIVER TUMORS: CORRESPONDENCE OF THE PREDICTED ABLATION ZONE WITH THE CLINICALLY OBTAINED ABLATION ZONE**

G.C.M. (Gonnie) van Erp

Student number : 4535561

09-12-2022

Thesis in partial fulfilment of the requirements for the joint degree of Master of Science in

*Technical Medicine*

Leiden University ; Delft University of Technology ; Erasmus University Rotterdam

Master thesis project (TM30004; 35 ECTS)

Dept. of Biomechanical Engineering, TUDELFT

March 2022 – December 2022

Supervisor(s):

Dr. M.C. (Mark) Burgmans

Dr. ir. A. (Alexander) Broersen

P. (Pim) Hendriks, MSc

Thesis committee members:

Ass. Prof. Dr. ir. S. (Stefan) Klein, Erasmus MC (chair)

Dr. M.C. (Mark) Burgmans, LUMC

Dr. ir. A. (Alexander) Broersen, LUMC

P. (Pim) Hendriks, MSc, LUMC

Dr. R.W.C. (Roderick) Scherptong, LUMC

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.



## Preface and Acknowledgement

This thesis has been written to fulfill the graduation requirements of the master Technical medicine. Within this study program, my interests in technical innovations and medicine was perfectly combined. During my internships, I was able to develop myself and gain a lot of experience. I soon realized that it was not a specific department that enthusiasts me, but especially the possibilities as a Technical physician within each department. Afterwards, I needed to decide where I would finish my master. Even before making the final discission, I got a call from Pim to ask me if I would like to start a PhD trajectory, after graduating within the radiology department. At first I was a bit flabbergasted, but I have accepted this offer gratefully. I took a leap of faith, but have learned much and developed myself throughout the year. The internship resulted in this master thesis consisting of my thesis research, literature review and the published PROMETHEUS protocol, the prospective study I have been coordinating during my graduation internship.

I would like express my sincere gratitude to my supervisors, Mark, Alexander & Pim for their valuable insights and enthusiastic guidance during this project. I am looking forward to extend our collaboration in the upcoming years with you and the other members of the PROMETHEUS team, which I also would like to thank for their support during my internship. Special thanks to Alexander for implementing the needle reconstructions within deLIVERed, which facilitated this research.

Lastly, I would like to thank my family, colleagues and friends for their support throughout my thesis.

## List of Abbreviations

BCLC	Barcelona Clinic Liver Cancer
BCLC-0	Very early stage HCC according to the Barcelona Clinic Liver Cancer system
BCLC-A	Early stage HCC according to the Barcelona Clinic Liver Cancer system
BCLC-B	Advances stage HCC according to the Barcelona Clinic Liver Cancer system
CECT	Contrast-enhanced computed tomography
CRLM	Colorectal liver metastasis
CT	Computed tomography
CTHA	CT hepatic angiography
deLIVERed	de Leiden Interactive Visualization en Registration Editor
DICOM	Digital Imaging and Communications in Medicine
DFS	Disease-free survival
DSC	Dice similarity coefficient
Emprint HP	Emprint high power
HCC	Hepatocellular carcinoma
IQR	Interquartile range
LR	Local recurrence
LRR	Local recurrence rate
LTP	Local tumor progression
LUMC	Leiden University Medical Center
MAM	Minimal ablation margin
MWA	Microwave ablation
NASH	Nonalcoholic steatohepatitis
PACS	Picture Archiving and Communication System
OR	Odds ratio
OS	Overall survival
RFA	Radiofrequency ablation
RVD	Relative volume deviation
SD	Standard deviation

# Table of Contents

<b>Preface &amp; Acknowledgement</b> .....	I
<b>List of abbreviations</b> .....	II
<b>Table of contents</b> .....	III
<b>Thermal ablation of primary liver tumors: Correspondence of the predicted ablation zone with the clinically obtained ablation zone</b> .....	<b>1</b>
Abstract .....	1
1. Introduction.....	1
2. Methods .....	2
2.1 Patients.....	3
2.2 Procedure .....	4
2.3 Data .....	4
2.4 Needle position reconstruction .....	5
2.5 Statistical analysis .....	6
2.6 Peripheral ablation zones .....	7
2.7 Reconstruction error .....	7
3. Results .....	8
3.1 Patients.....	8
3.2 Ablation zone dimensions comparison .....	8
3.3 Ablation zones overlap comparison.....	10
3.4 Reconstruction error .....	10
4. Discussion .....	11
5. Conclusion .....	15
References.....	16
Appendix A: Supplementary results .....	19
<b>Computational modeling of thermal ablation zones in the liver: a systematic review</b> .....	<b>21</b>
Abstract .....	21
1. Introduction.....	22
2. Methods .....	22
2.1 Search strategy .....	22
2.2 Study selection .....	22
2.3 Data analysis.....	22
3. Results .....	22
3.1 Study selection .....	22
3.2 Data analysis.....	22

4. Discussion .....	39
5. Conclusion .....	41
References.....	42
Appendix A: Search Strategies.....	47
<b>Study Protocol PROMETHEUS: PROspective Multicenter study to Evaluate the correlation between safety margin and local recurrence after THERmal ablation USing image co-registration in patients with hepatocellular carcinoma .....</b>	<b>48</b>
Abstract .....	48
1. Introduction.....	49
2. Materials and Methods .....	49
2.1 Trial design and study setting.....	49
2.2 Participants.....	49
2.3 Interventions.....	49
2.4 Follow-up .....	50
2.5 MAM analysis .....	50
2.6 Outcomes .....	51
2.7 Statistical methods .....	51
3. Discussion .....	52
References.....	54
<b>General discussion.....</b>	<b>55</b>
References.....	57

# Thermal ablation of primary liver tumors: Correspondence of the predicted ablation zone with the clinically obtained ablation zone

## Abstract

**Purpose** Manufacturer's predictions of ablation zone dimensions are the current directives for treatment planning in thermal ablation, while they are mostly based on ex vivo experiments making its reliability questionable. The aim of this study is to determine the correspondence in dimensions, volume, shape and overlap of the manufactures' predicted with the clinically realized ablation zones following thermal ablation of hepatocellular carcinoma (HCC). The secondary objective is to determine the effect of tumor- and liver characteristics on this correspondence.

**Methods** Data was retrospectively collected from two prospective studies. A registered pre-ablation and post-ablation computed tomography scan with liver, tumor and ablation zone segmentations were available for analysis. Needle position reconstruction was performed based on image visual assessment (e.g., gas formation, tumor location and subcapsular hemorrhage) using in-house developed software. The dimensions of the predicted ablation zone were derived from the manufacturer's chart corresponding to treatment settings used during ablation. The long axis diameter (LAD), short axis diameter (SAD) and volume of the realized and predicted ablation zone were compared. The overlap was determined using the Dice similarity coefficient (DSC) and the average surface deviation between the realized and predicted ablation zone. The effect of tumor location, vascular proximity and liver cirrhosis on the overlap was quantified using the DSC and average surface deviation.

**Results** Nineteen patients and 21 ablations were included for analysis. The median realized volume did not significantly differ from the predicted volume, 25.7 cm<sup>3</sup> and 22.6 cm<sup>3</sup> respectively (p = 0.526). The median LAD and SAD of the realized ablation zone differed significantly from the manufacturer's prediction (51.9 mm vs 40.0 mm, p<0.001 and 36.9 mm vs 35.0 mm, p<0.001 respectively). The predicted ablation zone corresponds to the realized ablation zone with a mean DSC of 0.73 and mean average surface deviation of 3.04 mm. Tumor location and vascular proximity did not affect the overlap between the realized and predicted ablation zone. The effect of liver cirrhosis was not assessed due to a low sample size of HCC in a non-cirrhotic liver (n=2).

**Conclusion** The manufacturer's predicted volume of liver ablation zones corresponds well to the clinically realized ablation volume. However, the LAD and SAD are underestimated by the manufacturers. The shape and overlap of the predicted and realized ablation zone were sufficient. Further studies evaluating the effect of tumor- and liver characteristics on the correspondence of the predicted with the realized ablation zone with a larger patient cohort is needed.

## 1. Introduction

Liver cancer is the sixth most common cancer and the second leading cause of cancer-related death worldwide [1]. Hepatocellular carcinoma (HCC) is the most prevalent type of primary liver cancer (90%)[2]. 80-90% of the patients diagnosed with HCC have underlying liver cirrhosis caused by chronic viral infection, i.e., hepatitis B or C, alcohol abuses, metabolic liver diseases or nonalcoholic fatty liver diseases [3, 4].

According to the Barcelona Clinic Liver Cancer (BCLC) staging system and treatment schedule, thermal ablation is the preferred treatment for very early stage (BCLC 0), i.e. a single nodules  $\leq 2$  cm, and early stage HCC (BCLC A), i.e. a maximum of 3 lesions of  $\leq 3$  cm each, if the patient is not a candidate for liver transplantation [5]. In case of early-stage HCC with a single nodule ( $>2$  cm), the preferred treatment option is surgical resection. However, due to inadequate liver function, unfavorable tumor location,

poor general condition or co-morbidity, resection may be infeasible. In that case, thermal ablation is preferred over resection.

Radiofrequency ablation (RFA) and microwave ablation (MWA) are currently the most widely applied thermal ablation techniques for HCC treatment. Both techniques aim to induce tissue heating of at least 55-60°C to necrotize the tumor along with a clinical safety margin of at least 5 mm [6, 7]. RFA is based on a rapidly alternating current which excites the ions in the liver tissue, causing frictional heating. In MWA, electromagnetic waves cause polar molecules, predominantly water, to realign with the oscillating field, which generates heat through kinetic energy [8].

Thermal ablation has several advantages over surgical resection. It is less invasive, associated with shorter hospitalization and has lower complication rates [9, 10]. Nevertheless, reported local recurrence rates (LRR) after thermal ablation tend to be higher compared to LRR after surgical resection for noduli >2 cm [9, 10]. Laimer et al. found the LRR to be associated with ablation margins, since for each millimeter increase in minimal ablation margin, a 30% risk reduction for local recurrence (LR) was found [7]. No LR occurred if a clinical safety margin of >5 mm was obtained, but this was only accomplished in 37.5% of the ablations. These results are in accordance with several other studies that investigated the correlation between ablation margin and LR in HCC and colorectal liver metastasis (CRLM) [6, 7, 11, 12]. In these studies, the percentage of patients in whom the intended safety margin of >5 mm was achieved varied between 2.7% and 51.4%, indicating a discrepancy between predicted and realized ablation zone.

Manufacturers of ablation systems provide estimations of ablation zone dimensions based on treatment settings (time and wattage). These estimates, together with the interventional radiologist's experience, are used in clinical practice to specify the treatment settings needed to create an ablation zone large enough for complete tumor coverage with a safety margin of at least 5 mm. The manufacturer-provided predictions are mostly based on in vivo or ex vivo preclinical animal experiments [13]. Factors like vascular proximity, perfusion, tumor conductivity, tumor location and differences in liver conductivity due to underlying liver diseases, e.g., cirrhosis and steatosis, are neglected in these experiments. Nevertheless, computational modelling and ex vivo experiments demonstrate that these tumor- and liver characteristics influence the ablation zone [14-17]. This might lead to unreliable predictions of the ablation zone dimensions, resulting in the discrepancy between predicted and realized ablation zones.

Previous studies compared the manufacturer's predicted ablation zones dimensions to realized ablation zone dimensions but did not research the correspondence in shape and location of ablation zones [13, 18]. Nevertheless, these factors are of clinical importance, since the aim of ablation is to cover the tumor including a safety margin, making the treatment success not only size, but also shape dependent. Contrarily, metrics for target overlap and boundary discrepancy are frequently used for in vivo validation of computational models [19]. Since these metrics are not used to evaluate manufacturer's predictions, it is challenging to compare the reliability of computer models to the current clinical practice [19].

The aim of this study is to determine the correspondence in dimensions, volume, shape and overlap of the manufacturer's predicted ablation zones with the clinically realized ablation zones following thermal ablation of HCC. The secondary objective is to determine the effect of tumor- and liver characteristics on this correspondence.

## 2. Methods

## 2.1 Patients

In this retrospective study data of one completed and one ongoing prospective trial was used, IAMCOMPLETE and PROMETHEUS respectively [20]. The IAMCOMPLETE study is a prospective single center, cohort study investigating the feasibility of co-registration of intraprocedural computed tomography (CT) scans acquired immediately before and after liver ablation in patients with BCLC stage 0/A HCC. PROMETHEUS is a prospective multi-center study to evaluate the correlation between safety margin and LR after thermal ablation using image co-registration in patients with HCC. Patients older than eighteen years who were candidate for thermal ablation as discussed in a multidisciplinary board with HCC BCLC-0 or BCLC-A and who signed informed consent were included. In addition, patients with intermediate stage HCC (BCLC-B) with a maximum of 2 tumors  $\leq 5$  cm were also allowed in the PROMETHEUS study. Both studies received approval from the Leiden, Den Haag, Delft, Medical Ethical committee and were registered (IAMCOMPLETE: Clinicaltrials.gov, NCT04123340, PROMETHEUS: International Clinical Trial Registry Platform, NL9713). Within the informed consent, patients could give consent for their data to be used in other studies regarding thermal ablation in HCC, which enabled this retrospective analysis.

Between December 2019 and March 2021, 23 patients were included in the IAMCOMPLETE study and from September 2021 to July 2022, 18 patients were included in the PROMETHEUS study. The study cohort contained two double-enrolled patients. One patient was excluded due to advanced disease (BCLC B with more than 2 tumors) at treatment time, one patient did not receive ablation therapy due to declining liver function with hepatic encephalopathy and one patient was excluded due to technical failure of the CT-system, which made it impossible to acquire a post-ablation scan. Therefore, 36 patients, with 52 ablations were screened for inclusion. Lesions treated with multi-needle ablations ( $n = 21$ ), overlapping ablation zones ( $n = 4$ ), needle repositioning during treatment ( $n = 3$ ) and ablation of a local recurrence ( $n = 1$ ) were excluded. Ultimately, 23 ablations in 19 patients were included in this retrospective study (Figure 1). All patients gave informed consent to use their data for further research in the field of thermal ablation for HCC and underwent percutaneous thermal ablation in the Leiden University Medical Center (LUMC).

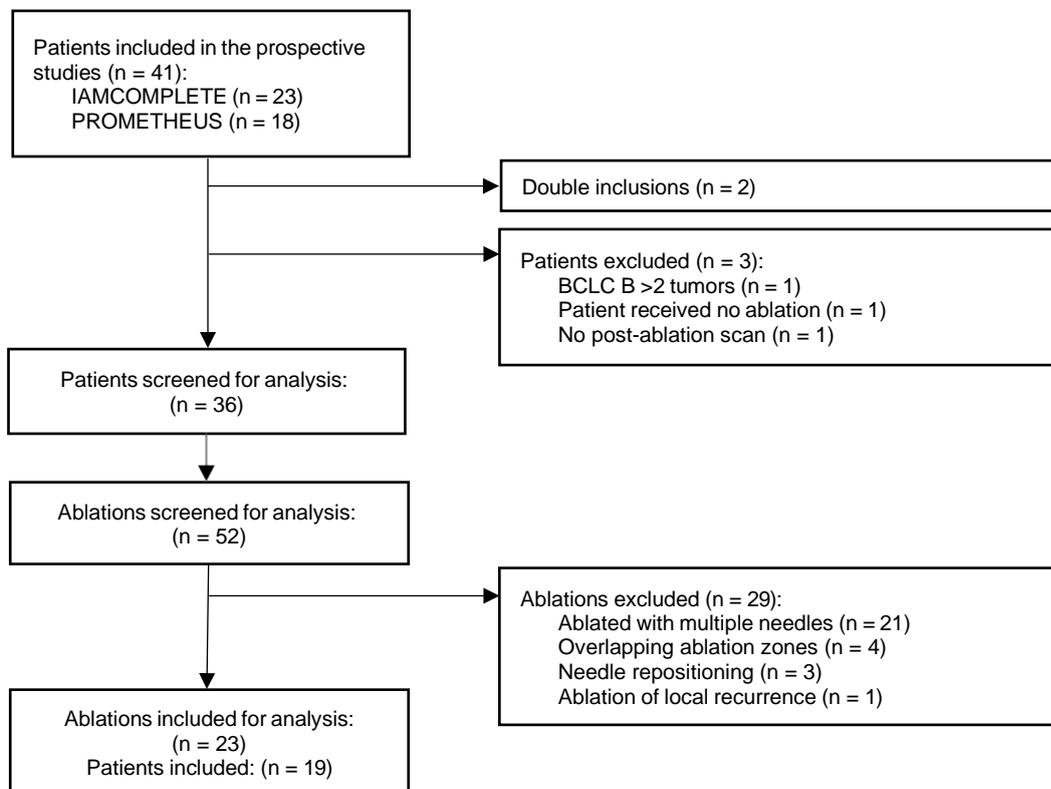


Figure 1: Flowchart of retrospective inclusion of ablations derived from two prospective studies for ablation zone analysis.



Figure 2: Example of a manufacturer’s predicted ablation zone, Emprint HP (Medtronic). The predicted long axis is 4.0 cm, short axes 3.5 cm, advancement 0.7 cm and the volume 25.7 cm<sup>3</sup> for a 10-minute in vivo liver ablation at 100 Watts.

## 2.2 Procedure

Ablations were performed under general anesthesia using ultrasonographic and/or CT guidance. All RFA procedures were performed using the single-needle Cool-tip Ablation system (Medtronic Trading NL B.V. (Covidien products), Eindhoven, the Netherlands). MWA ablations were performed using either the Amica (HS Hospital Service, Rome, Italy), Emprint or Emprint high power (HP) (Medtronic Trading NL B.V. (Covidien products), Eindhoven, the Netherlands) system.

Before start of the ablation, a dual-phase contrast-enhanced CT (CECT) with intravenous contrast was acquired. Two enhancement phases, arterial (bolus triggering in descending aorta + 10s) and portal-venous (bolus triggering in descending aorta + 50s) were acquired using a weight-dependent dose of the Xenetix 350 contrast medium. In case of an invisible lesion on ultrasound or a lesion with a challenging location, a CT hepatic arteriography (CTHA) was considered instead of a CT with intravenous contrast to improve lesion conspicuity. Prior to treatment, the patient was admitted to the angiography suite for insertion of a catheter in either the proper, left or right hepatic artery. In the CT room, a two-phase CTHA (+ 8s and + 22s) was acquired after direct infusion of 40 ml mixed bolus of Xenetix 350 and saline (1:1) at an injection rate of 3 – 3.5 ml/s. Both the CECT’s were acquired using a 320-slice spiral CT-scanner (Aquilion One, Canon Medical Systems Europe B.V., Zoetermeer, the Netherlands) in apnea and treatment position with a maximum slice thickness of 1.0 mm.

Image guided needle positioning was performed using ultrasound, CT or fusion of ultrasound and diagnostic imaging upon discretion by the interventional radiologist with over 10 years of experience. Settings of the ablation system (time and wattage) were determined based on manufacturer’s prediction, considering the tumor size with an anticipated ablative margin of at least 5 mm. After ablation treatment, tract ablations were performed to reduce the risk of bleeding and tumor seeding. Afterwards, a dual phase CECT, either a CT with intravenous contrast or CTHA, was acquired. At the end of the procedure, the interventional radiologist used the pre- and post-ablation CECT to determine whether the ablation was technically successful by visual inspection and 2D measurements. The ablation was deemed technically successful in case of complete tumor ablation with a margin of at least 5 mm. All included cases were deemed technically successful, since re-ablations were excluded from analysis due to needle repositioning or overlapping ablation zones.

## 2.3 Data

Patient and tumor characteristics were extracted from the study databases together with treatment specifications. In addition, the ablation procedure logs of the included patients were reviewed to determine whether needle repositioning took place. The following tumor- and liver characteristics were used for analysis:

- Tumor location: It was determined whether the tumor was subcapsular or centrally located. The tumor was considered subcapsular if the distance to the liver capsule was  $\leq 5$  mm, as measured in 2D on diagnostic imaging.

- Vascular proximity: The interventional radiologists classify the vascular proximity of tumors as  $> 5$  mm,  $< 5$  mm or tumor lies against the vascular structure in the ablation procedure logs. For this study, vascular proximity was defined as  $<5$  mm or  $>5$  mm according to the procedure log.
- Cirrhosis. The diagnosis of cirrhosis is based on medical history, blood tests, imaging and in some cases confirmed with pathology.

The dual-phase CT with intravenous contrast or CTHA acquired right before and after ablation were available for all patients. The scans were anonymized and saved as Digital Imaging and Communications in Medicine (DICOM) files in the study database.

Segmentations of the liver, tumor and ablation zone were available. Segmentations performed on arterial phase scan were used for analysis, according to the PROMETHEUS study protocol [20]. In case segmentation on the arterial phase scan was deemed unfeasible, contours created on the portal venous phase were used. Additionally, registration of pre- and post-ablation CECT was available based on semi-automatic intensity-based rigid registration using the liver contours of the post-ablation CECT (moving image) and pre-ablation CECT (fixed image). Where necessary, manual corrections were made to improve registration results. The CECT's were acquired in identical patient positioning during apnea. Therefore, rigid registration was sufficient since liver deformations caused by respiration and patient positioning were minimized, which usually hamper accurate rigid liver registration.

Image processing was performed using the in-house developed software deLIVERed (de Leiden Interactive Visualization en Registratie Editor), which includes an elastix-based registration algorithm implemented in a MeVisLab environment [21]. Supervised image processing was performed by a researcher and experienced interventional radiologist.

#### *2.4 Needle position reconstruction*

The manufacturer's predicted ablation zone is represented by an ellipsoid. This ellipsoid is centered around the needle and the advancement is indicated, representing the distance between the needle tip and the anterior part of the ablation zone (Figure 2). The manufacturer assumes a symmetric ellipsoid. Consequently, the second short axis, orthogonal to the first short axis, is of equal length. To locate the predicted ablation zone, the actual needle position should be specified. A needle position scan is not acquired as per protocol and might only be available in case of CT-guided procedures or if hydrodissection is applied. Therefore, a new extension in the deLIVERed software was created to perform needle position reconstructions.

The axial, sagittal and coronal view of the post-ablation scan, including the liver contour, ablation zone and the registered tumor of the pre-ablation scan formed the environment for this reconstruction (Figure 3a). The needle was assumed to be a rigid straight object. Therefore, only the entry point and needle tip were required for needle position reconstruction, which can be identified based on image visual assessment (Figure 3b and 3c). To locate the entry point, subcapsular liver hemorrhages or subcutaneous hemorrhages were identified. Gas bubbles formed by water vaporization during MWA were the main indicator for needle tip position. Additionally, tumor location was used, as in clinical practice needle tips were positioned just behind the tumor. Furthermore, the needle trajectory might be visible as a hypodense trajectory caused by tract ablation, mainly in the portal venous post-ablation scan. The clinical feasibility of needle trajectories was taken into consideration in needle position reconstruction. In case the image lacks features to identify the entry point, the reconstruction was based on an estimation of the clinically most likely used needle position.

