research article

MRI reduces variation of contouring for boost clinical target volume in breast cancer patients without surgical clips in the tumour bed

Noora Al-Hammadi, Palmira Caparrotti, Saju Divakar, Mohamed Riyas, Suparna Halsnad Chandramouli, Rabih Hammoud, Jillian Hayes, Maeve Mc Garry, Satheesh Prasad Paloor, Primoz Petric

Department of Radiation Oncology, National Center for Cancer Care and Research, Hamad Medical Corporation, Doha, Qatar

Correspondence to: Primož Petrič, M.D., M.Sc., Department of Radiation Oncology, National Center for Cancer Care and Research, Hamad Medical Corporation, P.O. Box 3050, Doha, Qatar. E-mail: ppetric@hamad.qa

Background. Omitting the placement of clips inside tumour bed during breast cancer surgery poses a challenge for delineation of lumpectomy cavity clinical target volume (CTV_{LC}). We aimed to quantify inter-observer variation and accuracy for CT- and MRI-based segmentation of CTV_{LC} in patients without clips.

Patients and methods. CT- and MRI-simulator images of 12 breast cancer patients, treated by breast conserving surgery and radiotherapy, were included in this study. Five radiation oncologists recorded the cavity visualization score (CVS) and delineated CTV_{LC} on both modalities. Expert-consensus (EC) contours were delineated by a senior radiation oncologist, respecting opinions of all observers. Inter-observer volumetric variation and generalized conformity index (CI_{gen}) were calculated. Deviations from EC contour were quantified by the accuracy index (AI) and inter-delineation distances (IDD).

Results. Mean CVS was 3.88 +/- 0.99 and 3.05 +/- 1.07 for MRI and CT, respectively (p = 0.001). Mean volumes of CTV_{LC} were similar: 154 +/- 26 cm³ on CT and 152 +/- 19 cm³ on MRI. Mean CI_{gen} and AI were superior for MRI when compared with CT (CI_{gen}: 0.74 +/- 0.07 *vs.* 0.67 +/- 0.12, p = 0.007; AI: 0.81 +/- 0.04 *vs.* 0.76 +/- 0.07; p = 0.004). CI_{gen} and AI increased with increasing CVS. Mean IDD was 3 mm +/- 1.5 mm and 3.6 mm +/- 2.3 mm for MRI and CT, respectively (p = 0.017).

Conclusions. When compared with CT, MRI improved visualization of post-lumpectomy changes, reduced inter-observer variation and improved the accuracy of CTV_{LC} contouring in patients without clips in the tumour bed. Further studies with bigger sample sizes are needed to confirm our findings.

Key words: breast cancer; contouring; contouring variation; MRI; CT

Introduction

There are two clinical scenarios in which delineation of the lumpectomy cavity (LC) is required during breast cancer radiotherapy: boost after whole breast irradiation (WBI) and accelerated partial breast irradiation (APBI). WBI after organ-sparing surgery reduces the risk of breast cancer recurrence and mortality.[1,2] Delivery of boost dose to the LC clinical target volume (CTV_{LC}) is an important component of this treatment. It has been

shown to improve local control at an increased risk of moderate to severe fibrosis.[3] APBI is becoming increasingly utilized in selected groups of patients. Shorter overall treatment time, reduced radiation exposure of the organs at risk and comparable disease control make it a good alternative to WBI for early stage disease.[4-6] High accuracy of contouring and precision of treatment delivery are needed to optimize the delicate therapeutic ratio between treatment benefit and side effects. This is especially important in the setting of highly conformal dose

delivery to a small volume, such as boost after WBI and becomes critical during APBI where the entire dose is delivered to the CTV_{LC}. Inter-observer variation (IOV) in contouring is one of the main contributors to the cumulative budget of uncertainties in radiotherapy.[7] It may undermine the gain of high-precision technologies, blur the dose-effect relations and compromise treatment comparisons. For the individual patient, geographical miss of the target volume leads to increased chance of relapse, while unnecessary irradiation of normal tissues increases the probability of side effects. Respecting the common contouring guidelines accompanied by adequate training and high quality imaging are the most important strategies to reduce contouring variation.[7-11]

