Author's response to reviews

Title: Radiofrequency Ablation of Liver Lesions: Quantitative Assessment of Treatment Completeness through CT Image Processing.

Authors:

Katia Passera (katiamarina.passera@marionegri.it)
Sabrina Selvaggi (sabrina.selvaggi@mail.polimi.it)
Davide Scaramuzza (davide.scaramuzza@istitutotumori.mi.it)
Francesco Garbagnati (francesco.garbagnati@istitutotumori.mi.it)
Daniele Vergnaghi (daniele.vergnaghi@istitutotumori.mi.it)
Luca Mainardi (luca.mainardi@biomed.polimi.it)

Version: 3 Date: 14 June 2012

Author's response to reviews: see over
Reviewer's Answer

Title: Radiofrequency Ablation of Liver Lesions: Quantitative Assessment of Treatment Completeness through CT Image Processing.

Reviewer: Gerlig Widmann
Reviewer’s report:
Comments to authors
This is a very innovative and promising method for (semi-automated) quantification of treatment success if RFA of liver tumours. 5 metastases and 5 HCCs smaller than 20cm^3 were treated with a single RFA electrode. The paper showed that the US-guided single electrode RFA was not sufficient in most cases and recurrence has to be expected. The importance of an ablative margin of 1 cm cannot be overemphasized and multi-electrode position or multi-needle approaches are required for optimal RFA treatment of even small lesions. It would be of greatest interest to retrospectively evaluate RFA cases with the presented method and to look how accurate the method could predict recurrence.

Major Compulsory Revisions

1) Material and Methods:
Which phase (all or only one specific) of CT-imaging was used for the registration, segmentation and evaluation? This is of particular interest for HCC which has a characteristic enhancement pattern.

The phases depend on the nature of the lesions and their pattern before treatment. For HCC and hypervascular metastases arterial phase was considered, for hypovascular metastases portal or equilibrium phases were evaluated.

2) In the HCC cases arterial occlusion was performed. Was there a difference between mets and HCC? In fact, we do not perform this technique but rather use multi-electrode approaches.

In our experience arterial occlusion plus RFA in HCC gave better results than multi-electrode approach.

3) How fast were the registration procedure and the semi-automatic segmentation algorithm?

The segmentation method is quite fast, taking about ten minutes for the semi-automatic segmentation considering the live-wire technique and the picking of pixels. The registration takes about 40 minutes, considering the non-linear registration. These times were evaluated on a PC (Intel Pentium III). The use of a more powerful or dedicated PC could improve this performance.
4) The problem of fictitious deformations that alter the true overlap of an RFA lesion may not be totally solved. RFA of larger lesions may produce significant deformations of the original shape. What was the mean percentage of liver deformation between two data-sets?

We did not measure the percentage of deformation of the original shape, but we evaluated through a visual inspection by clinicians the reliability of our registration procedure for compensating these deformations. In none of the analysed cases this effect resulted to be problematic.

4) Numerical indices for the RFA evaluation:
In addition to the presented extremely valuable indices, an automated calculation of the largest axial diameter and difference of the HU-units (before and after RFA) would be an interesting additional information, in order to provide the surrogates for the modified RECIST criteria.

We thank the reviewer for the suggestion to add two additional indices, which could be investigated as surrogate for modified RECIST criteria. While those indexes could be easily computed on our images, we do not believe the paper will benefit by this information. In fact, we are interested in introducing a new technique and its related metrics. A detailed comparison of these indexes with other ones (included those suggested by the reviewer) is not the aim of this work and would be delegated to a future study.

**Reviewer: Nagy Habib**

Reviewer’s report:

- **Level of interest:** An article whose findings are important to those with closely related research interests
- **Quality of written English:** Acceptable
- **Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Reviewer: Eren Berber**

Reviewer’s report:

- **Major Compulsory Revisions**
This is a study aiming at determining objective criteria to evaluate the completeness of RFA. My comments include:

1. Although, the idea is novel, it seemed to be a complicated method to propose. In how many of these cases was the incomplete ablation obvious with conventional assessment? What is the proposed benefits of this methodology over the conventional assessment?