The reconstructions were performed by the researcher and supervised by either an interventional radiologist or technical physician experienced in the field of thermal ablation. After reconstruction, the needle and an ellipsoid representing the manufacturer's predicted ablation zone were projected on the image (Figure 3d), which can be visualized as a 3D structure (Figure 4). deLIVERed automatically

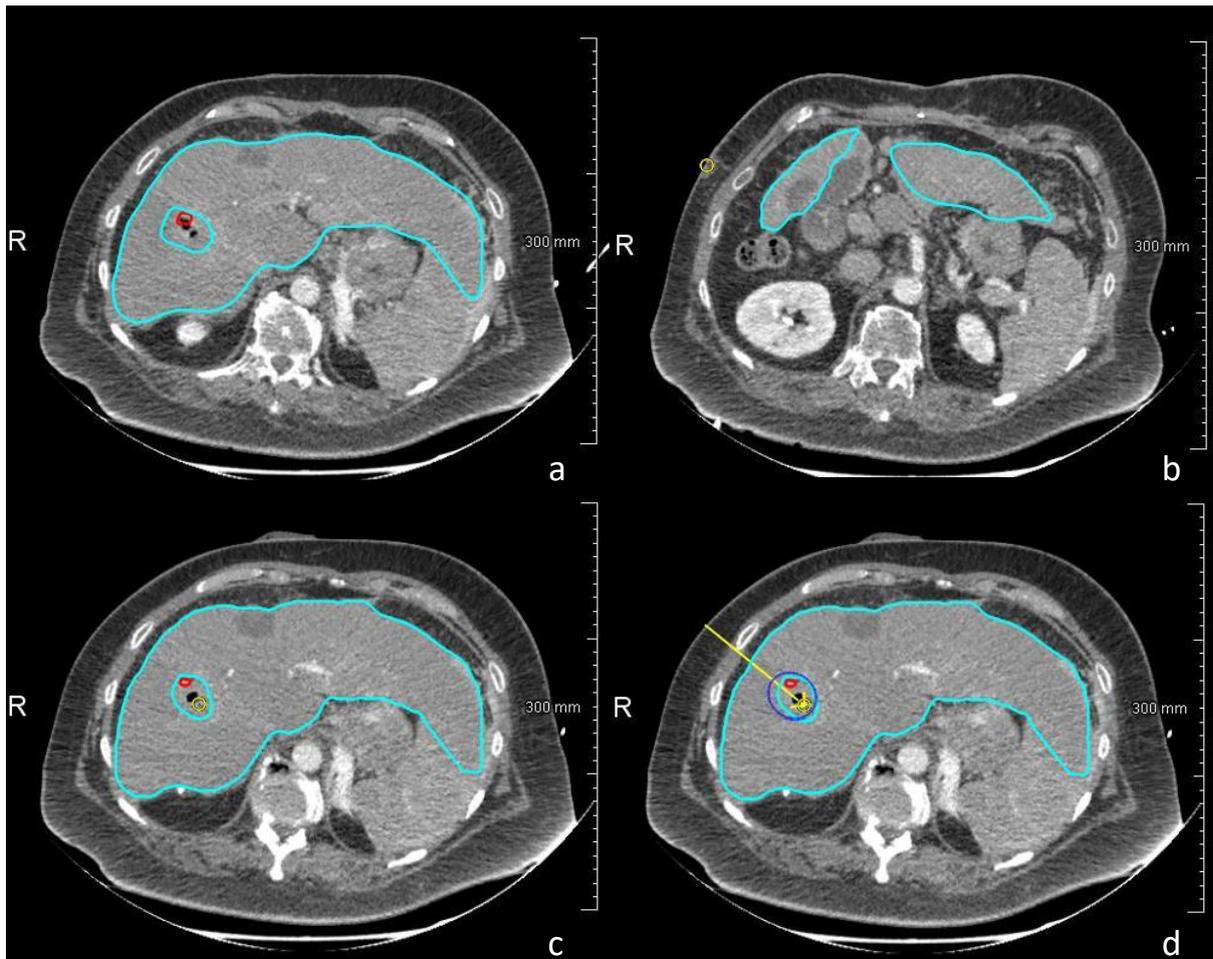


Figure 3. Overview of needle position reconstruction in deLIVERed on an arterial phase axial post-ablation scan. a) Overview of the liver contour with the segmented ablation zone (*cyan*) and the tumor contour (*red*). b) Identification of needle entry point (*yellow circle*) based on small subcutaneous hemorrhage. c) Identification of needle tip position (*yellow circle*) using the gas bubbles created by water vaporization and tumor location. d) Reconstructed needle trajectory (*yellow line*) and the projected predicted ablation zone (*dark blue*)

determines the longest axis of the ablation zone and its size; long axis diameter (LAD). The short axis diameter (SAD) and medium axis diameter, orthogonal to the SAD, were also determined.

### 2.5 Statistical analysis

The segmented ablation zone dimensions and volume (in  $\text{cm}^3$ ) were compared to the manufacturer's predicted dimensions and volume. The differences between the predicted and realized diameters were calculated. Since the manufacturer assumes equal length of the two short axes, the short and medium axis of the realized ablation zone were collectively compared as the SAD.

The overlap of the predicted and realized ablation zone is evaluated with the Dice similarity coefficient (DSC). The average surface deviation was defined as the directed absolute average Hausdorff distance between the realized and predicted ablation zone. These outcome measures were used to evaluate if tumor location, vascular proximity and image presence of liver cirrhosis influence the correspondence between the manufacturer's prediction and clinically realized ablation zone.

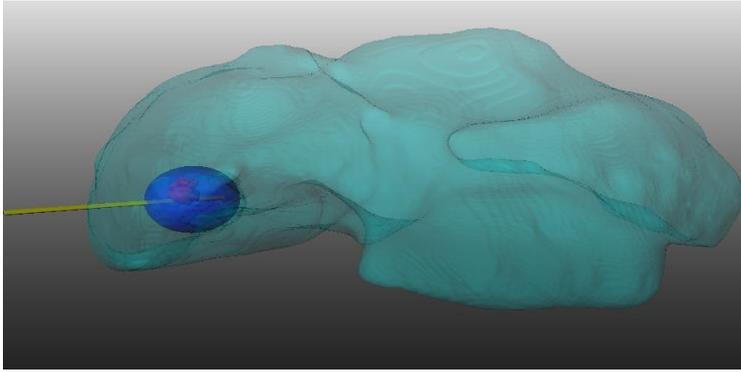


Figure 4. 3D model of the liver (*cyan*), tumor (*red*), reconstructed needle trajectory (*yellow*) and the predicted ablation zone as results of the needle reconstruction shown in Figure 3.

Surface deviation occurs in two directions: either the realized ablation zone is larger than predicted zone (positive direction) or it is smaller (negative direction). In addition to the average surface deviation, the maximum distance in the negative direction was determined. The number of cases was determined in which this negative maximum surface deviation is equal to or more than 5 mm, since treatment success is at risk in these cases. The relative undertreated ablation volume, defined as the volume within ablation zone prediction, but not ablated, relative to the predicted ablation zone volume was determined. These outcome measures are illustrated in Figure 5.

Dimension based variables were described using the median and interquartile range (IQR) while the overlap outcome measures were described using the mean and standard deviation (SD). Differences were assessed with the Mann-Whitney U test in case of unpaired data and the Wilcoxon signed-rank test for paired data. A p-value <0.05 was considered significant.

### 2.6 Peripheral ablation zones

For peripherally located ablation zones where the predicted ablation zone extended past the liver contour, the outcome measures are adjusted. In these cases, the DSC and average surface deviation were based on the part of the predicted ablation zone that is projected within the liver contour (Figure 6). deLIVERed determines the volume of the predicted ablation zone within the liver contours, which is then assumed to be the predicted volume of the manufacturer.

### 2.7 Reconstruction error

In case a per-procedural CT-scan was acquired with ablation needle position, these scans were anonymized and extracted from the Picture Archiving and Communication System (PACS) as DICOM

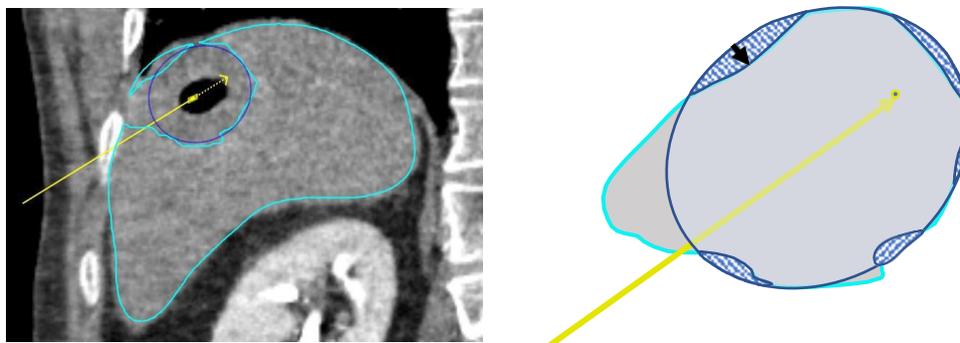


Figure 5. Outcome measures. Left: coronal arterial phase post-ablation CT-scan with projected needle reconstruction (*yellow*) and the predicted ablation zone projected over the segmented realized ablation zone (*cyan*). Right: schematic illustration of the maximum negative surface deviation (*black arrow*) and undertreated ablation volume (*blue shaded areas*). Note: Visualization is in 2D while all outcome measures were based on 3D measurements.

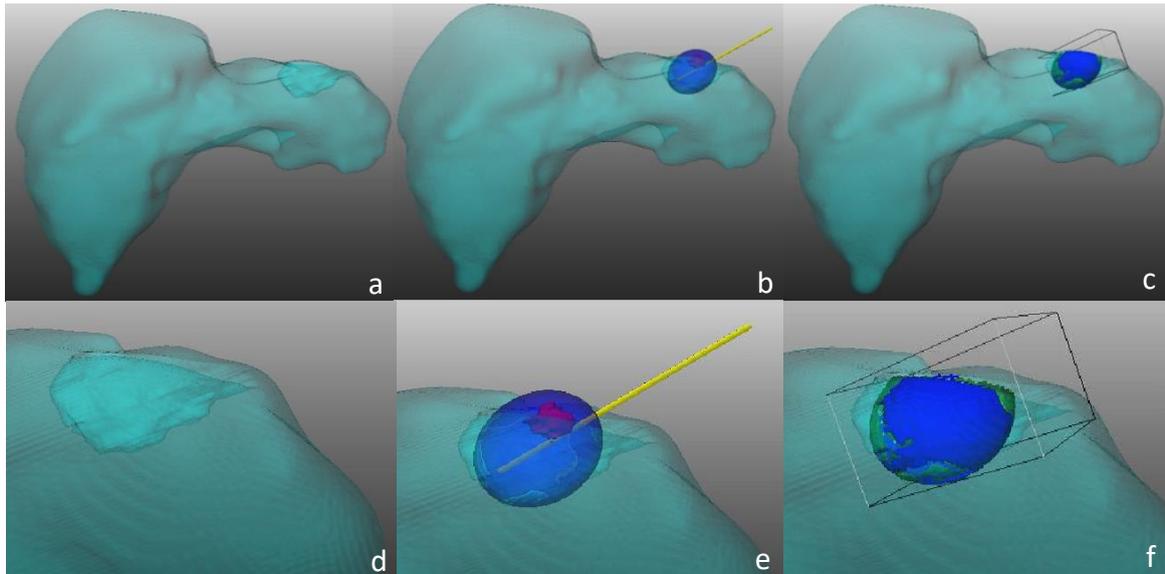


Figure 6. Example of predicted ablation zone extending past the liver contour and adapted outcome measures. a) 3D model of the liver with segmented realized ablation zone. b) 3D model of the liver with segmented ablation zone (*cyan*), tumor (*red*), the reconstructed needle (*yellow*) and predicted ablation zone (*blue*). c) 3D model of the liver with segmented ablation zone (*cyan*). In blue and green the surface of the ablation zone within the liver contour is projected. Only this surface is used for surface deviation calculation. The DSC is based only on this predicted volume. d, e, f) zoomed in situation of the 3D model above.

files. Rigid registration was performed to match the needle position scan with the post-ablation scan. Discrepancies in needle tip position were determined using the Euclidean distance between both needle tips. The DSC and average surface deviation using the scan-based needle position were calculated and compared to the outcomes of the reconstructed needle. Based on these results, the effect of needle position reconstruction error on the outcome measures were evaluated.

### 3. Results

#### 3.1 Patients

Needle position reconstruction was feasible for all ablations, supported by a needle position scan in eight cases (Table 1, Appendix A). The needle entry point was identified by either subcutaneous hemorrhage ( $n = 9$ ), subcapsular hemorrhage ( $n = 2$ ) or based on the needle position scan ( $n = 8$ ). In four cases no reference point was visible (Table 1, Appendix A). In two cases the ablation zone contained a hypoperfused area next to the ablation zone as a result of thermal vascular damage. It is a challenge to distinguish between those two regions. These cases were excluded from data analysis, resulting in nineteen patients included with a total of 21 ablations.

Baseline and treatment characteristics of the included patients can be found in Table 1 and 2 respectively. The cohort consisted of fourteen males and five females with a mean age of  $70 \pm 9.75$  years. Seventeen of the included patients had liver cirrhosis with an etiology of alcohol abuses (53%), hepatitis C (6%), nonalcoholic steatohepatitis (NASH) (29%), autoimmune hepatitis (6%) or cryptogenic (6%). Thirteen (61.9%) tumors with a subcapsular location were included, while the distance to the vascular structure was  $>5$  mm for twelve tumors (57%) and  $<5$  mm for nine tumors (43%). Nineteen ablations (90%) were performed using MWA, while RFA was used in two ablations (10%). Emprint and Emprint HP (33% and 52%, respectively) were the most used ablation systems.

#### 3.2 Ablation zone dimensions comparison

The median volume of the realized ablation zone did not significantly differ from the predicted volume ( $25.7 \text{ cm}^3$  versus  $22.6 \text{ cm}^3$ ,  $p = 0.526$ ). The LAD and SAD of the realized ablation zone differed

**Table 1. Patient characteristics (n=19)**

Age (years), mean (SD)	70	10.0
Sex, n (%)		
Male	14	74%
Female	5	26%
Child Pugh score, n (%)		
A	15	79%
B	3	16%
Unknown <sup>†</sup>	1	5%
BCLC stage, n (%)		
Very early stage (O)	7	37%
Early stage (A)	10	53%
Intermediate stage (B)	2	11%
Portal Hypertension, n (%)		
Yes	13	68%
No	6	32%
Cirrhosis, n (%)		
Yes	17	89%
No	2	11%
Etiology cirrhosis, n (%)		
Alcohol abuses	9	53%
Autoimmune hepatitis	1	6%
Cryptogenic	1	6%
Hepatitis C	1	6%
NASH	5	29%
Amount of lesions <sup>‡</sup> , n (%)		
1	12	63%
2	3	16%
3	3	16%
5	1	5%

<sup>†</sup>Unknown due to missing Albumin level

<sup>‡</sup>Total amount of lesions differs from number of ablations analyzed, since not all lesions fall within inclusion criteria (Figure 1)

BCLC = Barcelona clinical Liver cancer; NASH = nonalcoholic steatohepatitis; SD = Standard deviation

**Table 2. Tumor characteristics and treatment specifications (n=21)**

Lesion size (mm), mean (SD)	18.0	6.6
Vascular proximity, n (%)		
<5 mm	9	43%
>5 mm	12	57%
Subcapsular located, n (%) <sup>†</sup>		
Yes	13	62%
No	8	38%
Ablation type, n (%)		
MWA	19	90%
RFA	2	10%
System used, n (%)		
Amica (MWA)	1	5%
Cooltip (RFA)	2	10%
Emprint (MWA)	7	33%
Emprint HP (MWA)	11	52%
Wattage used (MWA), n (%)		
60	1	5%
75	1	5%
100	10	53%
150	7	37%

<sup>†</sup>Subcapsular tumors were defined as tumors located at <5 mm from the liver capsule

MWA = Microwave ablation; RFA = Radiofrequency ablation; SD = Standard deviation; HP = High power

**Table 3: The median + interquartile range (IQR) of the predicted versus realized ablation zone dimensions**

	Manufacturer's prediction	Realized ablation zone	p value <sup>†</sup>
SAD (mm)	35.0 [34.0 - 39.0]	36.92 [31.64 - 42.97]	<0.001
LAD (mm)	40.0 [39.0 - 47.0]	51.93 [46.48 - 56.99]	<0.001
Volume (cm <sup>3</sup> )	22.6 [19.7 - 38.7]	25.73 [14.85 - 35.81]	0.526

<sup>†</sup> Assessed with Wilcoxon signed-rank test

SAD = Short axis diameter; LAD = Long axis diameter

Table 4: The mean (SD) of the DSC and Average surface deviation (mm) between the created and predicted ablation zone for all cases (n=21). For the Emprint and Emprint HP cases (n=18) the effect of tumor location and vascular proximity on the DSC and Average surface deviation is determined.

	n	DSC			Average surface deviation (mm)		
		Mean	SD	p-value <sup>†</sup>	Mean	SD	p-value <sup>†</sup>
<b>Total</b>	21	0.73	0.09		3.04	1.31	
Emprint (HP)	18	0.75	0.08		2.81	0.84	
<i>Tumor location</i>							
Subcapsular	10	0.75	0.08	0.762	2.72	0.75	0.829
Central	8	0.73	0.09		2.93	0.99	
<i>Vascular proximity</i>							
>5 mm	10	0.75	0.08	0.897	2.54	0.75	0.146
<5 mm	8	0.74	0.09		3.16	0.87	

<sup>†</sup>Assessed with Mann-Whitney U test

DSC = Dice similarity coefficient; SD = standard deviation; HP = High Power

significantly from the manufacturer's prediction (51.9 mm versus 40.0 mm,  $p < 0.001$  and 36.9 mm versus 35.0 mm,  $p < 0.001$  respectively) (Table 3).

The manufacturer's prediction for both the SAD and LAD was smaller than the realized ablation zone, with a median deviation of 1.57 mm (IQR: -3.06 mm – 6.79 mm) and 10.65 mm (IQR: 6.28 mm – 16.93 mm). Figure 7 visualizes the differences in millimeters between the predicted ablation zone and the realized ablation zone for the SAD and LAD.

A complete overview of the ablation zone diameters and volume for each individual case is presented in Table 2 in Appendix A.

### 3.3 Ablation zones overlap comparison

The predicted ablation zone corresponded to the realized ablation zone with a mean DSC of  $0.73 \pm 0.09$  and mean average surface deviation of  $3.04 \text{ mm} \pm 1.31 \text{ mm}$ . When omitting other ablation systems then Emprint or Emprint HP a DSC of  $0.75 \pm 0.08$  and average surface deviation of  $2.81 \text{ mm} \pm 0.84$  were obtained (Table 4).

No significant difference was found in DSC or average surface deviation for tumors located subcapsular or central (DSC: 0.75 vs 0.73,  $p = 0.762$ , average surface deviation: 2.72 mm vs 2.93 mm,  $p = 0.829$ ). Vascular proximity had no significant effect on the DSC and average surface deviation (DSC 0.75 vs 0.74,  $p = 0.897$ ; average surface deviation: 2.54 mm vs 3.16 mm,  $p = 0.146$ ) (Table 4). Only two ablations were performed in a non-cirrhotic liver. Therefore, this characteristic was excluded from analysis (Table 1).

In fifteen out of 21 cases (71.4%) the maximum negative surface deviation of the ablated contour and predicted ablation zone was  $\geq 5.0 \text{ mm}$ . The DSC and average surface deviation did not significantly differ for a maximum negative surface deviation  $< 5.0 \text{ mm}$  and  $\geq 5.0 \text{ mm}$  (DSC  $0.75 \pm 0.05$  versus  $0.73 \pm 0.10$ ,  $p = 0.850$ , average surface deviation;  $2.71 \pm 0.45 \text{ mm}$  versus  $3.18 \pm 1.52 \text{ mm}$ ,  $p = 0.6222$ ) (Table 5). The relative undertreated ablation volume differed significantly for a maximum negative surface deviation  $< 5.0 \text{ mm}$  and  $\geq 5.0 \text{ mm}$  ( $30.36 \pm 16.67\%$  versus  $4.65 \pm 5.20\%$ ,  $p < 0.001$ ).

### 3.4 Reconstruction error

All eight needle position scans were registered to the post-ablation scans. The tip error varied between 5.30 mm and 37.95 mm (Table 6). An increase in DSC and average surface deviation using the needle position based on the needle scan was observed in only one case. In case 23 the outcome measures stayed nearly the same, despite a tip error of 7.58 mm. In all other cases, the scan-based needle position resulted in lower correspondence between the predicted and realized ablation zone.

Table 5: The mean (SD) of the DSC, Average surface deviation (mm) and relative undertreated ablation volume for all cases (n=21). A distinction is made for maximum negative surface deviation < 5mm and ≥5 mm.

	n	DSC			Average surface deviation (mm)			Relative undertreated ablation volume (%)		
		Mean	SD	p-value†	Mean	SD	p-value†	Mean	SD	p-value†
<b>Total</b>	21	0.73	0.09		3.04	1.31		23.01	18.52	
<i>maximum negative surface deviation</i>										
<5 mm	6	0.75	0.05	0.850	2.71	0.45	0.622	4.64	5.20	<0.001
≥5 mm	15	0.73	0.10		3.18	1.52		30.36	16.67	

†Assessed with Mann-Whitney U test

DSC = Dice similarity coefficient; SD = standard deviation

Table 6: Tip error (mm) in needle reconstruction of the eight cases in which a needle position scan was acquired. The DSC and Average surface deviation (mm) for the scan-based needle position and reconstructed needle position are presented.

Case	Slice thickness (mm)	Number of axial slices	Tip error (mm)	Scan based needle position		Reconstructed needle position	
				DSC	Average surface deviation (mm)	DSC	Average surface deviation (mm)
9	4.0	3	19.13	0.28	7.19	0.74	3.0
13	1.0	171	5.30	0.70	2.92	0.64	3.4
16	4.0	3	15.04	0.51	4.27	0.79	1.8
17	1.0	228	37.95	0.02	14.86	0.69	2.7
18	1.0	202	11.68	0.76	3.34	0.81	2.1
19	1.0	168	6.47	0.56	4.54	0.76	3.0
22	1.0	260	5.91	0.59	4.09	0.64	3.8
23	4.0	3	7.58	0.80	2.55	0.80	2.6

DSC: Dice similarity coefficient

#### 4. Discussion

The main study objective was to determine the correspondence in dimensions, volume, shape and overlap of the manufacturer's predicted ablation zones with the clinical realized ablation zones following thermal ablation of HCC. The realized volume did not differ from the predicted ablation zone volume. However, the LAD and SAD are both larger compared to the manufacturer's prediction, with a median difference of 10.65 mm and 1.57 mm, respectively. The shape and overlap of the predicted and realized ablation zones were sufficient with a DSC of 0.73 and average surface deviation of 3.04 mm. However, it is questionable whether the predictions are reliable enough for clinical demands, since the maximum negative surface deviation was ≥5 mm in fifteen out of 21 cases.