Currently, CT is the standard imaging modality for CTV_{LC} contouring. Due to its poor ability for soft tissue depiction, placement of surgical clips or markers at the edges of LC is recommended to improve tumour bed delineation.[12-16] But reliability of inserted markers as a surrogate for tumour bed is a matter of debate[9,16-18] and omission of their placement in some patients poses a special challenge to the radiation oncologist during CTV_{LC} delineation.[19] The role of MRI for contouring in breast cancer radiotherapy is controversial[11,20-23] and the evidence to support its use in patients without markers in the tumour bed is scarce.[24,25] In our present study, we aimed to (1) quantify the IOV and (2) assess the accuracy of CT- and MRI-based CTV_{LC} contouring in patients without clips in the LC. Our null hypothesis was that there is no statistically significant difference between MRI- and CT- based contouring in this subgroup of patients.

Patients and methods

Patients and images

Anonymized image data sets of patients with pathology-proven unilateral invasive ductal carcinoma of the breast, treated by breast conserving surgery and adjuvant radiotherapy in 2013 were considered for this study. Cases without surgical clips in the LC and available CT- and MRI-simulator data sets were eligible for inclusion. Adjuvant radiotherapy had to consist of WBI followed by CTV_{LC} boost. Patients who underwent oncoplastic surgery were excluded. All radiotherapy was completed before initiation of the study and the presented work did not interfere with routine management of our patients. The study protocol was reviewed and given ethical approval by the Institutional Medical Research Centre which governs our Institutional Review Board (Trial registration number: 15329/15).

Acquisition of CT and MRI simulator images

During simulation and treatment, patients were placed in comfortable and reproducible supine position with arms abducted over the head. For CT simulation patients were placed on breast board and wires were used to identify the surgical scars and drainage sites. Non-contrast volumetric CT study with contiguous slices of 5 mm thickness was obtained from the level of the body of the mandible to at least 5 cm below the inframammary fold (Siemens Somatom Sensation ® 16-slice scanner, 120 kVp, approximately 90 mAs, voxel size of 1.26 x 1.26 x 5 mm, matrix size of 512 x 512). MR images were obtained on a dedicated wide-bore 1.5T 450w MRI simulator (General Electrics Optima ®) equipped with radiotherapy applications. The MRI in this study was a simulation procedure and was acquired supine as per CT planning with efforts made to replicate the positioning as much as achievable. The arms were elevated and cradled, and external alignment lasers used to align the tattoos, albeit the incline was not applied due to limitations of the MRI bore diameter. The supine positioning achieved a more similar deformation of the breast tissue to the planning CT than a prone diagnostic arrangement. General purpose Flex coils were used. Our breast MRI protocol included T2 weighted FSE propellor, proton density with fat saturation, Dixon type LAVA-Flex and balanced steady state gradient echo FIESTA imaging sequences. All sequences were acquired axially with matrix size of 288 x 288, approximately 42 cm field of view and slice thickness of 5 mm. For the T2 FSE sequence, mean system related geometric distortions after the application of the vendor-provided correction algorithms were 0.5, 0.9 and 1.9 mm for radial distances of 100, 200, and 250 mm respectively. Anonymized non-registered CT and MRI data-sets were imported to the ECLIPSE workstation (Varian, Medical Systems ®) for contouring.

Cavity visualization score and contouring

Cavity visualization score (CVS) was recorded by each observer for all cases and both modalities, using the standardized numeric scale ranging from 1 (cavity not visualized) to 5 (all cavity margins

clearly visualized).[26] CTV_{LC} was contoured separately on CT and MRI by five experienced radiation oncologists (observers), who were blinded for each other's delineations. The observers had access to clinical and imaging findings at time of diagnosis and to surgical and pathology reports. They were asked to respect the following instructions during delineation:

1. Adjust window level to optimize visualization of the region of interest.
2. Contour on axial images.
3. When contouring on the MRI, use the T2 weighted FSE images as primary data set and take the information from other sequences into account.
4. Allow for a minimum interval of 2-weeks between CT- and MRI-based contouring to minimize bias resulting from familiarity with the cases.
5. Create CTV_{LC} according to our departmental guidelines:
   A. First, delineate the lumpectomy cavity (LC) as intra-mammary post-lumpectomy changes. During delineation, compare findings with contralateral anatomy to identify differences in geometry, tissue architecture, formation of seroma, hematoma or scar tissue, fat replacement on CT and decreased signal intensity on MRI. While contouring, take all available information into account to identify the LC (tumour location on preoperative imaging, pathology reports, lumpectomy scar on the skin, etc.).
   B. To define CTV_{LC}, add a 15 mm uniform margin around the LC and edit it to exclude the chest wall and skin.