The conventional assessment is a visual assessment. There are some cases in which it is easy to see that the ablation is not complete (case 6), however it is not possible to objectively quantify it in all situations. The present method
would be of help to clinicians to better understand the spatial location of incomplete ablation and its amount. The method represents a first attempt to objectively quantify and localize the result of an ablation.

2. What was the clinical follow up of the lesions presumed to be incompletely ablated?

The purpose of this work was only to introduce a novel method for a quantitative assessment of RFA treatment. We applied the method to only 10 lesions (5 HCC and 5 metastasis), therefore to study the clinical follow up of these patients would lead to results which would be not statistically significant. The clinical validation will be postponed to a future study designed to assess efficacy of treatment with respect to the presented parameters with a larger dataset, grouped for type of lesions. Since this aspect of our study is not clear, we decided to stress it into the text (Study limitations).

**Reviewer: Christof M Sommer**

**Reviewer's report:**
The submitted manuscript documents an experimental work designed to evaluate the RFA completeness of focal liver lesions applying numerical quantification after CT image registration and segmentation before and after treatment. The study addresses an issue that is clinically very relevant. Overall readability and logical progression in the text are good.

My major points of criticism (from a radiological/oncological point of view) are:
1. basis of CT image processing are very thick slices (5mm), which are not used any more in our clinical routine. Although those images reduce image noise compared to thinner slices, partial volume effects are significant. Especially for the study described herein, with results in the sub dimensions of 1 mm (such as T.F.M.), this might be a major limitation.

The reviewer was right. At the time of study, the protocol in our hospital was limited to 5 mm slice thickness. We recognized that limitation in the discussion section. A thinner slice would have improved all the computation process.

Moreover, important details of the CT scan (such as pitch, kVp and mAs) and of the image reconstruction (such as increment, kernel, window/level) are not given. The contrast protocol is old fashioned (no bolus triggering, instead fixed acquisition of the different CT phases). Additionally, it is not clear, which CT phases were analyzed (for HCC, the arterial phase might be optimal, for metastasis the venous phase).

We added details of CT scans (pitch 1-1.5, 120 kVp, 250-300 mAs) and image reconstruction (increment 0.6 mm, reconstruction 1-5 mm, medium smooth kernel (i30F), window/level 300/40) within the text.

The phases depend on the nature of the lesions and their pattern before
treatment. For HCC and hypervascular metastases arterial phase was considered, for hypovascular metastases portal or equilibrium phases were evaluated.

2. the RFA technique is insufficient: antique needle designs and a weak generator (with a maximum power of only 50W) were used. Consequently, the relatively short ablation cycles (between 12-20 min) in your relatively big liver lesions seem to be rather short.

Thanks to the reviewer we noticed a text error, in fact the maximum power of the generator is 110W. We apologized for this error.

3. the oncological success of your ablations is frustrating. Only 2/8 ablations had no residual tumor, and not a single one showed the intended safety margin of 1 cm. The reasons for this have to be evaluated in detail. Who selected the RFA cases analyzed in this study? Was it a negative selection of cases, and especially problematic RFA cases were included to emphasize the need and benefit of a complex registration/segmentation approach? What was the experience of the RFA operators? How many RFA cases are performed per year in your center?

From a clinical point of view, it would be also very interesting to follow up the cases with incomplete tumor necrosis. Is it possible with your approach to predict the location of residual tumor/recurrence applying the given numeric quantifications? This would be a real benefit of your study with lasting clinical relevance. You should consider this in your revision.