Shyn et al. observed moderate correlations between realized and predicted ablation zone sizes using the Emprint system [22]. They found better correlations for the SAD and volume compared to the LAD, which is in line with our results. A study by Young et al. compared the realized ablation zone diameters using MWA in liver malignancies with the prediction of the Emprint system [18]. Their study demonstrates that the manufacturer overestimates the ablation zone diameters in all three directions, which is contrary to our results. In another study, no differences between the realized and predicted ablation zone diameters of the Emprint system were found [23]. Both study methods used post-ablation scans acquired one month after ablation to determine the ablation zone diameters in 2D. Therefore, ablation zone shrinkage likely had an impact on their results, as ablation zones usually shrink over time [24, 25].

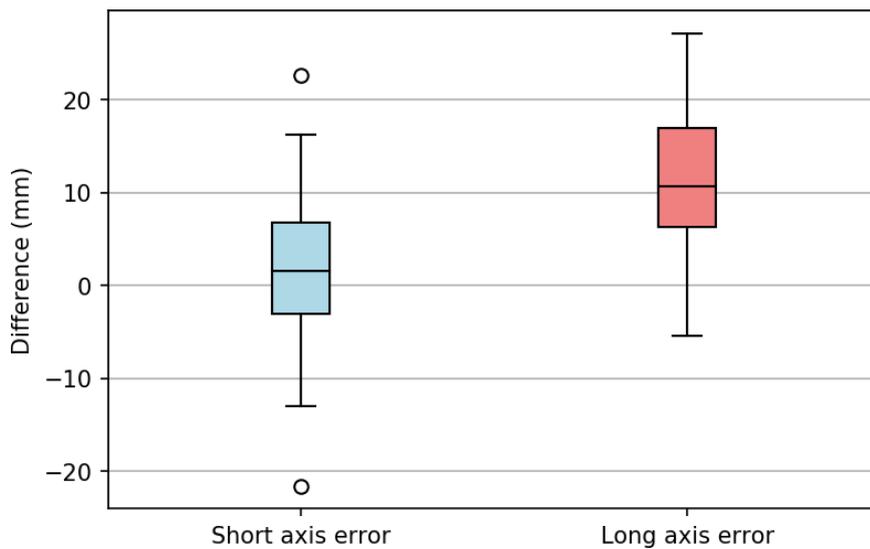


Figure 7. Boxplot of the difference in mm between the realized axis and the predicted axis. A negative difference indicates that the realized axis is smaller than the predicted axis. SAD median difference is 1.57 mm [-3.06 mm – 6.79 mm], LAD median difference is 10.65 mm [6.28 mm – 16.93 mm].

Tumor location and vascular proximity do not influence the correspondence between the manufacturer’s prediction and the clinical realized ablation zone. These results are in accordance with previous research evaluating the effect of vascular proximity on ablation zone dimensions in MWA [18, 23]. It is known that MWA is less sensitive to the heat-sink effect compared to RFA, which is confirmed by these results [26, 27]. However, this does not imply that vascular proximity has no effect on ablation zones at all. One of the Emprint HP ablation zones with the lowest DSC (0.61) visually seems to be affected by vascular proximity (Figure 1, Appendix A).

Unfortunately, the lack of inclusions of ablations in non-cirrhotic livers resulted in the inability to compare the reliability of the manufacturer’s prediction in cirrhotic and non-cirrhotic livers. In literature, the effect of cirrhosis on MWA is inconsistent. Computer modelling showed that tissue characteristics of surrounding liver parenchyma, mainly tissue conductivity, influences the ablation zone created with MWA, resulting in larger ablation zones in cirrhotic and fatty livers [28, 29]. A retrospective study by Heerink et al. supported these findings and found larger ablation volumes in 45 HCC ablations in cirrhotic livers compared to 45 CRLM ablations in non-cirrhotic livers [30]. Contrarily, a prospective study comparing 32 HCC ablations to 19 ablations of secondary liver malignancies observed smaller ablation zone diameters in HCC ablations compared to CRLM (LAD: 52 mm versus 65 mm, SAD: 37 mm versus 41 mm)[31]. The difference was hypothesized to be caused by less effective active heating in cirrhotic livers, due to a reduced water content. Young et al. also found that the reliability of manufacturer’s prediction is dependent on liver parenchyma characteristics. The ablation zone LAD in fatty livers, as diagnosed on imaging, corresponded better to the manufacturer’s predicted LAD, compared to non-fatty livers. They found a larger LAD in fatty livers compared to non-fatty livers [18]. Whereas De Cobelli et al. did not find any evidence of pathophysiological liver condition effecting the reliability of Emprint predictions [23]. These contradicting results require more research into the effect of cirrhotic and fatty liver parenchyma on the ablation zone dimensions and shape and the reliability of manufacturer’s prediction.

The effect of tumor- and liver characteristics on ablation zone dimensions has been extensively studied using computational modeling [19]. Patient-specific ablation therapy planning is assumed to improve the reliability of ablation zone predictions. A recently published article of Hoffer et al.

compared the accuracy of a computational model with the manufacturer's prediction in six in vivo ablations in porcine livers [32]. The average surface error as well as the maximum surface error were significantly smaller for the computational model compared to the manufacturer's prediction (average error: 1.1 mm versus 2.5 mm, maximum error: 5.2 mm versus 7.8 mm). However, this model is not clinically validated yet. A prospective study evaluating a computational model in primary and secondary liver tumor ablation, concluded that the model was accurate enough for clinical demands with a DSC of  $0.62 \pm 0.14$  and average surface deviation of  $3.4 \pm 1.7$  mm [33]. Other clinical retrospective studies found DSCs ranging from 0.42 – 0.73 for their model [34-36]. However, none of these studies compared their model to the manufacturer's prediction. Further research regarding computational modelling should incorporate a comparison to the manufacturer's prediction to conclude on the clinical value of computational modelling.

In this study, a new method for needle position reconstruction was applied. A strength of the current method is its handling of peripherally located ablations. In previous research, the diameter reaching the liver capsule was neglected in analysis or inclusions were limited to central ablations [18, 30, 31]. As low inclusion rates are a challenge in these studies, extension of inclusion criteria is advantageous. The LAD and SAD are compared to the diameters of the prediction chart of the manufacturer. These diameters are not adjusted based on the predicted ablation zone within the liver contour as is done with the predicted volume. The results indicate an underestimation of the LAD and SAD by the manufacturer's, which might be larger when neglecting the predicted ablation zone outside the liver contour in diameter comparison.

The proposed method is limited to single needle ablations, which is another limitation. Multi-needle ablations are more frequently used in RFA. Focusing on MWA should be considered for further research.

The outcome measures are based on those used in literature to make our results comparable with previous research. The DSC is the most used metric in image segmentation, but is size dependent, making it more susceptible to small differences in peripheral or small ablations [37, 38]. Nevertheless, the combination of the DSC with the average surface deviation, which is a frequently used outcome measure for boundary errors and not size dependent, gives a good representation of overlap [37, 38]. The maximum negative surface deviation is sensitive to outliers. To overcome this limitation, the relative undertreated ablation volume was also calculated as additional measurement (Figure 5). This outcome measure significantly differs for maximum negative surface deviation  $<5$  mm and  $\geq 5$  mm. However, the DSC and average surface deviation do not differ between these two groups. Hence, the DSC and average surface deviation are presumably not related to technical success of the ablation. The outcome measures are likely influenced by tissue shrinkage during ablation. Within literature, a great variation in tissue shrinkage is reported based on in vivo and ex vivo experiments, varying between 5 – 30% for RFA and 20 – 65% for MWA [39-41]. There is insufficient knowledge on how tissue contraction is influenced by factors such as cirrhosis, ablation power and treatment time.

Needle position reconstruction is based on image visual assessment, introducing insecurities in the method. In four cases no reference point for needle entry point identification was present. Availability of a needle position scan helped to reconstruct the needle. Nevertheless, found tip varied between 5.30 and 37.95 mm in the eight cases in which a needle position scan was acquired. Case 17 and case 9 have the largest tip errors, 37.95 and 19.13 mm respectively. These large deviations are dedicated to needle repositioning after CT-scan acquisition and before start of ablation. In both cases, the location of the ablation zone and needle tip do not correspond. This is confirmed by the low DSC obtained with the scan-based needle position, 0.02 and 0.28, respectively. These needle tip errors are not representative for the inaccuracy in needle position reconstruction. An axial scan of only 3 slices was acquired in three cases, hampering an accurate registration. When omitting the cases of a 3-slice needle scan and needle repositioning after needle scan acquisition, four cases are left (13, 18, 19 and

22) with a reconstruction tip error varying between 5.30 and 11.68 mm. In one case, the scan-based needle position resulted in a higher DSC and average surface deviation (0.70 and 2.92 mm versus 0.64 and 3.4 mm), while in the other three cases lower outcome measures were obtained. Needle position reconstructed might be biased by visibility of the segmented ablation zone on the post-ablation scan. This might have influenced the needle position reconstruction, explaining these favorable results with the reconstructed needle position compared to the scan-based needle position.

However, the scan-based needle position also contains several uncertainties. The needle position scans were acquired without contrast and under breathing. Therefore, accurate registration is hampered, resulting in registration errors. Additionally, needle artifacts make precise needle identification challenging. These factors might contribute to the found tip errors.

In further research, needle position scans could help to overcome insecurities in needle position reconstruction. However, only a full liver scan which includes treatment needle position, favorably acquired in apnea, will increase the reliability of the results.

A limitation of this study is the small patient cohort, which also led to the inability to evaluate and compare different ablation systems. Only two RFA cases and one ablation with the Amica system were included. Therefore, an analysis per ablation system would not be meaningful. However, the reliability of ablation zone predictions differs between each ablation system. For a long time, MWA was limited with unpredictable ablation zones. The Emprint system uses Thermosphere™ Technology to overcome this limitation [23, 42, 43]. The limitation in predictability of the Amica system reflected in the one Amica ablation analyzed (DSC: 0.52, average surface deviation: 7.4 mm). Therefore, it was chosen to separately evaluate the Emprint cases on its overlap.

Segmentation inaccuracies might have influenced the results. Ischemic regions, caused by thermal vascular damage, adjacent to the ablation zone cannot be distinguished from the coagulated ablation zone [44]. Additionally, although needle repositioning during ablation was an exclusion criterion, needle retraction while ablating could have occurred without registration in the procedure log. These needle movements as well as segmentation errors might have contributed to the large deviation observed in LAD (10.65 mm) (Figure 8).

Further research with a larger patient cohort is needed. To enlarge the inclusion of ablations in non-cirrhotic livers, inclusion of CRLM should be considered. This study revealed that in fifteen out of the 21 cases the negative maximum surface deviation was larger than 5 mm. These locations might be at risk for insufficient ablation margins and local tumor progression (LTP). However, these results also might be influenced by tissue shrinkage. Further studies should consider correlating these locations to found ablation margins and LTP to investigate whether these negative surface deviations have clinical consequences.

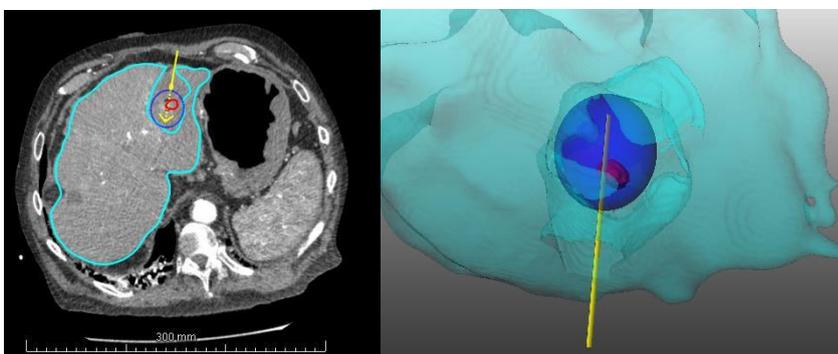


Figure 8. Example of possible needle retraction. Axial arterial post-ablation scan and 3D model containing the liver with segmentation ablation zone (cyan), tumor (red), reconstructed needle position (yellow) and projection of predicted ablation zone (blue). A bulge of the realized ablation zone in line with the needle trajectory can be observed.

## 5. Conclusion

The manufacturer's predicted volume of liver ablation zones corresponds well to the clinically realized ablation volume. The LAD and SAD are underestimated by the manufacturers. The shape and overlap of the predicted and realized ablation zone were sufficient with a DSC of 0.73 and average surface deviation of 3.04 mm. However, the maximum negative surface deviation was  $\geq 5$  mm in fifteen out of 21 cases, making it questionable if the predictions are of sufficient reliability for clinical demands. Further studies evaluating the effect of tumor- and liver characteristics on the correspondence of the predicted with the realized ablation zone with a larger patient cohort is needed.

## References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136(5):E359-86.
2. Grandhi MS, Kim AK, Ronnekleiv-Kelly SM, Kamel IR, Ghasebeh MA, Pawlik TM. Hepatocellular carcinoma: From diagnosis to treatment. *Surg Oncol*. 2016;25(2):74-85.
3. Wallace MC, Preen D, Jeffrey GP, Adams LA. The evolving epidemiology of hepatocellular carcinoma: a global perspective. *Expert Rev Gastroenterol Hepatol*. 2015;9(6):765-79.
4. Fattovich G, Stroffolini T, Zagni I, Donato F. Hepatocellular carcinoma in cirrhosis: incidence and risk factors. *Gastroenterology*. 2004;127(5 Suppl 1):S35-50.
5. Reig M, Forner A, Rimola J, Ferrer-Fàbrega J, Burrel M, Garcia-Criado Á, et al. BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. *J Hepatol*. 2022;76(3):681-93.
6. Kim YS, Lee WJ, Rhim H, Lim HK, Choi D, Lee JY. 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(3):758-65.
7. Laimer G, Schullian P, Jaschke N, Putzer D, Eberle G, Alzaga A, 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):2463-72.
8. Chu KF, Dupuy DE. Thermal ablation of tumours: biological mechanisms and advances in therapy. *Nat Rev Cancer*. 2014;14(3):199-208.
9. Shin SW, Ahn KS, Kim SW, Kim T-S, Kim YH, Kang KJ. Liver Resection Versus Local Ablation Therapies for Hepatocellular Carcinoma Within the Milan Criteria: A Systematic Review and Meta-analysis. *Annals of Surgery*. 2021;273(4):656-66.
10. Xu XL, Liu XD, Liang M, Luo BM. Radiofrequency Ablation versus Hepatic Resection for Small Hepatocellular Carcinoma: Systematic Review of Randomized Controlled Trials with Meta-Analysis and Trial Sequential Analysis. *Radiology*. 2018;287(2):461-72.
11. 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(1):166-75.
12. Wang CZ, Yan GX, Xin H, Liu ZY. Oncological outcomes and predictors of radiofrequency ablation of colorectal cancer liver metastases. *World J Gastrointest Oncol*. 2020;12(9):1044-55.
13. Ruiter SJS, Heerink WJ, de Jong KP. Liver microwave ablation: a systematic review of various FDA-approved systems. *Eur Radiol*. 2019;29(8):4026-35.
14. Huang HW. Influence of blood vessel on the thermal lesion formation during radiofrequency ablation for liver tumors. *Med Phys*. 2013;40(7):073303.
15. Ahmed M, Liu Z, Humphries S, Goldberg SN. Computer modeling of the combined effects of perfusion, electrical conductivity, and thermal conductivity on tissue heating patterns in radiofrequency tumor ablation. *Int J Hyperthermia*. 2008;24(7):577-88.
16. Liu Z, Ahmed M, Gervais D, Humphries S, Goldberg SN. Computer modeling of factors that affect the minimum safety distance required for radiofrequency ablation near adjacent nontarget structures. *J Vasc Interv Radiol*. 2008;19(7):1079-86.
17. Siriwardana PN, Singh S, Johnston EW, Watkins J, Bandula S, Illing RO, et al. Effect of Hepatic Perfusion on Microwave Ablation Zones in an Ex Vivo Porcine Liver Model. *J Vasc Interv Radiol*. 2017;28(5):732-9.
18. Young S, Rivard M, Kimyon R, Sanghvi T. Accuracy of liver ablation zone prediction in a single 2450MHz 100 Watt generator model microwave ablation system: An in human study. *Diagn Interv Imaging*. 2020;101(4):225-33.
19. Erp GCMv. Computer modelling of ablation zones in Radio-frequency ablation and Microwave ablation in liver tissue: a systematic review. 2022.

20. Oosterveer TTM, van Erp GCM, Hendriks P, Broersen A, Overduin CG, van Rijswijk CSP, et al. Study Protocol PROMETHEUS: Prospective Multicenter Study to Evaluate the Correlation Between Safety Margin and Local Recurrence After Thermal Ablation Using Image Co-registration in Patients with Hepatocellular Carcinoma. *Cardiovasc Intervent Radiol*. 2022;45(5):606-12.
21. Klein S, Staring M, Murphy K, Viergever MA, Pluim JP. elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging*. 2010;29(1):196-205.
22. Shyn PB, Bird JR, Koch RM, Tatli S, Levesque VM, Catalano PJ, et al. Hepatic Microwave Ablation Zone Size: Correlation with Total Energy, Net Energy, and Manufacturer-Provided Chart Predictions. *J Vasc Interv Radiol*. 2016;27(9):1389-96.
23. De Cobelli F, Marra P, Ratti F, Ambrosi A, Colombo M, Damascelli A, et al. Microwave ablation of liver malignancies: comparison of effects and early outcomes of percutaneous and intraoperative approaches with different liver conditions : New advances in interventional oncology: state of the art. *Med Oncol*. 2017;34(4):49.
24. Zimmermann M, Kuhl C, Keil S. Characteristic changes of the ablation zone on contrast-enhanced computed tomography after radiofrequency ablation of hepatic metastases. *Indian J Radiol Imaging*. 2018;28(3):320-6.
25. Ridouani F, Ghosn M, Cornelis F, Petre EN, Hsu M, Moskowitz CS, et al. Ablation Zone Involution of Liver Tumors Is Faster in Patients Treated with Irreversible Electroporation Than Microwave Ablation. *Medicina (Kaunas)*. 2021;57(9).
26. Lubner MG, Brace CL, Hinshaw JL, Lee FT, Jr. Microwave tumor ablation: mechanism of action, clinical results, and devices. *J Vasc Interv Radiol*. 2010;21(8 Suppl):S192-203.
27. Brace CL. Radiofrequency and microwave ablation of the liver, lung, kidney, and bone: what are the differences? *Curr Probl Diagn Radiol*. 2009;38(3):135-43.
28. Deshazer G, Merck D, Hagemann M, Dupuy DE, Prakash P. Physical modeling of microwave ablation zone clinical margin variance. *Med Phys*. 2016;43(4):1764.
29. Servin F, Collins JA, Heiselman JS, Frederick-Dyer KC, Planz VB, Geevarghese SK, et al. Fat Quantification Imaging and Biophysical Modeling for Patient-Specific Forecasting of Microwave Ablation Therapy. *Front Physiol*. 2021;12:820251.
30. Heerink WJ, Solouki AM, Vliegenthart R, Ruiter SJS, Sieders E, Oudkerk M, et al. The relationship between applied energy and ablation zone volume in patients with hepatocellular carcinoma and colorectal liver metastasis. *European Radiology*. 2018;28(8):3228-36.
31. Amabile C, Ahmed M, Solbiati L, Meloni MF, Solbiati M, Cassarino S, et al. Microwave ablation of primary and secondary liver tumours: ex vivo, in vivo, and clinical characterisation. *International Journal of Hyperthermia*. 2017;33(1):34-42.
32. Hoffer EK, Borsic A, Patel SD. Validation of Software for Patient-Specific Real-Time Simulation of Hepatic Radiofrequency Ablation. *Acad Radiol*. 2022;29(10):e219-e27.
33. Moche M, Busse H, Futterer JJ, Hinestroza CA, Seider D, Brandmaier P, et al. Clinical evaluation of in silico planning and real-time simulation of hepatic radiofrequency ablation (ClinicIMPPACT Trial). *Eur Radiol*. 2020;30(2):934-42.
34. Mariappan P, Weir P, Flanagan R, Voglreiter P, Alhonnoro T, Pollari M, et al. GPU-based RFA simulation for minimally invasive cancer treatment of liver tumours. *Int J Comput Assist Radiol Surg*. 2017;12(1):59-68.
35. Voglreiter P, Mariappan P, Pollari M, Flanagan R, Blanco Sequeiros R, Portugaller RH, et al. RFA Guardian: Comprehensive Simulation of Radiofrequency Ablation Treatment of Liver Tumors. *Sci Rep*. 2018;8(1):787.
36. Audigier C, Mansi T, Delingette H, Rapaka S, Mihalef V, Carnegie D, et al. Efficient Lattice Boltzmann Solver for Patient-Specific Radiofrequency Ablation of Hepatic Tumors. *Ieee Transactions on Medical Imaging*. 2015;34(7):1576-89.
37. Müller D, Soto-Rey I, Kramer F. Towards a guideline for evaluation metrics in medical image segmentation. *BMC Res Notes*. 2022;15(1):210.

38. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging*. 2015;15:29.
39. Farina L, Weiss N, Nissenbaum Y, Cavagnaro M, Lopresto V, Pinto R, et al. Characterisation of tissue shrinkage during microwave thermal ablation. *Int J Hyperthermia*. 2014;30(7):419-28.
40. Amabile C, Farina L, Lopresto V, Pinto R, Cassarino S, Tosoratti N, et al. Tissue shrinkage in microwave ablation of liver: an ex vivo predictive model. *Int J Hyperthermia*. 2017;33(1):101-9.
41. Liu D, Brace CL. Evaluation of tissue deformation during radiofrequency and microwave ablation procedures: Influence of output energy delivery. *Med Phys*. 2019;46(9):4127-34.
42. Winokur RS, Du JY, Pua BB, Talenfeld AD, Sista AK, Schiffman MA, et al. Characterization of in vivo ablation zones following percutaneous microwave ablation of the liver with two commercially available devices: are manufacturer published reference values useful? *J Vasc Interv Radiol*. 2014;25(12):1939-46.e1.
43. Hendriks P, Berkhout WEM, Kaanen CI, Sluijter JH, Visser IJ, van den Dobbelsteen JJ, et al. Performance of the Emprint and Amica Microwave Ablation Systems in ex vivo Porcine Livers: Sphericity and Reproducibility Versus Size. *Cardiovasc Intervent Radiol*. 2021;44(6):952-8.
44. Lee DH, Lee JM, Yoon JH, Kim YJ, Han JK. Thermal Injury-induced Hepatic Parenchymal Hypoperfusion: Risk of Hepatocellular Carcinoma Recurrence after Radiofrequency Ablation. *Radiology*. 2017;282(3):880-91.