Finally, the expert consensus (EC) contours of CTV_{LC} were delineated on CT and MRI for all cases. EC contouring was led by the senior radiation oncologist, taking the opinions of all five observers into account.

Analysis of contouring uncertainties

Contouring uncertainties on CT and MRI were analysed from two perspectives, reflecting our study objectives: (1) to quantify the IOV, global variability between delineations was assessed and (2) to quantify contouring accuracy, deviations of observers from the EC contours were analysed. Contour analysis tool 1 (CAT 1) software and related methodology[27,28] was used for volumetric and distance-based computations.

Inter-observer variation: Mean volumes and standard deviations (SD) of CTV_{LC} were calculated for each study case on CT- and MRI-based approach. Inter-observer coefficients of variance (CoV – ratio between SD and mean value) and ratios between the smallest and largest volume were determined for each case and modality. Inter-observer conformity index was calculated based on the generalized formalism (CI_{gen}), which is independent of the number of the analysed volumes.[29] It equals the sum of intersections of all possible volume-pairs divided by the sum of their unions.

Contouring accuracy: We used the EC as a surrogate for the "ground truth" contour. Deviations from EC were measured on CT and MRI for all cases and observers. Accuracy index (AI) was determined according to paired CI formalism.[29] AI was calculated as the ratio between common and encompassing volume for each pair of EC and observer's contour. Further, mean absolute distances between contours of individual observers and EC were calculated in contouring plane. This method has been used before and is described in detail.[27,28,30] Briefly, the inter-delineation distances (IDD) were calculated between each voxel of observer's contour and nearest voxel of the EC contour in 72 angular steps of 5 degrees for all slices.[27,28,30]

Statistical analysis

Statistical design of the study did not entail calculation of the sample size and the number of observers. Instead, all evaluable cases satisfying the inclusion criteria to the point of study initiation and all available observers from our department were included to maximize the statistical power. Continuous variables were presented as mean values with standard deviations. Paired sample t-test was used to compare mean values of analysed variables between CT and MRI. P-value of < 0.05 was considered as the limit for statistical significance. SPSS for windows (*SPSS Inc., 1989–2015, Chicago, Illinois) was used for data analysis.

Results

Cavity visualization score

The use of MRI improved the cavity visualization in 11 out of 12 (92%) cases (Figure 1). In the remaining one case, mean CVS was equal (3.0) on both modalities. Mean CVS was 3.88 +/- 0.99 and 3.05 +/- 1.07 for MRI and CT, respectively (p = 0.001). Correlation of CI_{gen} and AI with CVS is shown in

Figure 1. CI_{gen} and AI improved with increasing CVS for both contouring approaches. Example of contouring variation for two selected cases with a high and low CVS is presented in Figure 2.

Inter-observer variation

The results of IOV analysis are presented in Table 1 and Figure 1A. Mean CI_{gen} for MRI was significantly superior to CI_{gen} for CT (0.74 +/- 0.07 $vs.$ 0.67 +/- 0.12, p = 0.007). CI_{gen} for MRI was higher than for CT in 10 (83 %) cases. In case number 9, CT-based CI_{gen} was superior to MRI (0.77 $vs.$ 071) and in case number 6 they were identical (0.76). Mean volumes of CTV_{LC} were 154 +/- 26 cm^3 on CT and 152 +/- 19 cm^3 on MRI (non-significant difference). Mean volumetric CoV was non-significantly lower for MRI when compared with CT (12% $vs.$ 18 %; p = 0.1). Similarly, average ratio between the smallest and largest delineated volume was non-significantly higher for MRI when compared with CT (0.8 +/- 0.1 $vs.$ 0.7+/- 0.1.; p = 0.1).