The purpose of this work was to introduce a quantitative tool for the assessment of RFA treatment. We applied the method to only 10 lesions (5 HCC and 5 metastasis) that would be considered by clinicians very difficult cases with no other chance of alternative therapeutical approach. Our aim was to show the potentiality of the tool in some example cases, but not to perform such a clinical validation that the reviewer asks for. This type of validation would require a study designed for this aim, with a larger dataset grouped for type of lesions, in order to be statistically significant.

In this paper, we performed only a validation of segmentation step, that is the main step of the image processing and we postpone to a further work the clinical validation. Since this aspect of our study is not clear, we decided to stress it into the text (Study limitations).

4. please include the detailed times necessary to realize the different steps for CT image processing (such as pre-processing, clustering, etc.).

The segmentation method is quite fast, taking about ten minutes for the semi-automatic segmentation considering the live-wire technique and the picking of pixels. The registration takes about 40 minutes, considering the non-linear registration. These times were evaluated on a PC (Intel Pentium III), The use of a more powerful or dedicated PC could improve this performance.

5. specific techniques (such Fuzzy-C-means approach, non-linear
B-splines-based algorithm, Live-Wire algorithm, etc.) have to be explained in such a manner, that also non-familiar experts have the chance to understand the procedure and to follow the central theme.

The mentioned approaches are quite standard and widely described in literature. For the sake of the reviewer, we included some information about them in this reply (see below). However, we think that the inclusion of this information into the text will make it less fluent and distract readers' attention from the main message for the manuscript. If the reviewer believes that these information are really necessary, we will add them into an appendix at the end of the manuscript.

**Live-wire:**
Livewire is a semi-automatic segmentation technique also known as Intelligent Scissors. The method is based on the lowest cost path algorithm (11): the image is firstly convolved with a Sobel filter to extract edges. Using this filtered image a graph is created. Each pixel of the resulting image is considered as a vertex of the graph and has edges going to the 4 pixels around it, as up, down, left, right. Edges are weighted with features gathered from the sobel filter making it less costly to stay on an edge. The edge costs are defined based on a cost function.
The user simply sets the starting point clicking on an image's pixel, known as an anchor. Then, as he starts to move the mouse over other points, the smallest cost path is drawn from the anchor to the pixel where the mouse is over, changing itself if the user moves the mouse. If he wants to choose the path that is being displayed, he simply clicks the image again.

**Fuzzy C-means algorithm:**
Fuzzy-C-means is a unsupervised clustering algorithms. These algorithms classify a given data set through a certain number of clusters fixed a priori. The fuzzy-C-means clusters data by iteratively computing a mean image intensity for each cluster and segmenting the image by classifying each pixel in the cluster with the closest mean. The fuzzy c-means algorithm generalizes the K-means algorithm, allowing for soft segmentations, that is it allows pixels to belong to two or more clusters.

**B-spline free-form deformation algorithm**
Given a reference and a template image, a deformation algorithm find a spatial transformation such that the deformed template matches the reference image subject to a suitable similarity measure. In this work the transformation was expressed by B-splines and the similarity measure was normalized mutual information.

6. P5, second paragraph: Once the ...
What are the different intensity patterns for necrosis and lesion – are they identical for HCC/metastasis in all cases? From my experience, it is sometimes extremely difficult to characterize a focal change in liver density as post-RFA lesion or as a tumor (not to mention as a HCC or as a metastasis). This point should be explained and discussed more detailed, inclusive of additional references.
The reviewer was right; it is extremely difficult to characterize a focal change in liver density and this was in fact the most complicated part of the work. The rationale was to use some reference pixels and the results of fuzzy-C-means clustering.