## Appendix A: Supplementary results

*Table 1: Needle reconstruction details per case*

Case	System used	CT phase used for reconstruction. Arterial (A) or portal-venous (PV)	Predicted ablation zone extending liver contour?	Realized ablation zone extending liver contour?	Reference entry point
1	Emprint	PV	No	No	No reference
2	Emprint	PV	No	Yes	Subcutaneous contamination
3	Emprint	A	No	Minimal	Subcapsular hemorrhage
4	Emprint	PV	Yes	Yes	Subcutaneous contamination
5	Cool-tip	PV	Yes	Yes	No reference
6	Emprint	PV	No	Minimal	Subcutaneous contamination
7	Emprint	PV	No	Yes	Subcutaneous contamination
8	Emprint	PV	Yes	Yes	Subcutaneous contamination
9	Emprint	PV	Yes	Yes	Needle CT
10	Amica	PV	Yes	Yes	No reference
11	Emprint	PV	Yes	Yes	No reference
12	Emprint HP	A	Yes	Yes	Subcutaneous contamination
13	Emprint HP	A (CTHA)	No	Minimal	Needle CT
14	Emprint HP	A	Yes	Yes	Subcutaneous contamination
15	Emprint HP	A	Yes	Yes	Subcutaneous contamination
16	Cool-tip	A	Yes	Yes	Needle CT
17	Emprint HP	A (CTHA)	Yes	Yes	Needle CT
18	Emprint HP	A	Yes	Yes	Needle CT
19	Emprint HP	A	No	Yes	Needle CT
20	Emprint HP	A	No	No	Subcutaneous contamination
21	Emprint HP	A	No	Yes	Subcapsular hemorrhage
22	Emprint HP	A	Yes	Yes	Needle CT
23	Emprint HP	A	Yes	Yes	Needle CT

CT = Computed tomography; HP = High power; CTHA = Computed tomography hepatic angiography

Table 2: Raw data of the realized and manufacturers' predicted SAD (mm), LAD (mm), volume (cm<sup>3</sup>), DSC and Average surface deviation (mm) per case

Case	SAD (mm)			LAD (mm)		Volume (cm <sup>3</sup> )		DSC	Average surface deviation (mm)
	Predicted	Realized SAD 1	Realized SAD 2	Predicted	Realized	Predicted <sup>†</sup>	Realized		
1	30	29.4	31.6	34	46.1	15.2	17.8	0.74	1.73
3	34	32.1	33.4	39	39.8	22.6	14.7	0.63	3.99
4	34	23.1	32.3	39	46.5	13.2	14.5	0.81	2.00
5	31	23.0	26.9	37	43.3	16.4	8.9	0.63	4.13
6	35	32.1	42.3	40	47.5	24.6	25.7	0.83	2.04
7	35	37.2	46.1	40	54.7	23.1	27.0	0.82	2.20
8	35	36.6	45.7	40	54.3	22.3	37.4	0.74	2.98
10	40	18.4	39.9	53	49.8	31.6	13.0	0.56	7.36
11	35	22.0	34.6	40	39.7	20.2	12.3	0.81	3.53
12	40	41.6	44.5	48	57.0	36.3	40.9	0.88	1.41
13	38	38.1	50.1	46	63.7	33.3	32.3	0.64	3.45
14	35	41.1	43.4	40	50.6	21.4	31.4	0.77	2.63
15	40	35.6	38.7	48	51.9	37.3	29.2	0.81	2.60
16	31	27.4	36.2	37	56.4	15.7	15.7	0.79	1.78
17	33	43.0	49.3	36	59.0	19.7	35.8	0.69	2.73
18	34	23.5	38.0	39	48.1	15.3	14.9	0.81	2.09
19	34	30.4	36.5	39	57.8	21.5	18.3	0.70	3.03
20	39	27.7	31.9	47	41.6	36.0	16.1	0.61	4.55
21	40	47.8	62.7	48	75.1	38.7	70.0	0.69	3.34
22	39	42.6	46.0	47	53.9	29.4	43.1	0.64	3.80
23	40	42.9	52.0	48	61.4	31.7	46.7	0.80	2.56

<sup>†</sup> Volume of manufactures predicted ablation zone within liver contours

SAD = short axis diameter; LAD = Long axis diameter; DSC = Dice similarity coefficient

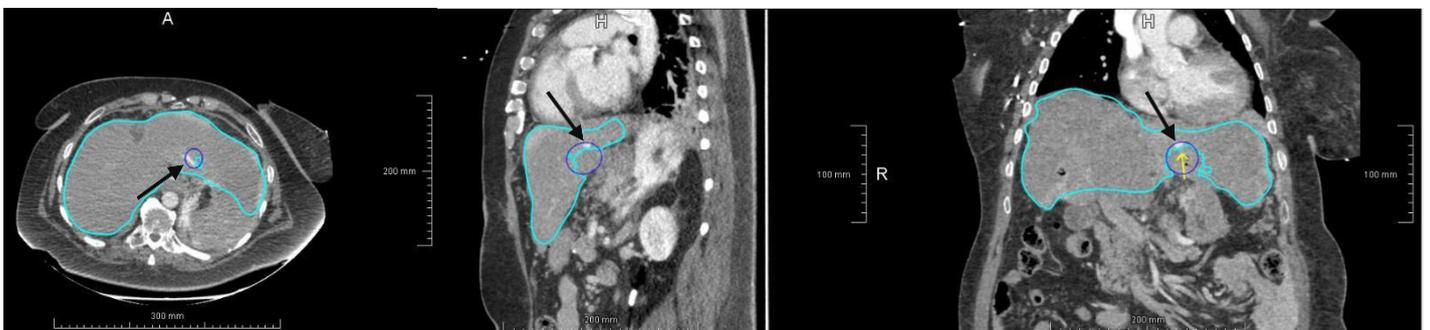


Figure 1. Axial, sagittal and coronal view of arterial phase post-ablation scan with segmented liver and ablation zone (cyan), reconstructed needle (yellow), predicted ablation zone (blue). Black arrow points out the artery influencing realized ablation zone. Emprint HP ablation with a DSC of 0.61 and average surface deviation of 3.0 mm in which the effect of a blood vessel on the realized ablation zone can be visually observed.

# Computational modeling of thermal ablation zones in the liver: a systematic review

## Abstract

**Purpose** This systematic review aims to give an overview of existing ablation zone computational models and compare their predictive capabilities.

**Methods** A systematic literature search was performed in the MEDLINE and Web of Science databases. Characteristics about the computational model and validation method of the included articles were retrieved.

**Results** The literature search identified 780 articles of which 38 are included. 22 articles focused on simulating radiofrequency ablation (RFA) zones and sixteen on microwave ablation (MWA) zones. 29 of the articles used the Pennes' bioheat equation as bioheat model, while the other nine articles made use of more advanced bioheat models. Out of the sixteen articles simulating MWA, only two used in vivo experiments to validate their simulations. Regarding the RFA simulations, nine of the 22 articles used in vivo validation. The Dice similarity coefficient between the ablation zone of in vivo experiments and RFA simulated ablation zones varied between 0.418 and 0.728 with mean surface deviations varying between 2.5 mm and 8.67 mm.

**Conclusion** Computational models to simulate ablation zones in MWA and RFA show considerable heterogeneity in model type and validation methods. It is currently unknown which model is most accurate and best suitable for use in clinical practice.

## 1. Introduction

Percutaneous thermal ablation is an established minimal invasive treatment for primary and secondary liver tumors [1, 2]. Radiofrequency ablation (RFA) and microwave ablation (MWA) are currently the most widely applied thermal ablation techniques to treat liver malignancies. Both techniques aim to induce tissue heating of at least 55-60°C to necrotize the tumor along with a clinical safety margin of minimal 5 mm [3, 4]. RFA is based on a rapidly alternating current which excites the ions in the liver tissue causing frictional heating. In MWA, electromagnetic waves cause polar molecules, predominantly water, to realign with the oscillating field, which generates heat through kinetic energy [5].

Thermal ablation has several advantages over surgical resection. It is less invasive, associated with shorter duration of hospitalization and has lower complication rates [6, 7]. Nevertheless, reported local recurrence rates (LRR) after thermal ablation tend to be higher than after surgical resection [6, 7]. Laimer et al. found the LRR to be associated with ablation margins: for each millimeter increase in minimal ablation margin, a 30% risk reduction for local recurrence was seen [4]. No recurrences occurred if a clinical safety margin > 5 mm was obtained, but this was only achieved in 37.5% of the ablations. These results are in accordance with several other studies that investigated the correlation between ablation margin and local recurrences [3, 4, 8, 9]. In these studies, the percentage of patients in whom the intended safety margin of > 5mm was achieved varied between 2.7% and 51.4%.

Currently, ablation zone size is predicted based on manufacturer's specifications concerning ablation time and power, mostly based on preclinical animal experiments [10]. However, computational modeling and ex vivo experiments demonstrate that tumor- and liver characteristics like tumor location, vascular proximity, liver cirrhosis and hepatic steatosis affect the heat conductivity and thus the ablation zone [11-14]. Therefore, patient-specific therapy planning of ablation zone, including aforementioned parameters, may allow greater safety margins to be obtained and reduce the risk of local recurrence.

Computational modeling of temperature distribution has already been used to simulate the ablation zones [15, 16]. Most of these models use the Pennes' bioheat equation, but they differ much in complexity [15-17]. This systematic review aims to give an overview of existing ablation zone computational models and compare their predictive capabilities.

## 2. Methods

### *2.1 Search strategy*

Studies were identified by searching the electronic databases MEDLINE and Web of Science on March 28, 2022. The search queries were based on synonyms of the keywords "Thermal ablation", "Liver neoplasm" and "Computational modeling". The complete search strategies used can be found in Appendix A.

### *2.2 Study selection*

After duplicate removal, abstracts were screened, followed by full-text assessment. Articles were found eligible if i) ablation zone simulation was performed based on either ii) percutaneous RFA or MWA in iii) liver tissue and if iv) the model was quantitatively validated using either ex vivo or in vivo experiments with v) ablation zone dimensions reported as outcome measure(s). Reviews, systematic reviews, letters to the editor, conference abstracts, articles written in other languages than English and articles with no full-text availability were excluded. The abstract screening and full-text assessment were performed by the first author (G.C.M. van Erp). In case of doubt, articles were discussed with a second assessor.

### *2.3 Data analysis*

Articles were sorted by thermal ablation technique used and characteristics about the computational model and validation method were retrieved. Regarding the computational model, the biological heat transfer model and the cell death model used were extracted. Furthermore, for each model it was noted whether perfusion, large blood vessels, water vaporization, temperature dependent thermal parameters and/or an imaging-based anatomical model were incorporated.

Regarding the validation of the computational model, the ex vivo or in vivo details, number of ablations, ablation settings, ground truth comparison, outcome measure and validation results were collected. Outcome measures were reported as homogeneous as possible. If the outcome measure included volumes or diameters of the simulated ablation zone and experimentally obtained ablation zone, the relative volume deviation (RVD) or relative diameter deviation (RDD) between the simulation and the actual ablation zone was calculated.

## 3. Results

### *3.1 Study selection*

The search strategy identified 780 articles after removal of duplicates. 618 articles were excluded after abstract screening. A total of 38 articles met the inclusion criteria and were included in this systematic review [18-55]. Figure 1 shows a flow diagram of the study selection process.

### *3.2 Data analysis*

Out of the 38 articles included, 22 focused on simulating RFA and sixteen on MWA. Table 1 and 2 contains details about the computational model used in the articles for MWA and RFA respectively.

#### *MWA computational models*

Sixteen articles presented computational models for MWA. Thirteen articles used the Pennes' bioheat equation as bioheat model [21, 24, 25, 29-31, 35, 42, 43, 45, 51, 53, 54], while one article used the

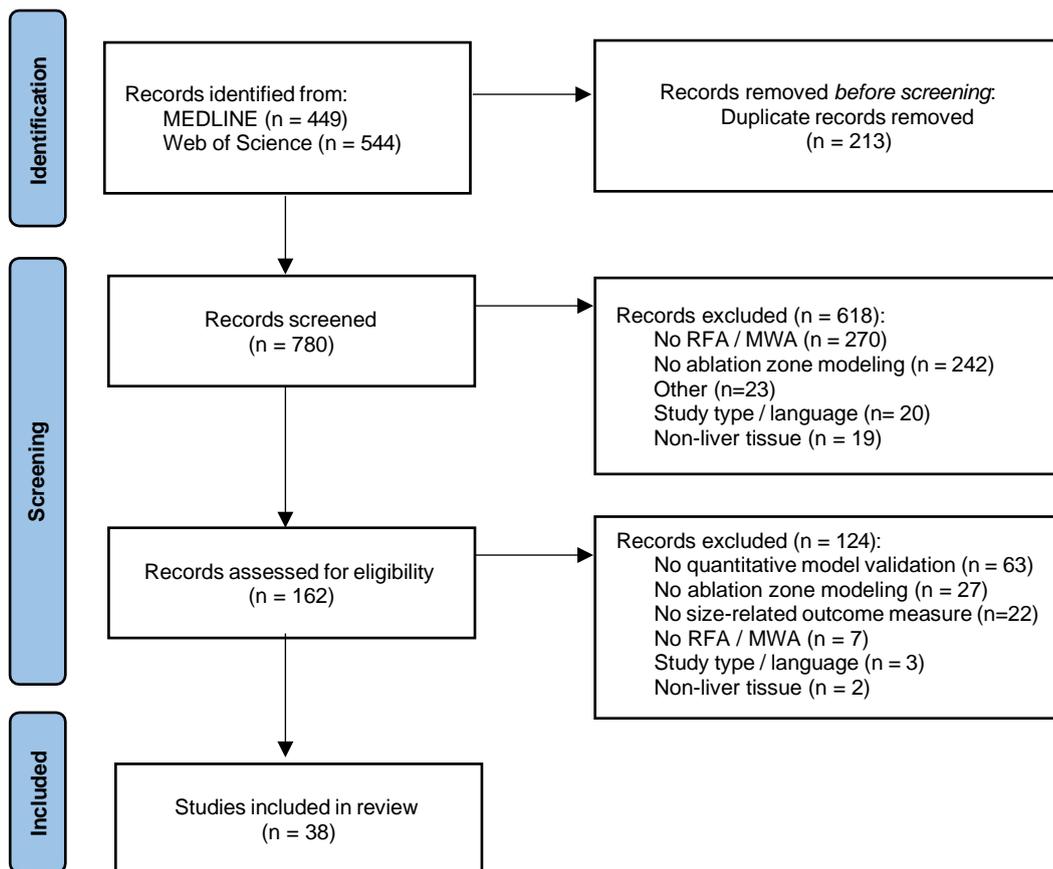


Figure 1: Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) flow diagram describing the study selection process.

transient heat transfer equation [28], one the local thermal non-equilibrium equation [47] and one used an own heat transfer model [26]. The Arrhenius thermal damage model (n=5), three-state cell death equation (n=2) or an isothermal contour (54°C or 60°C) (n=8) were used as cell death model. Deshazer et al. looked at both the Arrhenius thermal damage model and 52°C isothermal contour [25]. Six articles included perfusion in their model, two articles large blood vessels, thirteen water vaporization and thirteen temperature dependent tissue parameters. Two models used CT-based anatomy models. Gao et al. used CT-data to extract tumor geometry in order to model tumor coverage, while Zhai et al. created a complete CT-based 3D model for simulating the ablation [30, 54].

#### *RFA computational models*

22 articles presented computational models for RFA. Sixteen articles used the Pennes' bioheat equation as bioheat model [22, 23, 27, 32-34, 36-40, 44, 48-50, 55], while one article used the heat transfer equation [52], one the split volume bioheat equation [41], one article compared three different bioheat-models, ea. the Pennes' bioheat equation, the local thermal equilibrium equation and the local thermal non-equilibrium equation [46]. Three articles of Audigier et al. used a combination of the Pennes' bioheat equation with the Wulff-Klinger model [18-20]. The Arrhenius thermal damage model (n=7), three-state cell death equation (n=7) or an isothermal contour (n=7) were used as cell death model, while Subramanian et al. used their own thermal damage formula [44]. Fourteen articles included perfusion in their model, eleven articles large blood vessels, seven water vaporization and ten temperature dependent tissue parameters. Eight articles created a CT-based anatomical model for their simulation. Seven of them segmented the liver, tumor and blood vessels, while Ooi et al. only derived a liver contour from the CT-scan [40]. Next to an anatomical model, Moche et al. used dynamic CT measurements to derive perfusion values [37].

Table 1: Characteristics of the used computational models for Microwave ablation zone modeling from the included articles. (x=incorporated in the model)

Author (year)	Bioheat model	Cell death model	Numerical method*	Perfusion	Large blood vessels	Water vaporization	Temperature dependent tissue parameters	CT-based anatomic model	Model remarks	
Cavagnaro et al.[21] (2015)	Pennes' bioheat equation	60°C isothermal contour	FDTD						BHE-S	
	Pennes' bioheat equation	60°C isothermal contour				x			BHE-V	
	Pennes' bioheat equation	60°C isothermal contour						x		BHE-ST_B & BHE-ST (two different equations for temperature dependent parameters)
	Pennes' bioheat equation	60°C isothermal contour				x		x		BHE-V-ST_B & BHE-V-ST. (two different equations for temperature dependent parameters), Only conductivity is temperature-dependent (not the di-electric parameters)
Collins et al.[24] (2020)	Pennes' bioheat equation	60°C isothermal contour	FEM			x			SAR-T-1min_B & SAR-T-1min (two different equations for temperature dependent parameters). Temperature-dependency of conductivity as well as dielectric parameters	
	Pennes' bioheat equation	60°C isothermal contour					x	x		Determine dielectric properties based on MRI fat quantification with inverse modeling strategy
Deshazer et al. [25] (2017)	Pennes' bioheat equation	Arrhenius thermal damage model (isocontour 63%) and 52°C isothermal contour	FEM	x, but not in experiments		x		x	Damage-dependent blood perfusion rate. Two different models tested (A & B), they only differ in dielectric parameter dependency of temperature.	
Deshazer et al. [26] (2017)	Own heat transfer model	60°C isothermal contour	FEM			x		x	Investigated the option of intra-procedural SAR measurement to model ablation zone.	
Faridi et al. [28] (2020)	Transient heat transfer equation	Arrhenius thermal damage model (isocontour 63%)	FEM			x		x	Added the Morris method to determine the sensitivity of the ablation zones to uncertainty in tissue physical properties	

Gao et al. [30] (2019)	Pennes' bioheat equation	54°C isothermal contour	FEM			x	x	x	Tried to model coagulation zone over time and incorporate tumor geometry to assess tumor coverage
Gao et al. [31] (2019)	Pennes' bioheat equation	54°C isothermal contour	FEM			x	x		Used parameter sensitivity analysis to optimize the temperature-based parameters
Gao et al. [29] (2017)	Pennes' bioheat equation	54°C isothermal contour	FEM						Used experiments to determine phantom parameters and SAR distribution, which is the basis of the FEM model
Lopresto et al. [35] (2017)	Pennes' bioheat equation	60°C isothermal contour	FDTD			x	x		Evaluate the effect of ±25% variations in dielectric and thermal parameters using the combined expanded uncertainty
Radosevic et al. [42] (2021)	Pennes' bioheat equation	Arrhenius thermal damage model (isocontour 99%)	FEM			x	x		Simulated continuous and pulsed power.
Singh et al. [43] (2019)	Pennes' bioheat equation with Dual phase lag model	Three-state cell death model	FEM	x, but not in experiments		x	x		Incorporates lot of complexities: damage-dependent blood perfusion rate, mechanical deformation (shrinkage) and Heat-flux model. Modelled RFA as well as MWA, however only validated MWA with experiments
Tehrani et al. [45] (2010)	Pennes' bioheat equation	Three-state cell death model	FEM	x		x			Used a multicompartment model including tissue, tumor and blood. Added a model for tumor shrinkage
Tucci et al. [47] (2022)	Local thermal non-equilibrium equation	Arrhenius thermal damage model (isocontour 99%)	FEM	x	x, 4 different diameters	x	x		Damage-dependent blood perfusion rate. Two compartment model with difference in porosity (and other factors) in tumor and surrounding liver tissue. Within the tumor they modelled the difference of porosity in the tumor core towards the tumor rim (increasing porosity)
Wang et al. [51] (2021)	Pennes' bioheat equation	54°C isothermal contour	FEM	x	x	x	x		Incorporated convection heat transfer condition and Newton formula for heat transfer between blood vessel and tissue

Wu et al. [53] (2013)	Pennes' bioheat equation	55°C isothermal contour	FDTD			x		x		Used GPUs to simulate in 3D. Did not quantify the Electrical field, but determined its contribution based on experiments.
Zhai et al. [54] (2008)	Pennes' bioheat equation	Arrhenius thermal damage model (isocontour 63%)	FEM	x				x	x	GPU-accelerated model for preoperative 3D simulation of necrotic zone in clinical setting. Incorporates effect of necrosis on blood perfusion

\* FEM = Finite Element Method, FDTD = Finite Difference Time Domain, FVM = Finite Volume Method

SAR = Specific absorption rate, CT = Computed Tomography, GPU = Graphics processing unit, MWA = microwave ablation

Table 2: Characteristics of the used computational models for Radiofrequency ablation zone modeling from the included articles. (x=incorporated in the model)

Author (year)	Bioheat model	Cell death model	Numerical method*	Perfusion	Large blood vessels	Water vaporization	Temperature dependent tissue parameters	CT-based anatomic model	Model remarks
Audigier et al. [19] (2015)	Combination of Pennes' bioheat equation and Wulff-Klinger model	Three-state cell death model	Lattice Boltzmann solver	x	x			x	Computational fluid dynamics and Darcy's equation are coupled tot the bioheat equation to model blood circulation and blood flow, two-compartment model (blood vessels and liver tissue)
Audigier et al. [18] (2013)	Combination of Pennes' bioheat equation and Wulff-Klinger model	Three-state cell death model	Lattice Boltzmann solver	x	x			x	Computational fluid dynamics and Darcy's equation are coupled tot the bioheat equation to model blood circulation and blood flow
Audigier et al. [20] (2017)	Combination of Pennes' bioheat equation and Wulff-Klinger model	Three-state cell death model	Lattice Boltzmann solver	x	x			x	Navier-stokes equation and computational fluid dynamics solver used to model blood flow. Blood flow determined using preoperative MRI, blood pressures are measured invasively, porosity map created on CT-image. Besides used intra-operative measurements to validate parameter values used.