Using the fuzzy-C-means algorithm we classify each pixel of the image as belonging to a cluster of the 7 possible clusters (this number will find after a tuning of the method). Cluster indexes range from 1 (the darkest) to 7 (the brightest). By inspection of clustered images we have noticed that both metastasis and HCC have a ring concentric structure. In the case of HCC the center is brightest and the periphery tends to have rings that progressively decrease in intensity. Metastasis on the contrary has a darkest center with respect to periphery. We used the mean of cluster indexes of selected reference pixels to identify the type of lesion. If the mean of cluster indexes of selected reference pixels is superior to 3.5, clusters involved are the brightest, therefore it is an HCC with the center with the highest index and decreasing cluster indexes for the periphery (see examples in figure 4). On the contrary, if the mean of cluster indexes of selected reference pixels is inferior to 3.5, clusters involved are the darkest, therefore it is a metastasis.

In all cases we have analysed this method efficiently works as shown by segmentation validation.

7. the importance of the inter-barycentric distance:
How was the barycenter defined (or calculated)? It is quite obvious, that an inter-barycentric distance with the same magnitude order of the target lesion indicates insufficient positioning of the RFA lesion. However, can the authors give numerical data (e.g. an equation) for adequate and non-adequate inter-barycentric distances with respect to oncological success?

The index is computed as the absolute value of the difference between lesion and necrosis barycentre. The reviewer asks for a numerical data to assess if this distance is adequate or not. In the ideal situation the distance should be 0 (that is the lesion and necrosis are fully superimposed), the comparison with the dimension of the original lesion gives, on the other hand, an idea of misalignment.

We analysed a limited population only. Thus we are not able to conclude anything about the relationships between inter-barycentric distances and oncological success. This aspect will be investigated in a successive study involving a larger dataset as mentioned before.

My minor points of criticism (from a radiological/oncological point of view) are:
1. P2, first paragraph: Image guided ...
   Please point also out, that RFA is a curative alternative to surgery in circumscribed tumors <3cm.

   We added this sentence into the text.

2. P2, fifth paragraph: After validation ...
   Point out the selection criteria (and expertise) for the RFA cases included in
We decided to explain better within the text (method section) that the selected cases were very difficult cases and that we would represent only as example of tool capability.

3. P2, sixth paragraph: The method ...
Please indicate also max. lesion diameters, not only volumes.

Lesion diameter range = 2.5 - 4 cm. Only one lesion is superior to 3 cm. We added this information within the text.

4. P3, fifth paragraph: For HCC ...
“11.5mm balloon” is unclear for me (please indicate diameter and length).

Diameter 11.5 mm, length 2 cm. We added this information within the text.

5. P4, first paragraph: Image noise ...
What is a 5 x 5 median filter – please explain.

The median filter uses a template mask large 5X5 pixels in order to calculate the median locally. It replaces each pixel value with the median of neighbouring pixel values (computed on a mask of 5X5 pixels).

6. Please use the same nomenclature throughout the entire manuscript: e.g. post-RFA necrosis and original liver lesion! (to minimize confusion).

The reviewer was right, we correct the manuscript using the same nomenclature.

7. P5, second paragraph: Liver clustering ...
Only one radiologist was involved? I would recommend to involve three radiologists and compare also their results!

We added into the text that parameters of clustering algorithm was tuned. Normally an expert radiologist recognizes 7 different tissues in this type of images. However, we tuned the method with respect number of cluster in a range of values (6-8) and we selected the best parameter. Following, the results of tuning tests (the graphs show that with cluster number equal to 7 we obtained the best results in terms of positive predictivity/percentage match and negative predictivity/percentage match).
8. P7, second paragraph: necrosis: ...
“classify them iteratively” – unclear for me, please explain more detailed.

In the case of hyperdense areas into the necrosis, these areas can be recognized by the selection of the reference pixels that have a cluster index greater than 2 (1 and 2 are the darkest cluster). These areas are then re-classified with lower indexes (1 or 2, the usual indexes for necrosis) in order to segment the entire necrosis region.

9. P8, third paragraph: The effect...
I wished you had treated more healthy liver tissue in your patients. Then probably, also the tumors might have been destroyed completely! From the oncological point of view, in case of doubt, we should sacrifice healthy liver tissue to cover the tumor completely.

The 19X19 pixels is the mask used before non linear alignment. This operation was performed both for the pre-RFA and the post-RFA images, independently. We take a square 19X19 pixel of healthy liver tissues, this mask was replicated with radial padding (see figure below) in order to cover ROI area that is different in the case of lesion and necrosis.
The same operation is performed for the post-RFA image in the area of necrosis. Figure 5 of the manuscript shows the resulted images after substitution of synthetic patterns and figure 6 shows the need of this operation in order to avoid problems with the non linear alignment.

10. Results paragraph:
Please include also the detailed CT image processing times.

We added into the text this information.

11. P11, fourth paragraph: The inter-barycentric ...
Who defined the critical range of 60-120 degrees for O.I.? Please explain the rational, or give a reference instead. We know RFA or microwave systems, which ablate in other configurations (non-elliptical). Please comment on this.

To compute the orientation index an ellipse was fitted to each ROI in each slice containing the lesion and the same for the necrosis. This is the common way to identify the major axis of a region of interest and so to identify its orientation. It is not needed that the ablation configuration is elliptical. The fitting simply identify the direction in which the lesion or necrosis mainly develop, therefore the angle between major axis of lesion and major axis of necrosis indicates if the ablation is able to follow or not the configuration of lesion. The range of 60-120 degrees is defined by us as critical. The ideal value is 0 or 180 (parallel orientation), while moving away (e.g 90 degrees) means that lesion and necrosis develop in opposite direction.

12. P15, first paragraph: Table 2 ...
“Lesion volume range of 1.5-9.5cm³” – I cannot reproduce those numbers in Table 2?? Please explain.
It is an error of the text. Now it is substituted by the right value (1.5 to 19.4).

13. Table 2:
Annotation: “post-RFA lesion – some people might misunderstand this measure as the volume of the necrosis, and not as the residual tumor volume.

The reviewer was right. We had a note in the table.

14. P 17, second paragraph: There were ...“the short interaction time” – you should not discuss an issue, which was not evaluated or indicated in the “results” or “materials and methods” section.

We explained in the method section that the segmentation method requires only 10 minutes, so now the sentence could be put in the discussion section.

15. P 17, fourth paragraph: There are ...
Which are “the vital anatomical structures“ in the context of malignant liver lesions? – I would be as aggressive as possible to destroy the tumor completely!
You have also to decide prior to intervention, if you perform a curative or palliative RFA! Please clarify and discuss.

We considered vital anatomical structures organs and tissues near the liver, such as bile ducts, cholecyst., and organs next to peripheral lesions (e.g. bowel loops, heart) at risk of being injured.
The considered cases were really complicated. Therefore, during RFA the plan could have been changed to avoid the involvement of this vital structure. This could also explain why it was not always possible to ablate with the safety margin.

16. P 17, last paragraph: Besides, there ...
The last sentence on the page is very interesting ("Thus, in this RFA ...”).
Please indicate some references, and also in this context, the question of local recurrence or residual tumor is very interesting. Please implement the follow-up findings in your work.

The last sentence is our own hypothesis. There is no reference in literature. In order to avoid misunderstanding we added into the text that this sentence is our hypothesis.

In summary, the paper describes an interesting experiment based on an important hypothesis. Language style is good. Presentation is good. The findings are interesting and have potential clinical relevance. From a radiological/oncological point of view, the study has some major limitations (especially CT technique and reconstruction technique (which is the entire basis for image processing with registration and segmentation), RFA technique and the very bad oncological success rate). Therefore, it is
unacceptable for me to recommend “manuscript acceptance”, also because people might get a completely wrong impression of the power of RFA in this manuscript/as presented in this manuscript.

We understand the Reviewer's point. To avoid any false messages, we clarify into the method section that the studied lesions were selected among those suspected not to be completely treated. This has been done in order to better evidence the potentiality of the proposed method.