Chang et al. [22] (2004)	Pennes' bioheat equation	Arrhenius thermal damage model (isocontour 63%)	FEM	x, but not in experiments		x	Damage-dependent blood perfusion rate
Chen et al. [23] (2021)	Simplified Pennes' bioheat equation	55°C isothermal contour	Simplified towards analytical solution				Ignored the heat source of the electrical current flow in the model.
Duan et al. [27] (2016)	Pennes' bioheat equation	Arrhenius thermal damage model (isocontour 63%)	FEM	x, but not in experiments		x	Using a pre-procedural determined needle position, the probability of several ablation zones is displayed by the model. Damage-dependent blood perfusion rate
Haemmerich et al. [32] (2001)	Pennes' bioheat equation	50°C isothermal contour	FEM	x			Modelled monopolar and bipolar use of the probe(s)
Haemmerich et al. [33] (2010)	Pennes' bioheat equation	50°C isothermal contour	FEM		x	x	
Lim et al. [34] (2010)	Pennes' bioheat equation	47°C and 64°C isothermal contour	FEM	x, but not in experiments	x		
Macchi et al. [36] (2014)	Pennes' bioheat equation	50°C isothermal contour	FVM		x	x	Used several experiments to generate different electrical conductivity models to determine the dielectric parameters.
Mariappan et al. [38] (2017)	Pennes' bioheat equation	Three-state cell death model	FEM	x	x		Used a GPU to accelerate FEM, focused on clinical application
Moche et al. [37] (2020)	Pennes' bioheat equation	Three-state cell death model	FEM	x	x		Used a GPU, more focused on clinical application. Simulation parameters involved a proportional-integral-derivative
Nolte et al. [39] (2021)	Pennes' bioheat equation	52 °C isothermal contour	FVM		x		Modelled heat transfer inside tissue, needle and blood separate
Ooi et al. [40] (2019)	Pennes' bioheat equation	Arrhenius thermal damage model (isocontour 99%)	FEM	x, but not in experiments		x	Modelled different boundary conditions

Payne et al. [41] (2011)	Split-volume bioheat equation (own model)	Three-state cell death model	FEM	x	x			x	Incorporated Newton's cooling law to model heat transfer between vessels and tissue, and Darcy's law for blood velocity.
Subramanian et al.[44] (2015)	Pennes' bioheat equation	Own thermal damage formula	FEM					x	Experimental-based values of the specific heat, thermal conductivity and electrical conductivity
	Pennes' bioheat equation			x			x		Damage-dependent blood perfusion rate
Tucci et al. [46] (2021)	Local thermal equilibrium equation (LTE)	Arrhenius thermal damage model (isocontour 99%)	FEM	x			x		Porous media-based model, damage-dependent blood perfusion rate, assumes equilibrium in temperature between blood and tissue
	Local thermal non-equilibrium equation (LTNE)			x			x		Porous media-based model, damage-dependent blood perfusion rate, separates vaporization phase for water, tissue and blood
Vaidya et al. [48] (2021)	Pennes' bioheat equation	Arrhenius thermal damage model	FVM		x	x		x	Multicompartment model incorporating tissue, tumor, blood and needle. Damage-dependent blood perfusion rate.
Voglreiter et al. [49] (2018)	Pennes' bioheat equation	Three-state cell death model	FEM	x	x			x	Used a GPU to accelerate FEM, focused on clinical application
Wang et al. [50] (2019)	Pennes' bioheat equation	54 °C isothermal contour	FEM					x	
Welp et al. [52] (2006)	Heat transfer equation	Arrhenius thermal damage model (isocontour 99%)	FEM		x	x		x	Incorporated the heat transfer between blood and tissue
Zhang et al. [55] (2015)	Pennes' bioheat equation	Arrhenius thermal damage model (isocontour 63%)	FEM	x			x	x	Modelled Pulsed RFA as well as constant RFA (used in validation). Damage-dependent blood perfusion rate. Used a two-compartment model of tumor and tissue

\* FEM = Finite Element Method, FDTD = Finite Difference Time Domain, FVM = Finite Volume Method  
CT = Computed Tomography, GPU = Graphics processing unit, RFA = Radiofrequency ablation

Scatterplot of the modelled ablation zone compared to the experiments for MWA

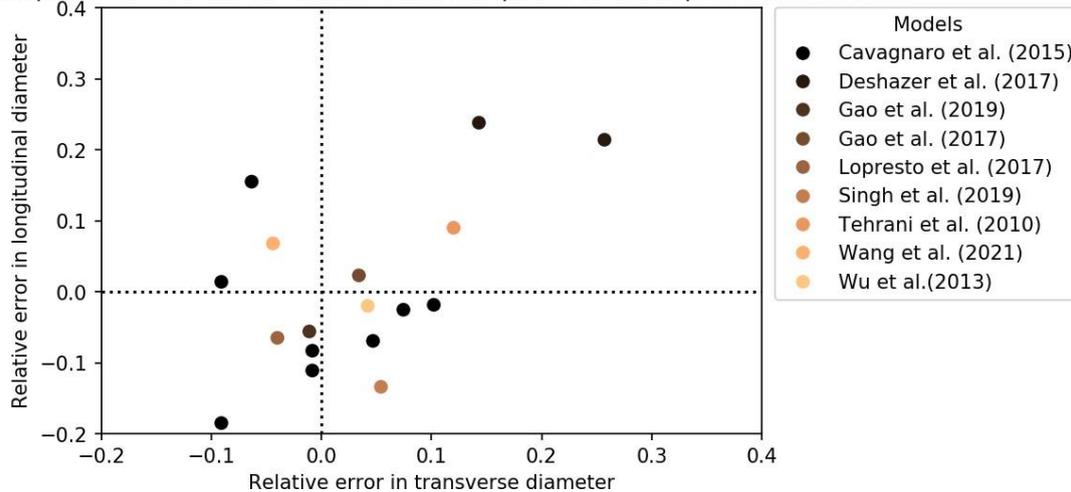


Figure 2. Scatterplot of the Relative error in longitudinal and transverse diameters of the modelled MWA zones compared to ex vivo validation [21, 25, 29, 31, 35, 43, 45, 51, 53]. In case of an experimental diameter of 30 mm, a relative error of 0.1 means the simulated diameter was 33 mm.

### MWA validation

Table 3 gives an overview of the predictive validation of the articles regarding MWA ablation zone simulation. Only two articles, Tucci et al. and Zhai et al. used in vivo validation [47, 54]. Tucci et al. modelled four different blood vessels and compared it to in vivo experiments of Amabile et al. [47, 56]. They conclude that their model including terminal arteries resembles well to the ablation zones achieved in the clinical study. Zhai et al. performed a study on nine patients [54]. Ablation simulation had a RVD of  $\pm 7.0\%$  to clinically obtained ablation volumes. However, only few details are known about the in vivo study, making it hard to conclude on these results. The other fourteen articles used ex vivo validation [21, 24-26, 28-31, 35, 42, 43, 45, 51, 53]. Two articles used the Dice similarity coefficient (DSC) to express their results and found similar scores between 0.74 and 0.82 [26, 28]. One article used the Jaccard similarity index, and found results of 86.6% and 93.4%, however these results might be biased, since the electrical and thermal conductivities were reconstructed to best fit the model after the experiments [24]. Sing et al. used the experiments of Wu et al. to validate their simulated ablation zone [43, 53]. The main difference between the two models was the use of the Three-state cell death model and incorporating tissue shrinkage within the model of Sing et al. The later one could explain why Sing et al. simulated a smaller longitudinal diameter (26.24 mm, RDD: -13.4%), compared to Wu et al. (29.7 mm, RDD = -2.0%). However, the transverse diameter of Sing et al. has a greater overestimation (RDD: 5.2% versus -4.7%). The better performance of the simulation of Wu et al. might result from bias, since the electrical field contribution in the simulation is altered based on an experiment-based coefficient. Figure 2 visually gives an overview of the models validated with the longitudinal and transverse RDD.

### RFA ex vivo validation

Table 4 contains the ex vivo validation of the RFA ablation zone models [22, 23, 27, 33, 34, 36, 39, 40, 44, 48, 50, 52, 55]. All experimental ablation zones are determined using sample section along the probe axis. However, the ablation settings for obtaining the ablation zone differ in all experiments. A visual overview of the longitudinal and transverse RDD is given in figure 3. Figure 4 contains a combined overview of the MWA and RFA ex vivo experiments.

### RFA in vivo validation

An overview of the nine RFA ablation zone simulation articles using in vivo validation is given in table 5. These consist of four in vivo animal experiments [20, 32, 41, 46], four retrospective clinical studies [18, 19, 38, 49] and one prospective clinical study [37]. The prospective study of Moche et al. found

a DSC of  $0.62 \pm 0.14$  with a surface deviation of  $3.4 \pm 1.7$  mm [37]. They concluded that the real-time simulation of RFA-induced tissue necrosis in the liver was fast ( $3.5 \pm 1.9$  min) and accurate enough for clinical demands. The results of the retrospective study of Mariappan et al., using the same computational model, are comparable [38]. They found a lower surface deviation of 2.50 mm with known CT-perfusion values and that the simulation accuracy increased by using patient-specific CT-based perfusion values. This study was a retrospective analysis though in only 23 ablations. The in vivo results of Audigier et al. are lower compared to the previous mentioned studies [18-20], reflecting a less accurate simulation. They found a DSC of 0.44 and surface deviation of  $5.3 \pm 3.6$  mm in their most recent study [20]. This difference in results might be explained by the reconstruction of the ablation probe location. Audigier et al. did not reconstruct the clinically used probe location, but assumed the center of the tumor as probe location in their simulation, which introduces an inaccuracy in the measurements. On the other hand, Moche et al. and Mariappan et al. reconstructed the used needle position using image registration, and their simulations are therefore based on the clinically used probe location [37, 38].

Scatterplot of the modelled ablation zone compared to the experiments for RFA

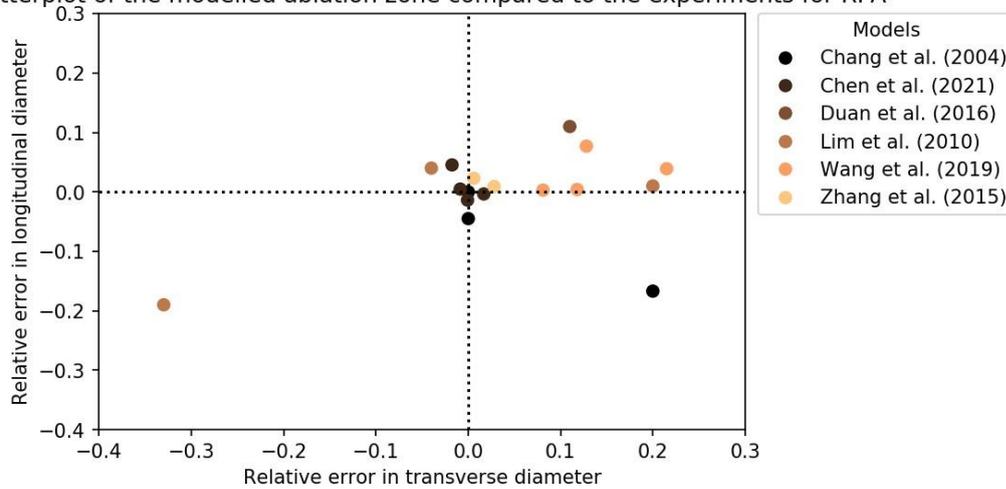


Figure 3. Scatterplot of the Relative error in longitudinal and transverse diameters of the modelled RFA zones compared to ex vivo validation [22, 23, 27, 34, 35, 43, 45, 50, 55]. In case of an experimental diameter of 30 mm, a relative error of 0.1 means the simulated diameter was 33 mm.

Scatterplot of the modelled ablation zone compared to the experiments for RFA & MWA

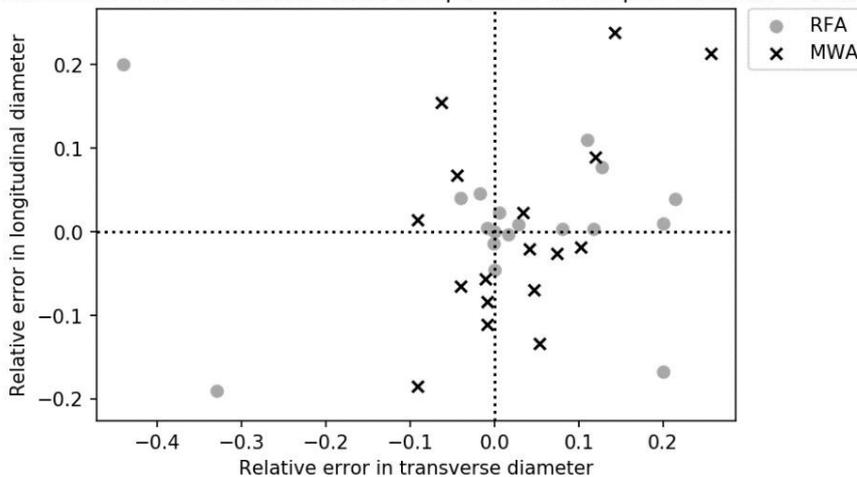


Figure 4. Scatterplot of the Relative error in longitudinal and transverse diameters of the modelled MWA & RFA zones compared to ex vivo validation. In case of an experimental diameter of 30 mm, a relative error of 0.1 means the simulated diameter was 33 mm.

Table 3: Experimental model validation of the articles regarding microwave ablation

Author (year)	Model	In vivo or ex vivo validation	Number of experiments	Ground truth	Ablation settings (time of ablation & power)	Outcome measure	Performance	Validation remarks
Cavagnaro et al. [21] (2015)	BHE-S	Ex vivo, bovine livers	6	Sectioning sample and measure ablation zone	10 min, 40 W, 2450 MHz	Longitudinal(L) and transverse(T) RDD*	L:-8.31% T: -0.83%	
	BHE-V						L:-18.5% T: -9.09%	
	BHE-ST_B						L:-1.85% T: 10.2%	
	BHE-ST						L:-2.54%, T:7.44%	
	BHE-V-ST_B						L:-6.93%, T:4.68%	
	BHE-V-ST						L:-11.1%, T:-0.83%	
	SAR-T-1min_B						L: 15.5%, T:-6.34%	
SAR-T-1min	L:1.39%, T:-9.09%							
Collins et al. [24] (2020)	Fat phantoms	Ex vivo, phantom	15	Sectioning sample, photographed and 2D segmentation of ablation zone	15 min, 60 W, 915 MHz,	Jaccard similarity index	86,6 +- 5.3%	For each phantom, the electrical and thermal conductivities were reconstructed to best fit the model.
	non-fat phantom		6			Jaccard similarity index	93.4 ± 2.2%	
Deshazer et al. [25](2017)	model A	Ex vivo, bovine livers	4	Sectioning sample and measure ablation zone	10 min, 30 W, 915 MHz	Longitudinal(L) and transverse(T) RDD*	L: 2.9%, T: 24.0%	A: linear temperature dependency of dielectric properties, B: similar to model A but added linear decrease in electrical conductivity above 95°C
	model B		4		10 min, 30 W, 915 MHz		L: 5.7%, T:12.0%	
	model A		8		15 min, 60 W, 915 MHz		L: 21.4%, T: 25.7%	
	model B		8		15 min, 60 W, 915 MHz		L:23.8%, T: 14.3%	
Deshazer et al. [26] (2017)	Short-tip, 1000 W/kg iso-SAR	Ex vivo, porcine livers	3	Segmentation on Infrared camera temperature measurements	6 min, 15 W, 915 MHz	DSC	0.74 ± 0.01	
	Short-tip, 500 W/kg iso-SAR						0.82 ± 0.04	
	Long-tip, 1000 W/kg iso-SAR		3				0.77 ± 0.03	
	Long-tip, 500 W/kilo iso-SAR						0.76 ±0.01	

Faridi et al. [28] (2020)		Ex vivo, bovine livers	4	Segmentation on MRT-derived Arrhenius thermal damage 3D maps.	10 min, 30 W, 2450 MHz	DSC	0.8 ± 0.0		
			8		5 min, 30 W, 2450 MHz		0.8 ± 0.08		
			3		5 min, 50 W, 2450 MHz		0.75 ± 0.06		
Gao et al. [30] (2019)		Ex vivo, porcine livers	20	Sectioning sample and measure ablation zone	40, 45, 50, 55 and 60W 2450 MHz		Error of transverse radius, advancement and backward longitudinal length	± 5%	
Gao et al. [31] (2019)		Ex vivo, porcine livers	20	Sectioning sample and measure ablation zone	6 min, 60 W, 2450 MHz		Longitudinal(L) and transverse(T) RDD*	L: 2.3%, T 3.4%	Optimized thermo-dependent parameters based on experiments
Gao et al. [29] (2017)		Ex vivo, phantom		Sectioning sample and measure ablation zone	10 min, 60W, 2450 MHz		Longitudinal(L) and transverse(T) RDD* Advancement	L: -5.6%, T:-1.1% 0.341 vs 0.3 ± 0.05 cm	
Lopresto et al. [35] (2017)		Ex vivo, bovine livers	4	Sectioning sample and measure ablation zone	10 min, 60 W 2,45 GHz		Longitudinal(L) and transverse(T) RDD* Advancement	L: -6.5%, T: -4.0% 7.4 mm (model) versus 7.5 ± 2.1 mm	
Radosevic et al. [42] (2021)	continuous	Ex vivo, bovine livers	24	Sectioning sample and measure ablation zone	2, 3, 4, 5 & 10 min, 60 & 80 W, 2450 MHz		Longitudinal(L) and transverse(T) RDD*	L: -5 to +13% T: +3 tot + 17%	
	pulsed power		12		2, 5 and 10 min, 60, 80, 100 and 125 W, 2450 MHz			L:-3% to + 14% T: +1% to +18%	
Singh et al. [43] (2019)		Ex vivo, porcine livers	10	Sectioning sample and measure ablation zone	2 min, 40 W, 2450 MHz		Longitudinal(L) and transverse(T) RDD*	L: -13.4% T: 5.4%	Used the experimental results of Wu et al. [53]
Tehrani et al. [45] (2010)		Ex vivo, porcine livers	56	Sectioning sample and measure ablation zone	10 min, 50 & 60 W, 2450 MHz & 80 W, 915 MHz		Longitudinal(L) and transverse(T) RDD*	L: 9%, T: 12%	Used the experimental results of Sun et al. [57]

Tucci et al. [47] (2022)	In vivo, patients	32	Segmentation on 24-h post-ablation CT	5 & 10 min (respectively), 60 W, 2450 MHz	Transverse RDD*	+24% (5 min) +43% (10 min)	Used the experimental results of Amabile et al.[56]	
					RVD*	31% (5 min), 93% (10 min)		
					Terminal arteries	Transverse RDD*		-4% (5 min) +8% (10 min)
					RVD*	-32% (5 min), -8% (10 min)		
					Terminal branches	Transverse RDD*		-42% (5 min), -43%(10 min)
					RVD*	-83% (5 min), -84% (10 min)		
Tertiary branches	Transverse RDD*	-18% (5 min), -13% (10 min)						
RVD*	-88% (5 min), -84% (10 min)							
Wang et al. [51] (2021)	Ex vivo, porcine liver	11	Sectioning sample and measure ablation zone	6 min, 60 W, 2450 MHz	Longitudinal(L) and transverse(T) RDD*	L: 6.8%, T: -4.4%	Used a peristaltic pump to simulate blood circulation and soft plastic tubes for blood vessels	
Wu et al. [53] (2013)	Ex vivo, porcine livers	10	Sectioning sample and measure ablation zone	2 min, 40 W, 2450 MHz	Longitudinal(L) and transverse(T) RDD*	L: -2.0%, T: 4.2%		
Zhai et al. [54] (2008)	In vivo, patients	9	Segmentation on 1-2 weeks post-ablation CT	Patient specific, 2450 MHz	RVD*	±7.0%	Article contains only small details on experiments. Study type unknown	

\* Relative differences are results of the computational model compared to the experiments

SAR = Specific absorption rate, DSC = Dice similarity coefficient, CT = Computed Tomography, MRT = Magnetic Resonance Thermometry, RDD = Relative diameter deviation, RVD = Relative volume deviation, Advancement = the distance from the antenna tip to the boundary of the ablated zone

Table 4: ex vivo validation of computational models modeling Radiofrequency ablation

Author (year)	Model	In vivo or ex vivo validation	Number of experiments	Ground truth	Ablation settings (time of ablation & power)	Outcome measure	Performance	Validation remarks
Chang et al. [22] (2004)		Ex vivo, porcine livers	2	Sectioning sample and placed in 2,3,5-triphenyltetrazolium chloride in order to color cell viability	15 minutes, 20 V	Longitudinal(L) and transverse(T) RDD*	L: 0.0%, T: 0.0%	
			2		15 minutes, 25 V		L: -16.7%, T: 20.0%	
			2		15 minutes, 30 V		L: -4.5%, T: 0.0%	
Chen et al. [23] (2021)	single probe	Ex vivo, porcine livers	5	Sectioning sample and measure ablation zone	12 min	Longitudinal(L) and transverse (T) RDD*  Longitudinal midline (Lm), longitudinal probe line (Lp) & transverse (T) RDD*	L: -0,35%, T: 1,68%	
	switching probe (10 mm)		5				Lm: -1,38%, Lp:-1,82%, T:-0,08%	
	switching probe (15 mm)		5				Lm: 0,47%, Lp: 0,05%, T: -0,87%	
	switching probe (20 mm)		5				Lm: 4,54%, Lp: 0,64%, T: -1,76%	
Duan et al. [27] (2016)		Ex vivo, porcine livers	20	Sectioning sample and measure ablation zone	5 min, temperature controlled (105°C)	Longitudinal(L) and transverse(T) RDD and relative area deviation (A)*	L: 11.1%, T:10.9%, A:1%	
Haemmerich et al. [33] (2010)		Ex vivo, porcine liver	12	Sectioning sample and measure ablation zone	12 min, impedance-controlled (50 V throughout ablation)	Transverse RDD*	13.8%	
Lim et al. [34] (2010)	T = 47°C	Ex vivo, bovine livers	6	Segmentation on converted grey-scale digital photographs of sectioned tissue	10 min, 25 V	Longitudinal(L) and transverse(T) RDD and relative area deviation (A)*	L: 4%, T: -4%, A: 14%	
	T = 64°C				10 min, 25 v		L: -20%, T: -44%, A: -44%	
	T = 47°C		6		15 min, 25 V		L: 1%, T: 20%, A: 5%	
	T = 64°C				15 min, 25 V		L: -19%, T: -33%, A: -38%	

Macchi et al. [36] (2014)	FH-NL	Ex vivo, porcine livers			constant power (10, 15 W), 480 kHz	Transverse lesion diameter	24 vs 18-26 mm	Fast (FH) and slow heating (SH) experimental-based model, two Arrhenius-based electrical conductivity models $\sigma(\vartheta, t)$ and a constant model. Compared to experiments of Gallati and Braschi et al.[58] non-English so experimental details are not complete.
	SH-NL						25.5 vs 18 - 26 mm	
	$\sigma(\vartheta, t)_1$						25 vs 18-26mm	
	$\sigma(\vartheta, t)_2$						24.5 vs 18-26 mm	
	constant model						17 vs 18-26 mm	
Nolte et al. [39] (2021)		Ex vivo, phantoms	12	Segmentation on photograph of sectioned tissue	10 min, Temperature-controlled (103°C) max power 35 W	Mean absolute error in area $R^2$ Pearson correlation coefficient	0.12 cm <sup>3</sup> 0.12 0.66	Used ink which colors irreversibly above threshold $T > 70^\circ\text{C}$ . Used a peristaltic pump and rods to simulate blood vessels
Ooi et al. [40] (2019)		Ex vivo, bovine livers	3	Sectioning sample and measure ablation zone	12 min, Impedance-controlled, 1800 mA	Transverse RDD*	-20.9%	Used the experimental results of Goldberg et al. [59]
Subramanian et al. [44] (2015)		Ex vivo, bovine livers	15	Segmentation on image of flatbed scanner after sectioning sample	500 KHz, 1-6 min, 31-34 V 60-80W	Relative area deviation*	-2.63%	Optimized tissue parameters based on experiments.
Vaidya et al. [48] (2021)		Ex vivo, phantom	1	Sectioning phantom, using temperature-sensitive ink to measure ablation zone	10 min, Temperature-controlled (103°C) max power 35 W	Relative area deviation*	17.03%	Used ink which colors irreversibly above threshold $T > 70^\circ\text{C}$

Wang et al. [50] (2019)	Ex vivo, porcine livers	3	Sectioning sample and measure ablation zone	Temperature- controlled (80°C), 330 kHz,	Longitudinal(L) and transverse(T) RDD*	L: 7.7%, T: 12.8%	Used a peristaltic pump to simulate blood circulation and soft plastic tubes for simulating blood vessels	
		3		Temperature- controlled (85°C), 330 kHz		L: 3.9%, T: 21.5%		
		3		Temperature- controlled (90°C), 330 kHz		L: 0.4%, T: 11.8%		
		3		Temperature- controlled (95°C), 330 kHz		L: 0.3%, T: 8.1%		
Welp et al. [52] (2006)	Ex vivo	10	Sectioning sample and measure ablation zone	12 min, impedance- controlled, 25 W	Transverse RDD*	Vessel $\varnothing$ = 4 mm, flow 25 ml/min	-5.7%	Used glass tubes to simulate blood vessels
						Vessel $\varnothing$ = 4 mm, flow 50 ml/min	-2.4%	
						Vessel $\varnothing$ = 4 mm, flow 75 ml/min	-8.7%	
						Vessel $\varnothing$ = 6 mm, flow 75 ml/min	1.9%	
						Vessel $\varnothing$ = 6 mm, flow 150 ml/min	1.9%	
Vessel $\varnothing$ = 6 mm, flow 300 ml/min	1.9%							
Zhang et al. [55] (2015)	In vitro, porcine livers	5	Sectioning sample and measure ablation zone	6 min, 30 V	Longitudinal(L) and transverse(T) RDD and relative area deviation (A)*	L: 0.9%, T: 2.8%, A:10.2%		
		5		12 min, 30V		L:2.3%, T:0.6%, A:8.7%		

\* Relative differences are results of the computational model compared to the experiments

$\varnothing$  = diameter, DSC = Dice similarity coefficient, CT = Computed Tomography, RDD = Relative diameter deviation, RVD = Relative volume deviation

Table 5: In vivo validation of computational models modeling Radiofrequency ablation

Author (year)	Model	In vivo or ex vivo validation	Number of experiments	Ground truth	Ablation settings (time of ablation & power)	Outcome measure	Performance	Validation remarks
Audigier et al. [19] (2015)		In vivo, patients	10 patients, 14 tumors	Segmentation on post-ablation CT-scan	Patient-Specific	DSC Sensitivity PPV	0.418 66.94% 38.30%	Retrospective study
Audigier et al. [18] (2013)		In vivo, patients	5 patients, 7 ablations	Segmentation on post-ablation CT-scan	Patient-specific	Surface deviation	8.67 mm	Retrospective study
Audigier et al. [20] (2017)		In vivo, porcine	5 pigs, 12 ablations	Segmentation on post-ablation CT-scan	6 min, Temperature controlled (105°C), two iterations for large tumors	Surface deviation DSC Sensitivity PPV	5.3 ± 3.6 mm 0.44 47% 53%	Surrogate tumors implanted,
Haemmerich et al. [32] (2001)	monopolar bipolar	In vivo , domestic pigs	3	Liver was cut into slices of 3-5 mm thickness and imaged. Segmentation took place on images	12 min, temperature controlled (95°C)	RVD	103.8% 54.0%	
Mariappan et al. [38] (2017)	unknown CT perfusion values known CT-perfusion values	In vivo, patients	6 patients, 10 ablations 12 patients, 23 ablations	Segmentation on 1 month post-ablation CT-scan	Patient-Specific, temperature controlled	DSC RVD Surface deviation DSC RVD Surface deviation	0.7286 5.11% 2.55 mm 0.691 17.93% 2.50 mm	Retrospective study
Moche et al. [37] (2020)		In vivo, patients	46 patients, 51 ablations	Segmentation on 1 month post-ablation CT-scan	Patient-Specific, temperature controlled	DSC Sensitivity PPV Surface deviation	0.62 ± 0.14 0.70 ± 0.21 0.66 ± 0.25 3.4 ± 1.7 mm	Prospective study
Payne et al. [41] (2011)		In vivo, pigs	2	Segmentation on post-ablation CT-scan	temperature controlled	RVD	39.6%	

Tucci et al. [46] (2021)	Pennes' LTE LTNE	In vivo, pigs	8	Sectioning sample and measure ablation zone	12 min, 90V, 500 KHz, impedance controlled.	Transverse RDD*	-32.4% -7.57% -7.57%	Compared to experiments of Goldberg et al.
Vogltreiter et al. [49] (2018)		In vivo, patients	21	Segmentation on post-ablation CT-scan	Patient specific	DSC RVD Sensitivity PPV Surface deviation	0.7003 ± 0.0937 13.77 ± 12.96% 69.70 ± 10.94% 71.73 ± 12.00% 2.44 ± 0.84 mm	Retrospective study

\* Relative differences are results of the computational model compared to the experiments

DSC = Dice similarity coefficient, PPV = Positive predictive value, RDD = Relative diameter deviation, RVD = Relative volume deviation,

## 4. Discussion

This systematic review gives an overview of existing computational models for ablation zone simulation and their predictive capabilities. A high variety within computation modeling and validation methods has been found, making a full-fledged comparison hard to perform.

Since the heating mechanism is substantially different among RFA and MWA, mechanism-based differences can be identified in the models. Twelve of the sixteen (75%) MWA models included the effect of water vaporization and temperature-dependent parameters. This is more than in the RFA models, where only five of the 22 included articles modelled both effects (22.7%). The temperature-dependency of dielectric parameters (electrical conductivity and permittivity) and thermal conductivity is two-folded. One aspect is protein denaturation at temperatures above 60°C. The protein denaturation causes the tissue parameters to change [60-62]. The other factor is water vaporization, which changes the liver tissue water content and creates gas, which has an isolating effect. This leads to a drop in conductivity and permittivity above 100°C. This effect is observed in RFA as well as in MWA [60-62]. However, RFA often makes use of temperature-controlled or impedance-controlled power to avoid vaporization and gas formation, while tissue temperatures in MWA often exceed 100°C. Therefore, the temperature-dependency of the parameters, as well as water vaporization, influences the ablation zone in MWA relatively more. This could explain the discrepancy in incorporating these factors in MWA models and RFA models.

On the other hand, inclusion of large blood vessels in their model was more common in RFA models as compared to MWA models (11 out of 22 (50%) and 2 out of 16 (12.5%), respectively). In a retrospective clinical study, local tumor recurrence in RFA was related to peritumoral vessels with a diameter of at least 3 mm [63]. The so called "heat-sink effect" is considered to have greater impact in RFA as compared to MWA [64, 65]. In MWA, the heat-sink effect has only minimal effect on the created ablation zone [66]. When comparing the ex vivo RFA experiments evaluating their model with the DSC to in vivo RFA experiments, it can be concluded that ex vivo validation leads to better results. This is probably because ex vivo experiments are performed in non-perfused tissue, and thus neglecting the heat-sink effect.

Since MWA heats tissue to higher temperatures and at a faster rate, it is known to create larger, more homogenous and more predictable ablation zones compared to RFA [67]. However, these characteristics cannot be related to the found results of the systematic review. Figure 4 compares the RDDs of MWA and RFA of ex vivo validation, but no consistency can be found in the simulation results and MWA ablation zones having an improved predictability.

The included models vary in complexity. Some models are purely based on the Pennes' equation, while others include more sophisticated (bioheat) models. This review does not present all possible parameters modelled, for example the modeling of cooled needles and anatomical location is neglected. This emphasizes the complexity of ablation zone prediction and the number of parameters affecting the ablation zone. Theoretically, inclusion of all model characteristics, i.e. perfusion, large blood vessels, temperature-dependent parameters, water vaporization and image based patient specific models would result in the most accurate simulation. Besides, tissue contraction models, two-compartment models, cooled needles and RFA power control feedback loops could increase the accuracy. However, incorporation of these modelled parameters also increases the complexity, resulting in a trade-off between complexity and accuracy.

In addition the choice in bioheat model and cell-death model affects the prediction. The Pennes' equation is most used due to its simplicity and feasibility. Despite the wide use, the model has an important limitation. The equation only considers microvascular perfusion assuming a constant blood temperature of 37°C without flow directionality, while blood temperature of the vessels within and surrounding the ablation zone will increase during ablation [40]. Tucci et al. studied the use of three different bioheat equations, the Pennes' bioheat equation and two equations based on the porous

media theory [46]. According to the porous media theory, the tissue can be divided in two phases: the tissue phase, the cellular structure including the interstitial space, and the blood phase, a fluid part representing the blood flowing through the tissue phase. The amount of blood phase is expressed by porosity. Since the liver is a highly vascularized tissue, this model might give a more accurate representation of the liver tissue than considering a constant blood volume as in the Pennes' bioheat equation. They conclude that the porous media-based models achieved a better agreement with experimental results, since the Pennes' bioheat equation led to smaller ablation zones in medium and high voltage RFA (65V & 90V).

Unfortunately, this is the only study comparing the different bioheat models. Due to the heterogeneity in validation, it is impossible to identify factors that have the largest impact on the size and shape of the ablation zone and the simulation accuracy. To decide on the best structure and complexity of the model, a comparative clinical study should be conducted. This could be a retrospective or prospective study in which the clinical obtained ablation zone is compared to different simulations. This study requires a needle position scan, in order to simulate the ablation at corresponding needle position as well as a post-ablation CT scan to determine the clinical obtained ablation zone. After image processing, the different models could be compared to find out which combination of included parameters and models results in the most accurate prediction.

Tissue composition of both tumor and liver parenchyma may affect the shape and size of the ablation zone. Collins et al. investigated the effect of fat content on electrical and thermal conductivity in MWA [24]. Their results demonstrate adding fat altered the phantom behavior and resulted in varying ablation outcomes. The thermal conductivity was found to significantly decrease with increasing fat content. Another study by Liu et al. determined the effect of different liver parenchyma thermal conductivity on RFA using ex vivo agar phantoms and computer modeling [68]. They also found a negative correlation between thermal conductivity and fat content. In addition, they observed the so called 'oven-effect'. In case of a lower liver parenchyma thermal conductivity, higher temperatures were observed at the tumor edge, while the liver parenchyma surrounding the tumor had a lower temperature increase. This effect was also correlated with fat content. Hypothetically, the 'oven-effect' would lead to higher ablation temperatures within the tumor but increases the risk of narrow safety margins.

Nevertheless, Deshazer et al. and Servin et al. modelled the ablation extents in MWA using a two-compartment model [69, 70]. They concluded that a greater fat content in the liver leads to larger ablation volumes. The lower conductivity leads to less heat loss to surrounding tissue, since low thermal conducting tissue retains high temperatures, creating a larger ablation zone.

In cirrhotic livers, the perfusion may decrease. Deshazer et al. simulated the difference in perfusion between normal liver tissue and cirrhotic liver tissue, and found larger ablation zones in cirrhotic livers compared to normal liver tissue [69].

In a cirrhotic liver, multiple tissue-specific parameters affecting the ablation zone changes, since cirrhotic livers are less perfused and have a lower thermal conductivity. The effect of cirrhosis on ablation zone extent is important in clinical practice, since 80-90% of the most common primary liver malignancy, hepatocellular carcinoma, arises in cirrhotic livers [71, 72]. It is hard to model all the changes between cirrhotic and normal liver tissue. Therefore, the effect of cirrhotic liver tissue on ablation zone extent should be studied using in vivo clinical data. Young et al. evaluated 103 MWA procedures and found smaller ablation zones in cirrhotic livers, while in non-cirrhotic fatty livers, the ablation zones were larger compared to healthy liver tissue [73]. In RFA experiments performed in porcine livers, as well as in vivo in patients, smaller ablation zones in cirrhotic livers are observed [74, 75]. On the other hand, Wang et al. conducted a prospective clinical study, and found larger ablation zones in cirrhotic livers compared to non-cirrhotic livers in bipolar RFA [76]. However, the study group with cirrhotic livers were all hepatocellular carcinoma patients, while the normal liver group consisted of patients with metastatic liver cancer, which might cause a bias due to differences in tumor conductivity parameters.

These contradicting results requires further in vivo research in the effect of cirrhosis in MWA as well as RFA.

The bottleneck within current clinical practice is that the manufacturer's prediction is mostly based on ex vivo experiments [10]. 27 of the included computational models are also validated with ex vivo experiments (71%). Besides, most ex-vivo validations only use one specific ablation setting. However, the aim of the models for predicting ablation zones should be applicable for all clinically used ablation settings. For clinically implementation of computational models, the focus should shift to in vivo clinical experiments.

Another drawback from the performed validation experiments are the outcome measures used. Many articles only compare the predicted longitudinal and transversal diameter to the obtained ablation zone diameters. However, not only the diameters are of importance, but also the shape of the ablation zone. The ablation zone should completely cover the tumor including a safety margin. Outcome measures such as the DSC and surface deviation are therefore more expressive. However, most of the included articles only looked into the mean surface deviation, while the maximum surface deviation and its direction are of great importance to express the boundary discrepancies and its clinical impact.

Three included articles originate from the ClinicIMPACT project, which aims to bring an existing RFA model for liver cancer treatment into clinical practice [37, 38, 49, 77]. One of their publications, Voglreiter et al., presented a planning and simulation tool for RFA and highlighted the clinical applications [49]. The clinical study of Moche et al. evaluated this application prospectively and concluded that the model was accurately enough for clinical implementation with a surface deviation of 3.4 mm [37]. Nevertheless, within ablation it is aimed to obtain a 5 mm safety margin around the tumor, but the exact relationship between ablation margins and local recurrence is still unknown. Therefore, it is questionable if a surface deviation of 3.4 mm is accurate enough to ensure clinical safety margins and reduce local recurrences. Besides, the current discrepancy between the manufacturer's prediction and clinically obtained ablation zone is unknown. Hence, the match between the manufacturer's prediction and the clinically obtained ablation zone should be quantified in order to conclude on the accuracy needed for clinical implementation of computational models.

Innovations in the field of thermal ablation are often focused on prevention of local recurrences by means of advanced image processing, ablation margin quantification, optimized needle positioning and treatment planning. Ultimately, the incorporation of patient-specific ablation zone simulation could play an important role in this process. As ablation margins of >5mm are generally pursued, robust simulation with high accuracy are required. Mariappan et al. concluded that the simulation yields better accuracy when personalized perfusion values are given as input in the simulation model [38]. Other patient-specific parameters could also increase the accuracy of the simulation like thermal conductivity and tumor geometry for an accurate two compartment model. In conclusion, in vivo clinical studies evaluating patient-specific ablation zone simulation are necessary using outcome measures which represent the target overlap as well as the boundary discrepancies. In order to determine the desired accuracy, the current clinical practice should be evaluated and taken into account when evaluating simulations.

## 5. Conclusion

Computational models to simulate ablation zones in MWA and RFA show considerable heterogeneity in model type and validation methods. It is currently unknown which model is most accurate and best suitable for use in clinical practice. However, several studies have demonstrated good correlation between simulated ablation zones and in vivo ablations. To establish the added value of computational model-based ablation zone simulation, the accuracy of the current clinical practice, i.e., the correspondence between manufacturer's prediction and clinically obtained ablation zone, should be studied.

## References

1. Reig M, Forner A, Rimola J, Ferrer-Fàbrega J, Burrel M, Garcia-Criado Á, et al. BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. *J Hepatol*. 2022;76(3):681-93.
2. Crocetti L, de Baére T, Pereira PL, Tarantino FP. CIRSE Standards of Practice on Thermal Ablation of Liver Tumours. *Cardiovasc Intervent Radiol*. 2020;43(7):951-62.
3. Kim YS, Lee WJ, Rhim H, Lim HK, Choi D, Lee JY. 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(3):758-65.
4. Laimer G, Schullian P, Jaschke N, Putzer D, Eberle G, Alzaga A, 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):2463-72.
5. Chu KF, Dupuy DE. Thermal ablation of tumours: biological mechanisms and advances in therapy. *Nat Rev Cancer*. 2014;14(3):199-208.
6. Shin SW, Ahn KS, Kim SW, Kim T-S, Kim YH, Kang KJ. Liver Resection Versus Local Ablation Therapies for Hepatocellular Carcinoma Within the Milan Criteria: A Systematic Review and Meta-analysis. *Annals of Surgery*. 2021;273(4):656-66.
7. Xu XL, Liu XD, Liang M, Luo BM. Radiofrequency Ablation versus Hepatic Resection for Small Hepatocellular Carcinoma: Systematic Review of Randomized Controlled Trials with Meta-Analysis and Trial Sequential Analysis. *Radiology*. 2018;287(2):461-72.
8. 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(1):166-75.
9. Wang CZ, Yan GX, Xin H, Liu ZY. Oncological outcomes and predictors of radiofrequency ablation of colorectal cancer liver metastases. *World J Gastrointest Oncol*. 2020;12(9):1044-55.
10. Ruiter SJS, Heerink WJ, de Jong KP. Liver microwave ablation: a systematic review of various FDA-approved systems. *Eur Radiol*. 2019;29(8):4026-35.
11. Huang HW. Influence of blood vessel on the thermal lesion formation during radiofrequency ablation for liver tumors. *Med Phys*. 2013;40(7):073303.
12. Ahmed M, Liu Z, Humphries S, Goldberg SN. Computer modeling of the combined effects of perfusion, electrical conductivity, and thermal conductivity on tissue heating patterns in radiofrequency tumor ablation. *Int J Hyperthermia*. 2008;24(7):577-88.
13. Liu Z, Ahmed M, Gervais D, Humphries S, Goldberg SN. Computer modeling of factors that affect the minimum safety distance required for radiofrequency ablation near adjacent nontarget structures. *J Vasc Interv Radiol*. 2008;19(7):1079-86.
14. Siriwardana PN, Singh S, Johnston EW, Watkins J, Bandula S, Illing RO, et al. Effect of Hepatic Perfusion on Microwave Ablation Zones in an Ex Vivo Porcine Liver Model. *J Vasc Interv Radiol*. 2017;28(5):732-9.
15. Chiang J, Wang P, Brace CL. Computational modelling of microwave tumour ablations. *Int J Hyperthermia*. 2013;29(4):308-17.
16. Singh S, Melnik R. Thermal ablation of biological tissues in disease treatment: A review of computational models and future directions. *Electromagn Biol Med*. 2020;39(2):49-88.
17. Pennes HH. Analysis of tissue and arterial blood temperatures in the resting human forearm. 1948. *J Appl Physiol (1985)*. 1998;85(1):5-34.
18. Audigier C, Mansi T, Delingette H, Rapaka S, Mihalef V, Sharma P, et al. Lattice Boltzmann method for fast patient-specific simulation of liver tumor ablation from CT images. *Med Image Comput Assist Interv*. 2013;16(Pt 3):323-30.
19. Audigier C, Mansi T, Delingette H, Rapaka S, Mihalef V, Carnegie D, et al. Efficient Lattice Boltzmann Solver for Patient-Specific Radiofrequency Ablation of Hepatic Tumors. *Ieee Transactions on Medical Imaging*. 2015;34(7):1576-89.

20. Audigier C, Mansi T, Delingette H, Rapaka S, Passerini T, Mihalef V, et al. Comprehensive preclinical evaluation of a multi-physics model of liver tumor radiofrequency ablation. *Int J Comput Assist Radiol Surg.* 2017;12(9):1543-59.
21. Cavagnaro M, Pinto R, Lopresto V. Numerical models to evaluate the temperature increase induced by ex vivo microwave thermal ablation. *Phys Med Biol.* 2015;60(8):3287-311.
22. Chang IA, Nguyen UD. Thermal modeling of lesion growth with radiofrequency ablation devices. *Biomed Eng Online.* 2004;3(1):27.
23. Chen R, Zhang J, Kong D, Lou Q, Lu F. Fast calculation of 3D radiofrequency ablation zone based on a closed-form solution of heat conduction equation fitted by ex vivo measurements. *Phys Med Biol.* 2021;66(5):055022.
24. Collins JA, Heiselman JS, Clements LW, Weis JA, Brown DB, Miga MI. Toward Image Data-Driven Predictive Modeling for Guiding Thermal Ablative Therapy. *Ieee Transactions on Biomedical Engineering.* 2020;67(6):1548-57.
25. Deshazer G, Hagmann M, Merck D, Sebek J, Moore KB, Prakash P. Computational modeling of 915 MHz microwave ablation: Comparative assessment of temperature-dependent tissue dielectric models. *Med Phys.* 2017;44(9):4859-68.
26. Deshazer G, Prakash P, Merck D, Haemmerich D. Experimental measurement of microwave ablation heating pattern and comparison to computer simulations. *Int J Hyperthermia.* 2017;33(1):74-82.
27. Duan B, Wen R, Fu Y, Chua KJ, Chui CK. Probabilistic finite element method for large tumor radiofrequency ablation simulation and planning. *Med Eng Phys.* 2016;38(11):1360-8.
28. Faridi P, Keselman P, Fallahi H, Prakash P. Experimental assessment of microwave ablation computational modeling with MR thermometry. *Med Phys.* 2020;47(9):3777-88.
29. Gao HJ, Wu SC, Wang XR, Hu R, Zhou ZH, Sun XC. Temperature simulation of microwave ablation based on improved specific absorption rate method compared to phantom measurements. *Computer Assisted Surgery.* 2017;22:9-17.
30. Gao H, Wang X, Wu S, Zhou Z, Bai Y, Wu W. Conformal coverage of liver tumors by the thermal coagulation zone in 2450-MHz microwave ablation. *Int J Hyperthermia.* 2019;36(1):591-605.
31. Gao HJ, Wang XR, Wu SC, Zhou ZH, Bai YP. 2450-MHz microwave ablation temperature simulation using temperature-dependence feedback of characteristic parameters. *International Journal of Rf and Microwave Computer-Aided Engineering.* 2019;29(1).
32. Haemmerich D, Staelin ST, Tungjitkusolmun S, Lee FT, Jr., Mahvi DM, Webster JG. Hepatic bipolar radio-frequency ablation between separated multiprong electrodes. *IEEE Trans Biomed Eng.* 2001;48(10):1145-52.
33. Haemmerich D. Mathematical modeling of impedance controlled radiofrequency tumor ablation and ex-vivo validation. *Annu Int Conf IEEE Eng Med Biol Soc.* 2010;2010:1605-8.
34. Lim D, Namgung B, Woo DG, Choi JS, Kim HS, Tack GR. Effect of input waveform pattern and large blood vessel existence on destruction of liver tumor using radiofrequency ablation: finite element analysis. *J Biomech Eng.* 2010;132(6):061003.
35. Lopresto V, Pinto R, Farina L, Cavagnaro M. Microwave thermal ablation: Effects of tissue properties variations on predictive models for treatment planning. *Medical Engineering & Physics.* 2017;46:63-70.
36. Macchi EG, Gallati M, Braschi G, Persi E. Dielectric properties of RF heated ex vivo porcine liver tissue at 480 kHz: measurements and simulations. *Journal of Physics D-Applied Physics.* 2014;47(48).
37. Moche M, Busse H, Futterer JJ, Hinestroza CA, Seider D, Brandmaier P, et al. Clinical evaluation of in silico planning and real-time simulation of hepatic radiofrequency ablation (ClinicMPPACT Trial). *Eur Radiol.* 2020;30(2):934-42.
38. Mariappan P, Weir P, Flanagan R, Voglreiter P, Alhonnoro T, Pollari M, et al. GPU-based RFA simulation for minimally invasive cancer treatment of liver tumours. *Int J Comput Assist Radiol Surg.* 2017;12(1):59-68.

39. Nolte T, Vaidya N, Baragona M, Elevelt A, Lavezzo V, Maessen R, et al. Study of flow effects on temperature-controlled radiofrequency ablation using phantom experiments and forward simulations. *Med Phys*. 2021;48(9):4754-68.
40. Ooi EH, Lee KW, Yap S, Khattab MA, Liao IY, Ooi ET, et al. The effects of electrical and thermal boundary condition on the simulation of radiofrequency ablation of liver cancer for tumours located near to the liver boundary. *Comput Biol Med*. 2019;106:12-23.
41. Payne S, Flanagan R, Pollari M, Alhonnoro T, Bost C, O'Neill D, et al. Image-based multi-scale modelling and validation of radio-frequency ablation in liver tumours. *Philos Trans A Math Phys Eng Sci*. 2011;369(1954):4233-54.
42. Radosevic A, Prieto D, Burdío F, Berjano E, Prakash P, Trujillo M. Short pulsed microwave ablation: computer modeling and ex vivo experiments. *Int J Hyperthermia*. 2021;38(1):409-20.
43. Singh S, Melnik R. Coupled thermo-electro-mechanical models for thermal ablation of biological tissues and heat relaxation time effects. *Physics in Medicine and Biology*. 2019;64(24).
44. Subramanian S, Mast TD. Optimization of tissue physical parameters for accurate temperature estimation from finite-element simulation of radiofrequency ablation. *Phys Med Biol*. 2015;60(19):N345-55.
45. Tehrani MHH, Soltani M, Kashkooli FM, Raahemifar K. Use of microwave ablation for thermal treatment of solid tumors with different shapes and sizes-A computational approach. *Plos One*. 2020;15(6).
46. Tucci C, Trujillo M, Berjano E, Iasiello M, Andreozzi A, Vanoli GP. Pennes' bioheat equation vs. porous media approach in computer modeling of radiofrequency tumor ablation. *Sci Rep*. 2021;11(1):5272.
47. Tucci C, Trujillo M, Berjano E, Iasiello M, Andreozzi A, Vanoli GP. Mathematical modeling of microwave liver ablation with a variable-porosity medium approach. *Comput Methods Programs Biomed*. 2022;214:106569.
48. Vaidya N, Baragona M, Lavezzo V, Maessen R, Veroy K. Simulation study of the cooling effect of blood vessels and blood coagulation in hepatic radio-frequency ablation. *Int J Hyperthermia*. 2021;38(1):95-104.
49. Voglreiter P, Mariappan P, Pollari M, Flanagan R, Blanco Sequeiros R, Portugaller RH, et al. RFA Guardian: Comprehensive Simulation of Radiofrequency Ablation Treatment of Liver Tumors. *Sci Rep*. 2018;8(1):787.
50. Wang XR, Gao HJ, Wu SC, Jiang T, Zhou ZH, Bai YP. Numerical evaluation of ablation zone under different tip temperatures during radiofrequency ablation. *Math Biosci Eng*. 2019;16(4):2514-31.
51. Wang J, Wu S, Wu Z, Gao H, Huang S. Influences of blood flow parameters on temperature distribution during liver tumor microwave ablation. *Front Biosci (Landmark Ed)*. 2021;26(9):504-16.
52. Welp C, Siebers S, Ermert H, Werner J. Investigation of the influence of blood flow rate on large vessel cooling in hepatic radiofrequency ablation. *Biomed Tech (Berl)*. 2006;51(5-6):337-46.
53. Wu TN, Li P, Shao Q, Hong J, Yang L, Wu SC. A simulation-experiment method to characterize the heat transfer in ex-vivo porcine hepatic tissue with a realistic microwave ablation system. *Numerical Heat Transfer Part a-Applications*. 2013;64(9):729-43.
54. Zhai W, Xu J, Zhao Y, Song Y, Sheng L, Jia P. Preoperative surgery planning for percutaneous hepatic microwave ablation. *Med Image Comput Comput Assist Interv*. 2008;11(Pt 2):569-77.
55. Zhang B, Moser MA, Zhang EM, Luo Y, Zhang W. Numerical analysis of the relationship between the area of target tissue necrosis and the size of target tissue in liver tumours with pulsed radiofrequency ablation. *Int J Hyperthermia*. 2015;31(7):715-25.
56. Amabile C, Ahmed M, Solbiati L, Meloni MF, Solbiati M, Cassarino S, et al. Microwave ablation of primary and secondary liver tumours: ex vivo, in vivo, and clinical characterisation. *International Journal of Hyperthermia*. 2017;33(1):34-42.

57. Sun Y, Cheng Z, Dong L, Zhang G, Wang Y, Liang P. Comparison of temperature curve and ablation zone between 915- and 2450-MHz cooled-shaft microwave antenna: results in ex vivo porcine livers. *Eur J Radiol.* 2012;81(3):553-7.
58. Gallati M, Braschi G, editors. On the simulation of radio frequency thermal lesions in porcine liver. Proceedings of the 32nd IASTED International Conference on Modelling, Identification and Control; 2013.
59. Goldberg SN, Stein MC, Gazelle GS, Sheiman RG, Kruskal JB, Clouse ME. Percutaneous radiofrequency tissue ablation: optimization of pulsed-radiofrequency technique to increase coagulation necrosis. *J Vasc Interv Radiol.* 1999;10(7):907-16.
60. Pop M, Molckovsky A, Chin L, Kolios MC, Jewett MA, Sherar MD. Changes in dielectric properties at 460 kHz of kidney and fat during heating: importance for radio-frequency thermal therapy. *Phys Med Biol.* 2003;48(15):2509-25.
61. Prakash P. Theoretical modeling for hepatic microwave ablation. *Open Biomed Eng J.* 2010;4:27-38.
62. Ji Z, Brace CL. Expanded modeling of temperature-dependent dielectric properties for microwave thermal ablation. *Phys Med Biol.* 2011;56(16):5249-64.
63. Lu DS, Raman SS, Limanond P, Aziz D, Economou J, Busuttill R, et al. Influence of large peritumoral vessels on outcome of radiofrequency ablation of liver tumors. *J Vasc Interv Radiol.* 2003;14(10):1267-74.
64. Lubner MG, Brace CL, Hinshaw JL, Lee FT, Jr. Microwave tumor ablation: mechanism of action, clinical results, and devices. *J Vasc Interv Radiol.* 2010;21(8 Suppl):S192-203.
65. Brace CL. Radiofrequency and microwave ablation of the liver, lung, kidney, and bone: what are the differences? *Curr Probl Diagn Radiol.* 2009;38(3):135-43.
66. Yu NC, Raman SS, Kim YJ, Lassman C, Chang X, Lu DS. Microwave liver ablation: influence of hepatic vein size on heat-sink effect in a porcine model. *J Vasc Interv Radiol.* 2008;19(7):1087-92.
67. Chiang J, Hynes K, Brace CL. Flow-dependent vascular heat transfer during microwave thermal ablation. *Annu Int Conf IEEE Eng Med Biol Soc.* 2012;2012:5582-5.
68. Liu Z, Ahmed M, Weinstein Y, Yi M, Mahajan RL, Goldberg SN. Characterization of the RF ablation-induced 'oven effect': the importance of background tissue thermal conductivity on tissue heating. *Int J Hyperthermia.* 2006;22(4):327-42.
69. Deshazer G, Merck D, Haggmann M, Dupuy DE, Prakash P. Physical modeling of microwave ablation zone clinical margin variance. *Med Phys.* 2016;43(4):1764.
70. Servin F, Collins JA, Heiselman JS, Frederick-Dyer KC, Planz VB, Geevarghese SK, et al. Fat Quantification Imaging and Biophysical Modeling for Patient-Specific Forecasting of Microwave Ablation Therapy. *Front Physiol.* 2021;12:820251.
71. Wallace MC, Preen D, Jeffrey GP, Adams LA. The evolving epidemiology of hepatocellular carcinoma: a global perspective. *Expert Rev Gastroenterol Hepatol.* 2015;9(6):765-79.
72. Fattovich G, Stroffolini T, Zagni I, Donato F. Hepatocellular carcinoma in cirrhosis: incidence and risk factors. *Gastroenterology.* 2004;127(5 Suppl 1):S35-50.
73. Young S, Rivard M, Kimyon R, Sanghvi T. Accuracy of liver ablation zone prediction in a single 2450MHz 100 Watt generator model microwave ablation system: An in human study. *Diagn Interv Imaging.* 2020;101(4):225-33.
74. Livraghi T, Goldberg SN, Lazzaroni S, Meloni F, Solbiati L, Gazelle GS. Small hepatocellular carcinoma: treatment with radio-frequency ablation versus ethanol injection. *Radiology.* 1999;210(3):655-61.
75. Wang LG, Jiang WJ, Fan WJ, Zheng YB, Song XP, Liu S, et al. Microwave Ablation: The Differences Between Biliary Cirrhosis and Normal Porcine Liver Using a Cooled-tip Electrode. *Anticancer Res.* 2016;36(3):1221-6.
76. Wang H, Lee JC, Cao K, Tang HW, Wang S, Zhang ZY, et al. What is the difference in ablation zone of multi-bipolar radiofrequency ablation between liver cirrhosis and normal liver background? - a prospective clinical study. *Int J Hyperthermia.* 2020;37(1):1248-59.

77. ClinicIMPPACT. ClinicIMPPACT 2017 [Available from: <https://www.clinicimppact.eu/>].

## Appendix A: Search Strategies

### String PubMed

["Ablation Techniques"[Mesh] OR "Ablation Technique\*"[tw] OR "Thermal ablation"[tw] OR "Radiofrequency ablation"[tw] OR "Microwave ablation"[tw] OR "ablation"[tw]]

AND

["Liver Neoplasms"[Mesh] OR "Carcinoma, Hepatocellular"[Mesh] OR "Liver"[Mesh] OR Hepatocellular carcinoma\*[tw] OR "hcc"[tw] OR liver tum\*[tw]]

AND

["Mathematical Computing"[Mesh] OR "Computer Simulation"[Mesh] OR "Thermal Conductivity"[Mesh] OR "Electric Conductivity"[Mesh] OR "Finite Element Analysis"[Mesh] OR predictive model\*[tw] OR prediction model\*[tw] OR computational model\*[tw] OR mathematical model\*[tw] OR "simulation"[tw]]

### String Web of Science

ALL= ("Ablation Techniques" OR "Ablation Technique\*" OR "Thermal ablation" OR "Radiofrequency ablation" OR "Microwave ablation" OR "ablation")

AND

ALL = ("Liver Neoplasms" OR "Carcinoma, Hepatocellular" OR "Liver" OR "Hepatocellular carcinoma\*" OR "hcc" OR "liver tum\*")

AND

ALL = ("Mathematical Computing" OR "Computer Simulation" OR "Thermal Conductivity" OR "Electric Conductivity" OR "Finite Element Analysis" OR "predictive model\*" OR "prediction model\*" OR "computational model\*" OR "mathematical model\*" OR "simulation")

# Study Protocol PROMETHEUS: PROspective Multicenter study to Evaluate the correlation between safety margin and local recurrence after THERmal ablation USING image co-registration in patients with hepatocellular carcinoma

Timo T.M. Oosterveer MSc<sup>1\*</sup>, Gonnie C.M. van Erp<sup>1\*</sup>, Pim Hendriks MSc<sup>1</sup>, Alexander Broersen PhD<sup>1,2</sup>, Christiaan G. Overduin PhD<sup>3</sup>, Carla S.P. van Rijswijk MD PhD<sup>1</sup>, Arian R. van Erkel MD PhD<sup>1</sup>, Rutger W. van der Meer MD PhD<sup>1</sup>, Maarten E. Tushuizen MD PhD<sup>4</sup>, Adriaan Moelker MD PhD<sup>5</sup>, Martijn R. Meijerink MD PhD<sup>6</sup>, Otto M. van Delden MD PhD<sup>6</sup>, Koert P. de Jong MD PhD<sup>7</sup>, Christiaan van der Leij MD PhD<sup>8</sup>, Maarten L.J. Smits MD PhD<sup>9</sup>, Thijs A.J. Urlings MD MSc<sup>10</sup>, Jeffrey P.B.M. Braak<sup>11</sup>, Elma Meershoek-Klein Kranenbarg MSc<sup>11</sup>, Bianca van Duijn-deVreugd<sup>1</sup>, Evelijn Zeijdner PhD<sup>12</sup>, Jelle J. Goeman PhD<sup>13</sup>, Jurgen J. Fütterer MD PhD<sup>3</sup>, Minneke J. Coenraad MD PhD<sup>4</sup>, Jouke Dijkstra PhD<sup>1,2</sup>, Mark C. Burgmans MD PhD<sup>1</sup>, The Dutch Hepatocellular and Cholangiocarcinoma Group

1. Department of Radiology, Leiden University Medical Center, The Netherlands
2. Division of Image Processing, Leiden University Medical Center, The Netherlands
3. Department of Medical Imaging, Radboud University Medical Center, The Netherlands
4. Department of Gastroenterology and Hepatology, Leiden University Medical Center, The Netherlands
5. Department of Radiology and Nuclear Medicine, Erasmus University Medical Center, The Netherlands
6. Department of Radiology, Amsterdam University Medical Center, The Netherlands
7. Department of Hepatobiliary and Pancreatic Surgery, University Medical Center Groningen, The Netherlands
8. Department of Radiology, Maastricht University Medical Center, The Netherlands
9. Department of Radiology, University Medical Center Utrecht, The Netherlands
10. Department of Radiology, Haaglanden Medical Center, The Netherlands
11. Department of Surgery, Leiden University Medical Center, The Netherlands
12. Dutch Oncology Research Platform, Utrecht, The Netherlands
13. Department of Biomedical Data Sciences, Leiden University Medical Center, The Netherlands

\* Shared first authorship.

## Abstract

**Purpose** The primary objective is to determine the minimal ablation margin required to achieve a local recurrence rate of <10% in patients with hepatocellular carcinoma undergoing thermal ablation. Secondary objectives are to analyze the correlation between ablation margins and local recurrence and to assess efficacy.

**Materials and Methods** This study is a prospective, multicenter, non-experimental, non-comparative, open-label study. Patients >18 years with Barcelona Clinic Liver Cancer stage 0/A hepatocellular carcinoma (or B with a maximum of two lesions <5cm each) are eligible. Patients will undergo dual-phase contrast-enhanced computed tomography directly before and after ablation. Ablation margins will be quantitatively assessed using co-registration software, blinding assessors (i.e. two experienced radiologists) for outcome. Presence and location of recurrence are evaluated independently on follow-up scans by two other experienced radiologists, blinded for the quantitative margin analysis. A sample size of 189 tumors (~145 patients) is required to show with 80% power that the risk of local recurrence is confidently below 10%. A two-sided binomial z-test will be used to test the null hypothesis that the local recurrence rate is ≥10% for patients with a minimal ablation margin ≥2mm. Logistic regression

will be used to find the relationship between minimal ablation margins and local recurrence. Kaplan-Meier estimates are used to assess local and overall recurrence, disease-free and overall survival.

**Discussion** It is expected that this study will result in a clear understanding of the correlation between ablation margins and local recurrence. Using co-registration software in future patients undergoing ablation for hepatocellular carcinoma may improve intraprocedural evaluation of technical success.

*Trial registration* The Netherlands Trial Register (NL9713), <https://www.trialregister.nl/trial/9713>.

**Keywords:** Hepatocellular carcinoma, thermal ablation, minimal ablation margin, local recurrence, co-registration

## 1. Introduction

Hepatocellular carcinoma (HCC) is the third most common cause of cancer-related death in the world [1]. Surgical resection is the preferred treatment for patients with HCC  $\geq 2$  cm, but resection may not be feasible or perilous as a result of unfavorable tumor location or underlying liver disease [1]. Thermal ablation is an alternative treatment with lower complication rates, lower costs and shorter hospital stay [2]. However, hepatic resection yields better results regarding local recurrence (LR) [2].

To reduce the risk of LR after thermal ablation, it is generally recommended to ablate a tumor with a minimal ablation margin (MAM) of  $>5$ mm [3]. A clear relation between MAM and LR seems evident, but the precise relationship still needs to be established. Also, there is no validated, standardized method to accurately determine a MAM. Commonly, margins are assessed through side by side positioning of pre- and post-ablation cross-sectional images and visual qualitative assessment. Over recent years, co-registration software has become available that allows immediate three-dimensional quantitative assessment of the MAM. It would potentially be the equivalent of the frozen section that is used for margin control during surgery. Yet, quantitative margin assessment during ablation has not been validated in large prospective studies, and is not common practice.

Quantitative margin assessment can only determine intra-procedural treatment decisions if the correlation between MAM and LR is clearly understood. In this study, obtained margins will be quantitatively assessed and correlated with clinical outcome. The primary objective is to determine the MAM required to achieve a local recurrence rate (LRR) of  $<10\%$  in patients with HCC [4]. Secondary objectives are to analyze the correlation between MAM and LR and to assess efficacy of thermal ablation in patients with HCC.

## 2. Materials and Methods

### 2.1 Trial design and study setting

The PROMETHEUS trial is a prospective, multicenter, non-experimental, non-comparative, open-label study. The sponsor of the study is the Leiden University Medical Center (LUMC). This study is a collaboration between Dutch academic centers and cancer organizations. The trial is funded by the Dutch Cancer Society and registered at <https://www.trialregister.nl> (ID: NL9713).

### 2.2 Participants

Patients over 18 years with Barcelona Clinic Liver Cancer stage 0-A HCC, or stage B with a maximum of two lesions  $<5$ cm each, are eligible. Full inclusion and exclusion criteria are provided in Table 1.

### 2.3 Interventions

All patients are discussed in a multidisciplinary tumor board for eligibility and consented prior to inclusion. The ablation procedure and follow-up will be according to local standard of care. Interventions and important time points are shown in Fig 1.

**Table 1.** Full inclusion and exclusion criteria.

Inclusion criteria	Exclusion criteria
Age 18 years or above	Estimated GFR <30 ml/min
HCC very early (0) or early stage (A) according to the BCLC staging system, OR HCC intermediate stage (B) according to the BCLC staging system with a maximum of two lesions of ≤5cm each	Known severe allergy to contrast medium ASA classification >3 Child Pugh C
Either de novo or recurrent HCC: prior locoregional therapy is allowed in the study *	Tumor related ECOG ≥1
Candidate for percutaneous thermal ablation as discussed in a multidisciplinary tumor board	Neoadjuvant transarterial therapy (TACE, TAE or TARE), i.e. combination therapy of transarterial therapy
Absence of any psychological, familial, sociological or geographical condition potentially hampering compliance with the study protocol and follow-up schedule	Portal vein tumor invasion Extrahepatic metastasis
Written informed consent	Uncorrectable coagulopathy

*ASA American Society of Anesthesiologists, BCLC Barcelona Clinic Liver Cancer, ECOG Eastern Cooperative Oncology Group, GFR glomerular filtration rate, HCC hepatocellular carcinoma, TACE transarterial chemoembolization, TAE trombo-endarterectomy, TARE transarterial radioembolization*

*\*Recurrence in an area with prior TACE or TARE treatment is considered to be combination therapy and thus excluded. In case of prior TACE/TARE treatment, only recurrence in another area of the liver may be included.*

Both radiofrequency ablation (RFA) and microwave ablation (MWA) are allowed in the study. All patients will undergo dual-phase contrast-enhanced computed tomography (CECT), i.e. arterial and venous phase, directly before and after the ablation. Ablation and CECT will be performed under general anesthesia. The pre-ablation and post-ablation CECT will be performed during apnea to minimize breathing artifacts. Alternatively, high-frequency jet ventilation may be used.

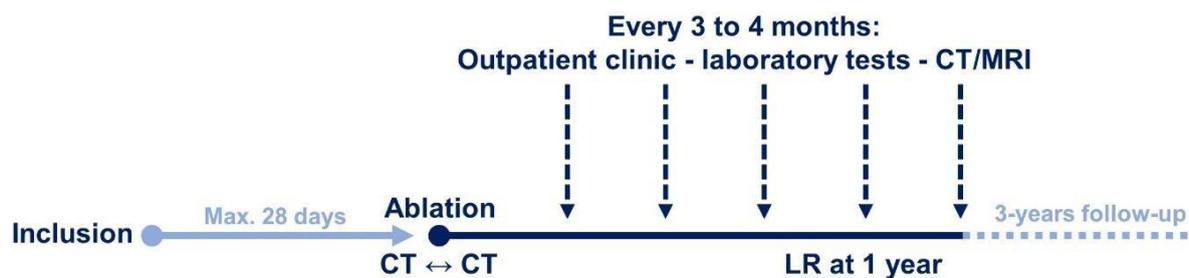
At the end of the procedure, the interventional radiologist will determine whether complete tumor ablation with sufficient margins was achieved. All patients will be treated with the intend to obtain complete tumor ablation with a >5mm margin and it is left at the discretion of the treating interventional radiologist to determine whether technical success has been achieved. Assessment will be performed as per current practice, i.e. visual qualitative assessment in most centers. Peri-procedural care will be in accordance with the protocol of the local institution.

#### 2.4 Follow-up

Patients will undergo physical examination, laboratory tests, contrast-enhanced magnetic resonance imaging of the liver and chest-CT every 3-4 months after ablation until liver transplantation, untreatable progression or death. Follow-up scans will be reviewed independently by two experienced interventional radiologists, other than the radiologists assessing the quantitative MAM (see MAM analysis), to determine the presence and location of recurrence. These radiologists will be blinded for the analyses of the quantitative MAM. Disagreement between the two radiologists will be resolved by consensus reading.

#### 2.5 MAM analysis

All pre- and post-ablation CECT images will be transferred on-line to the LUMC using ALEA Clinical (FormsVision, The Netherlands). Two experienced interventional radiologists, blinded for outcome, will independently perform delineation of the tumor and ablation zone, on the pre- and post-ablation CECT respectively. The pre- and post-ablation CECT will be co-registered using post-processing



**Fig 1.** Schematic overview of interventions and major time points for participants.

software (deLIVERed, LUMC) to quantitatively assess the MAM (Fig 2). Discordances of >3mm between both radiologists will be resolved by consensus reading, otherwise the mean MAM will be calculated. The mean MAM will then be correlated with the presence and location of LR. The results of deLIVERed will be compared with SAFIR (Fraunhofer-Gesellschaft) software to determine whether results are reproducible using different co-registration software. All clinical data will be entered in Castor Electronic Data Capture and subsequently analyzed using appropriate software packages (SPSS or R).

## 2.6 Outcomes

### Primary endpoint

The primary endpoint is the MAM that results in an LRR <10% at 1-year follow-up.

### Secondary endpoints

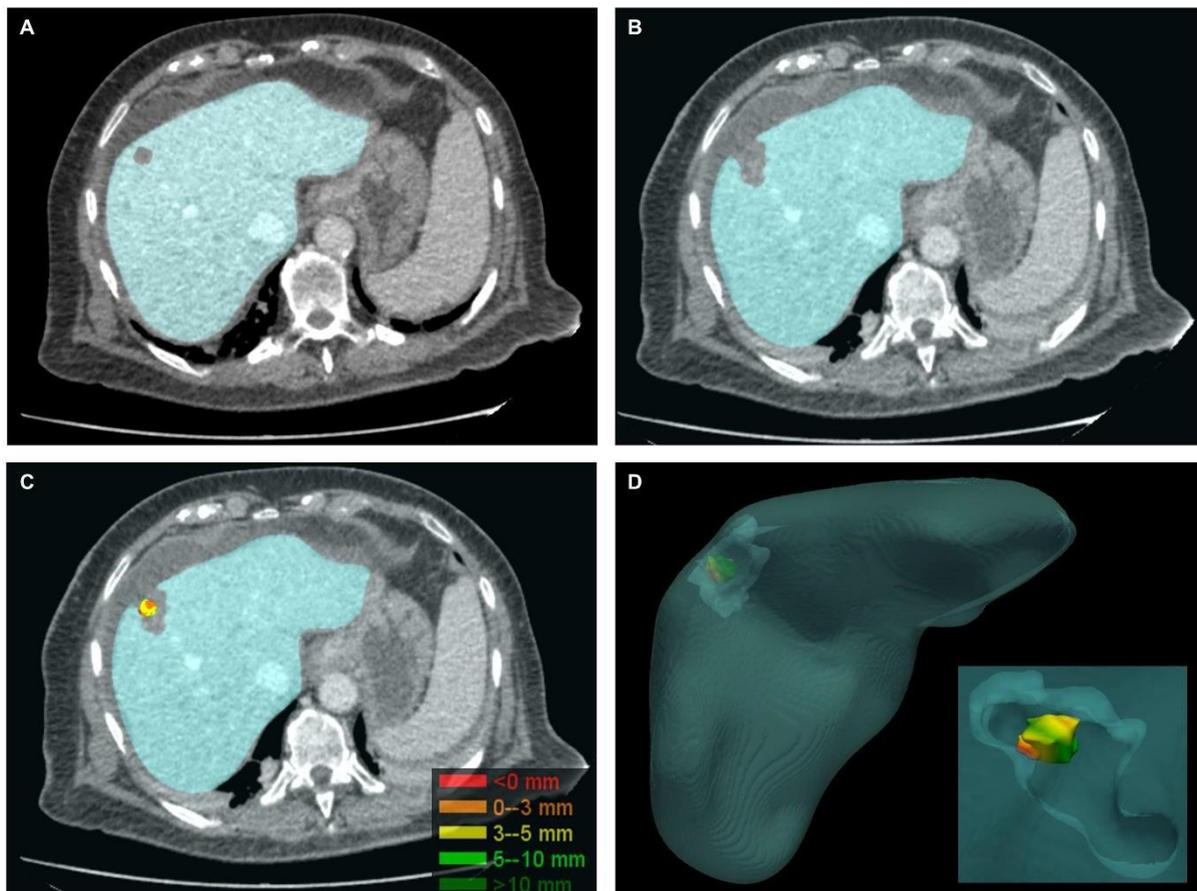
LR at 1 year will be analyzed for different MAM categories: <0mm, 0-3mm, 3-5mm and ≥5mm. Also, local and overall recurrence rates and disease-free survival (DFS) and overall survival (OS) at 1, 2 and 3 years will be analyzed. Finally, the relation between LR and DFS and OS will be investigated.

### Sample size

A sample size of 189 tumors would be sufficient to show with 80% power that the risk of LR is confidently below 10%, assuming a true risk of 4% for tumors ablated with a MAM ≥2mm, based on a study by Kim et al [5]. Other retrospective studies into local recurrence provide similar numbers as provided by Kim et al. [5-7]. The calculation is based on the normal approximation of the binomial distribution ( $d=0.10-0.04=0.06$ ;  $s=\sqrt{0.04*0.96}=0.20$ ;  $z=1.96+0.84$  (80% power);  $N=(s*z/d)^2=84$  tumors with ablation margin ≥2mm; total number of tumors  $N=84/.445=189$ ). Based on our own data, patients will have an average of 1.3 tumors per patient. Thus for 80% power, a sample size of 145 patients is indicated. Taking a potential drop-out rate of 10% into account, the needed sample size is 165 patients.

## 2.7 Statistical methods

For the primary objective, a two-sided binomial z-test will be used to test (reject) the null hypothesis that the LRR is ≥10% for patients with a MAM ≥2mm. Additionally, as a secondary analysis, a logistic regression model will be used to find the relationship between MAM and LR. Kaplan-Meier estimates will be used to assess local and overall recurrence, DFS and OS. Survival data will be censored at the date of last follow-up if patients are still alive. The log-rank test will be used to compare recurrence for different MAM categories. Logistic regression analyses will be performed to determine possible independent predictors for local and overall recurrence, DFS and OS. A p-value <0.05 will be considered significant.



**Fig 2.** 3D Quantitative MAM assessment using deLIVERed in a 78 years old female with a single HCC of 1 cm. **A** Axial slice showing the liver mask based on delineation of the liver and tumor on the pre-ablation venous phase CT-images. **B** Axial slice showing the liver mask based on delineation of the liver and ablation zone in the venous phase post-ablation CT-images. **C** 2D representation of the 3D tumor model with color coded ablation margins after co-registration. Ablation margins are calculated in 3D, hence showing unexpected tight medial ablation margins which are not visible in 2D. **D** 3D model of the tumor and ablation zone with color coded ablation margins after co-registration

### 3. Discussion

Thermal ablation is on its way to replace surgical resection as the treatment of first choice for patients with early stage HCC. Thermal ablation offers clear advantages over surgery in an era with rising healthcare costs and an aging population. The task that lies ahead is to bring the efficacy of thermal ablation up to par with surgical resection. Various studies have demonstrated that LRRs comparable to resection can be achieved if sufficient ablation margins are obtained.

Recommendations to ablate a liver tumor with a MAM >5mm are mainly based on expert opinion and pertain to treatment intent rather than the actual obtained margins. Ablation systems have predefined algorithms, based on in vitro experiments, to predict the size and shape of the ablation, but tissue factors influence the actual ablation volume and size. Several studies have demonstrated that true margins are often narrower than intended and often misjudged by conventional side-by-side evaluation of pre- and post-ablation images [5, 6, 8-11].

Retrospective studies have demonstrated the potential of quantitative MAM assessment using image co-registration. In a study including 110 patients with 176 HCCs, the MAM was assessed using CECT-CECT co-registration and proved to be the only significant independent predictor of local tumor

progression (LTP) [6]. For each millimeter increase of the MAM, a 30% reduction of the relative risk for LTP was found (OR=0.7, 95%CI 0.5-0.98, p=0.036). No LTP was detected in lesions with a MAM >5mm, but only in 37.5% of tumors this MAM was obtained. Similar results were reported in a study by Kim et al., which included 103 patients with 110 HCCs [5]. MAM was also assessed with CECT-CECT co-registration and strongly correlated with LRRs: 22.7%, 18.9%, 5.9% and 0% for margins of  $\geq 0$ mm,  $\geq 1$ mm,  $\geq 2$ mm, and  $\geq 3$ mm, respectively. Remarkably, in only 2.7% of the ablations the MAM was >5mm. Park et al. found that the cumulative incidence of LR was twice as high in patients with a MAM <2mm, compared with a MAM  $\geq 2$ mm [12]. Another retrospective study by Jiang et al. found similar results, but the post-ablation CT used for co-registration was obtained 1 months after ablation and shrinkage of the ablation zone within this period may have led to underestimation of margins [7].

As PROMETHEUS is a prospective study with a standardized imaging protocol, it is expected that this study will result in a clearer understanding of the correlation between MAM and LR and in validation of quantitative margin analysis. Knowledge provided will be important for the implementation of image co-registration as an intraprocedural decision-making tool in clinical practice. In future patients, it may help to objectively identify areas at risk of LR and instigate re-ablation during the same treatment session if margins are deemed to be insufficient. Following the above-mentioned retrospective studies, PROMETHEUS is the next step towards clinical use of image co-registration as an intraprocedural decision-making tool.

Our study has several limitations. The study is designed as a prospective, single arm observational study without control group. However, this is also a strength, as the PROMETHEUS study allows optimal and standardized imaging of the tumor and ablation zone during the same session. In addition, it might be that the optimal ablation margin is dependent on ablation size and type of ablation system used. It is allowed to include patients with intermediate stage HCC with a maximum of two HCCs <5cm. However, it is common practice in most participating centers to treat patients with HCC >3cm with combined transarterial chemoembolization and ablation. These patients are not eligible for inclusion and we thus expect that the vast majority of patients will have tumors <3cm. In post-hoc analysis, we will investigate whether differences in optimal MAM exist between patients treated with various ablation systems.

Furthermore, tissue contraction may pose an important challenge when interpreting our study results. Tissue contraction during ablation may result in calculated margins being smaller than they actually are. Currently, there is insufficient knowledge on how contraction is influenced by factors such as cirrhosis, tumor cellularity, ablation systems, power settings and ablation times. Most studies on tissue contraction have been performed in healthy ex-vivo animal livers. Brace et al. studied the difference in contraction for MWA compared to RFA based on multiple markers in sections of healthy unperfused ex-vivo bovine livers [13]. The mid and peripheral markers, placed at a distance of 10 and 15 mm from the ablation applicator respectively, showed a significant difference in contraction between RFA and MWA. This difference was not seen for the inner markers, placed at 5 mm from the ablation applicator. Two in-vivo animal model studies report a tissue contraction up to 12% [14, 15]. However, this was also in normal liver tissue. One retrospective in-vivo human study was performed by Lee et al [16]. In contrast to the study by Brace et al., they found a limited relative tumor and ablation zone contraction of -9.95% and -7.1%, respectively, for tumors treated with MWA [16]. The exact amount of tissue contraction in patients with HCC treated with thermal ablation remains unknown, may vary between patients and depends on liver consistency. However, as tissue contraction is present in all patients, it is indirectly taken into account in the cut-off value for the MAM.

Last, it is assumed that LRs for different tumors in the same patient are independent [4].

## References

1. European Association for Study of L, European Organisation for R, Treatment of C. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. *Eur J Cancer*. 2012;48(5):599-641.
2. Weis S, Franke A, Mossner J, Jakobsen JC, Schoppmeyer K. Radiofrequency (thermal) ablation versus no intervention or other interventions for hepatocellular carcinoma. *Cochrane Database Syst Rev*. 2013(12):CD003046.
3. Lencioni R, de Baere T, Martin RC, Nutting CW, Narayanan G. Image-Guided Ablation of Malignant Liver Tumors: Recommendations for Clinical Validation of Novel Thermal and Non-Thermal Technologies - A Western Perspective. *Liver Cancer*. 2015;4(4):208-14.
4. Tiong L, Maddern GJ. Systematic review and meta-analysis of survival and disease recurrence after radiofrequency ablation for hepatocellular carcinoma. *Br J Surg*. 2011;98(9):1210-24.
5. Kim YS, Lee WJ, Rhim H, Lim HK, Choi D, Lee JY. 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(3):758-65.
6. Laimer G, Schullian P, Jaschke N, Putzer D, Eberle G, Alzaga A, 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):2463-72.
7. Jiang C, Liu B, Chen S, Peng Z, Xie X, Kuang M. 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):1135-41.
8. Kim KW, Lee JM, Klotz E, Kim SJ, Kim SH, Kim JY, 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(5):W565-72.
9. Hendriks P, Noortman WA, Baetens TR, van Erkel AR, van Rijswijk CSP, van der Meer RW, et al. Quantitative Volumetric Assessment of Ablative Margins in Hepatocellular Carcinoma: Predicting Local Tumor Progression Using Nonrigid Registration Software. *Journal of Oncology*. 2019;2019:4049287.
10. Laimer G, Schullian P, Putzer D, Eberle G, Goldberg SN, Bale R. 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. *Int J Hyperthermia*. 2020;37(1):1362-7.
11. Sibinga Mulder BG, Hendriks P, Baetens TR, van Erkel AR, van Rijswijk CSP, van der Meer RW, et al. Quantitative margin assessment of radiofrequency ablation of a solitary colorectal hepatic metastasis using MIRADA RTx on CT scans: a feasibility study. *BMC Med Imaging*. 2019;19(1):71.
12. Park J, Lee JM, Lee DH, Joo I, Yoon JH, Park JY, et al. Value of Nonrigid Registration of Pre-Procedure MR with Post-Procedure CT After Radiofrequency Ablation for Hepatocellular Carcinoma. *Cardiovasc Intervent Radiol*. 2017;40(6):873-83.
13. Brace CL, Diaz TA, Hinshaw JL, Lee FT, Jr. Tissue contraction caused by radiofrequency and microwave ablation: a laboratory study in liver and lung. *J Vasc Interv Radiol*. 2010;21(8):1280-6.
14. Bressemer KK, Vahldiek JL, Erxleben C, Poch F, Shnaiyen S, Geyer B, et al. Exploring Patterns of Dynamic Size Changes of Lesions after Hepatic Microwave Ablation in an In Vivo Porcine Model. *Sci Rep*. 2020;10(1):805.
15. Erxleben C, Niehues SM, Geyer B, Poch F, Bressemer KK, Lehmann KS, et al. CT-based quantification of short-term tissue shrinkage following hepatic microwave ablation in an in vivo porcine liver model. *Acta Radiol*. 2021;62(1):12-8.
16. Lee JK, Siripongsakun S, Bahrami S, Raman SS, Sayre J, Lu DS. Microwave ablation of liver tumors: degree of tissue contraction as compared to RF ablation. *Abdom Radiol (NY)*. 2016;41(4):659-66.

## General discussion

In 1993 the first cases of in-human RFA for liver malignancy treatment were performed [1, 2]. Since then, the treatment technique has been constantly improved. At the beginning, thermal ablation was only an optional treatment in case resection or liver transplantation was contra-indicated. Technical developments and innovations improved the performance of RFA and broadened the use of other ablative technologies. Recently, thermal ablation is incorporated as first-line treatment for very early stage HCC in the guidelines [3]. Although the technique has already evolved during the last decades, there still lies potential in improving thermal ablations for liver malignancies. The current challenge is lowering the local recurrences for lesions >2 cm to compete with resection. Therefore, research in the field of thermal ablation is focused on advanced image processing, ablation margin quantification, optimized needle positioning and treatment planning to prevent local recurrences. The work presented in this thesis corresponds to this trend, as the rationale of all parts are based on increasing treatment success of thermal ablation.

The systematic review focuses on computational modeling for thermal ablation zone simulation. The impact of patient-specific variables, such as vascular proximity, perfusion rate and underlying liver diseases are addressed by computational modeling. Incorporation of patient-specific parameters could yield better accuracy in ablation zone dimensions prediction. However, patient-specific simulation is challenging due to the number of factors affecting the ablation zone.

Nowadays, treatment planning is based on the manufacturer's predictions and experience of the interventional radiologist, making it operator dependent. Vascular proximity is taken into consideration while determining treatment settings, but no other patient-specific parameters are involved. Advanced patient-specific ablation zone modelling could standardize the ablation treatment. However, accurate computer modelling of the ablation zone is still a bridge too far, but small improvements to the clinical practice could be achieved. Next to vascular proximity, other patient-specific parameters should be considered when selecting treatment settings. To realize this, the correlation between ablation zone sizes and these patient-specific parameters such as cirrhosis and perfusion should be established.

Our retrospective study performed on the correspondence of clinically obtained ablation zones with the manufacturer's prediction already indicate a sufficient overlap. When comparing the results of the retrospective study with the ex vivo validated computational models of the systematic review, smaller relative errors are found for the computational models (Figure 1). On the other hand, a better DSC was found in the retrospective study compared to the clinically validated computational models (0.73 vs 0.42 - 0.73), indicating computational modelling would not improve reliability yet. Limited data was available to investigate the effect of several patient-specific parameters on the reliability of manufacturer's predictions. Further clinical research should clarify if these patient-specific factors actually increase the reliability of ablation zone dimensions predictions.

The PROMETEUS study focuses on ablation margin quantification [4]. The ablation margin is known to be of importance for minimizing the local recurrence rate (LRR). In the last several years a lot of research have been performed to the minimal ablation margin (MAM) and ablation margin quantification [5]. Although technical solutions for ablation margin quantification already exists, they are still not implemented within standard clinical practice [6-8]. A clear relation between ablation margins and LR seems evident, but the precise relationship between ablation margins and risks of recurrence still needs to be established. With the PROMETEUS study it is aimed to correlate ablation margins with local recurrence to set the optimal threshold for MAM.

Next to determination of this correlation, the PROMETEUS study will also take another step towards implementation of ablation margin quantification software. The image analysis performed with our in-house software will be compared to the analysis using a commercially available software executed by the Radboud UMC. This validation will also generate knowledge into the interuser and

intersoftware reproducibility of MAM. A good reproducibility will lead to a step forward in clinical implementation of ablation margin qualification software.

The other focus in research of thermal ablations is optimized needle positioning using stereotactic thermal ablation and robotic needle positioning [9]. A systematic review, reviewing 34 articles, concludes that this advanced technology allows highly precise and safe tumor targeting, leading to enhanced primary treatment efficacy [10]. Comparative studies showed enhanced targeting accuracy and reduced probe readjustments compared to free-hand targeting [11, 12]. In addition, this technique enables ablation of large lesions following a multi-needle approach with 3D needle planning and precise stereotactic needle placement. The primary technical success rate in a retrospective study with 41 stereotactic liver ablations with a median lesion size of 9.0 cm was 80.5% [13]. This technique might overcome the size-related limitations of thermal ablation. Nevertheless, the number of coaxial needles used during stereotactic needle ablation was an independent predictor for major complications. These advanced needle placement techniques, in the end, rely on the manufacturer provided charts for ablation zone volume prediction. The full potential of this new technology can be exploited with incorporation of reliable patient-specific ablation zone predictions.

Scatterplot of the relative error in axis between created and predicted ablation zone

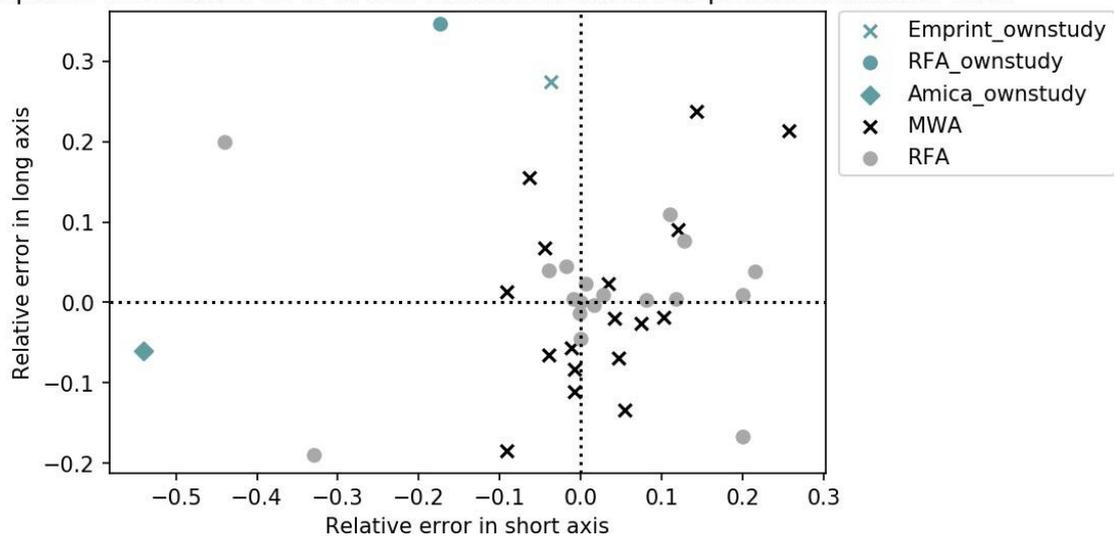


Figure 1. Scatterplot of the relative error in long and short axis diameter of the modelled MWA & RFA zones compared to ex vivo validation analyzed in the systematic review. In addition, the relative error in long and short axis diameter of the clinically obtained ablation zone compared to the manufacturer’s prediction are presented. In case of an experimental diameter of 30 mm, a relative error of 0.1 means the simulated diameter was 33 mm.

*Note: only the short axis diameter of the clinically realized diameter is included within this analysis leading to a negative relative error in the short axis for the Emprint cases of the thesis study.*

## References

1. van Sonneveld E, McMullen WN, Solbiati L. Tumor Ablation Principles and Practice Livraghi L, Mueller PR, Silverman SG, editors. New York: Springer New York; 2005. 542 p.
2. Gennaro N, Schiaffino S, Mauri G, Monfardini L. The What, the Why, and the How of Liver Ablations: A Practical Guide for the Medical Oncologist. *Oncology*. 2021;99(11):722-31.
3. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. *J Hepatol*. 2018;69(1):182-236.
4. Oosterveer TTM, van Erp GCM, Hendriks P, Broersen A, Overduin CG, van Rijswijk CSP, et al. Study Protocol PROMETHEUS: Prospective Multicenter Study to Evaluate the Correlation Between Safety Margin and Local Recurrence After Thermal Ablation Using Image Co-registration in Patients with Hepatocellular Carcinoma. *Cardiovasc Intervent Radiol*. 2022;45(5):606-12.
5. Minier C, Hermida M, Allimant C, Escal L, Pierredon-Foulongne MA, Belgour A, et al. Software-based assessment of tumor margins after percutaneous thermal ablation of liver tumors: A systematic review. *Diagn Interv Imaging*. 2022;103(5):240-50.
6. Hendriks P, Noortman WA, Baetens TR, van Erkel AR, van Rijswijk CSP, van der Meer RW, et al. Quantitative Volumetric Assessment of Ablative Margins in Hepatocellular Carcinoma: Predicting Local Tumor Progression Using Nonrigid Registration Software. *J Oncol*. 2019;2019:4049287.
7. Faber RA, Burghout KST, Bijlstra OD, Hendriks P, van Erp GCM, Broersen A, et al. Three-dimensional quantitative margin assessment in patients with colorectal liver metastases treated with percutaneous thermal ablation using semi-automatic rigid MRI/CECT-CECT co-registration. *Eur J Radiol*. 2022;156:110552.
8. Sandu RM, Paolucci I, Ruiters SJS, Sznitman R, de Jong KP, Freedman J, et al. Volumetric Quantitative Ablation Margins for Assessment of Ablation Completeness in Thermal Ablation of Liver Tumors. *Front Oncol*. 2021;11:623098.
9. Laimer G, Schullian P, Bale R. Stereotactic Thermal Ablation of Liver Tumors: 3D Planning, Multiple Needle Approach, and Intraprocedural Image Fusion Are the Key to Success-A Narrative Review. *Biology (Basel)*. 2021;10(7).
10. Tinguely P, Paolucci I, Ruiters SJS, Weber S, de Jong KP, Candinas D, et al. Stereotactic and Robotic Minimally Invasive Thermal Ablation of Malignant Liver Tumors: A Systematic Review and Meta-Analysis. *Front Oncol*. 2021;11:713685.
11. Heerink WJ, Ruiters SJS, Pennings JP, Lansdorp B, Vliegenthart R, Oudkerk M, et al. Robotic versus Freehand Needle Positioning in CT-guided Ablation of Liver Tumors: A Randomized Controlled Trial. *Radiology*. 2019;290(3):826-32.
12. Zhang Z, Shao G, Zheng J, Wen S, Zeng H, Hao W, et al. Electromagnetic navigation to assist with computed tomography-guided thermal ablation of liver tumors. *Minimally Invasive Therapy & Allied Technologies*. 2020;29(5):275-82.
13. Schullian P, Johnston EW, Putzer D, Eberle G, Laimer G, Bale R. Safety and efficacy of stereotactic radiofrequency ablation for very large ( $\geq 8$  cm) primary and metastatic liver tumors. *Scientific Reports*. 2020;10(1):1618.