Contouring accuracy

Results of analysis of deviations from EC contours are shown in Table 2 and Figure 1B. Observers placed all contours in the correct breast quadrant. Mean AI was higher for MRI when compared with CT (0.81 +/- 0.04 $vs.$ 0.76 +/- 0.07; p = 0.004). MRI-based mean AI was superior to CT in 10 (83 %) cases. In case number 9, CT-based AI was slightly superior to MRI (0.81 +/- 0.04 $vs.$ 0.8 +/- 0.05) and in case number 1, AI was the same for both modalities (0.88 +/- 0.1) (Table 2, Figure 1B). There was small but significant difference in mean IDD between CT and MRI (3.6 mm +/- 2.3 mm $vs.$ 3 mm +/- 1.5 mm; p = 0.017). Corresponding mean CoV for CT was higher than for MRI (61 % $vs.$ 49 %; p = 0.003). The mean value of maximal IDD was 13 +/- 6 mm for CT and 10 +/- 4 mm for MRI (p = 0.06).
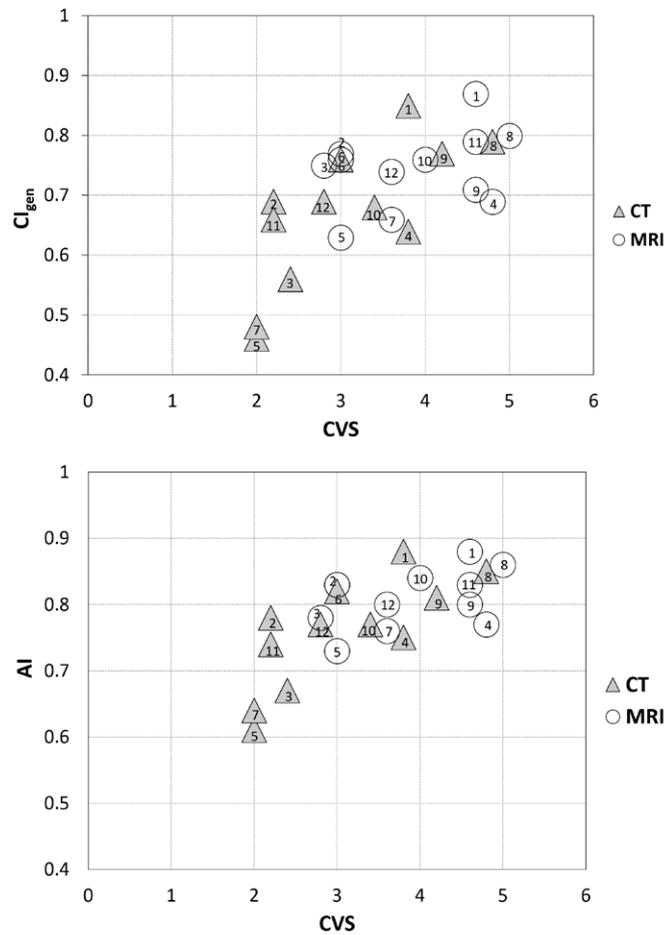

FIGURE 1. (A) Generalized conformity index (CI_{gen}) and **(B)** accuracy index (AI) as a function of the cavity visualization score (CVS) for CT and MRI based contouring of lumpectomy cavity clinical target volume. None of the patients had surgical clips inserted in the tumor bed. Case numbers are indicated for each modality.

FIGURE 2. CT and MRI based contouring in two examples with high and low cavity visualization scores (CVS). Observers' delineations are white and expert consensus (EC) contours black. **(A)** Case with a CVS of 4.8 on CT and 5 on MRI: mean generalized conformity index (CI_{gen}), accuracy index (AI) and inter-delineation distance (IDD) were 0.79, 0.85 and 2.4 mm on CT and 0.80, 0.86 and 2.2 mm on MRI. **(B)** Case with a CVS of 2 on CT and 3 on MRI: mean CI_{gen}, AI and IDD were 0.46, 0.61 and 6 mm for CT and 0.63, 0.73 and 4.5 mm for MRI.
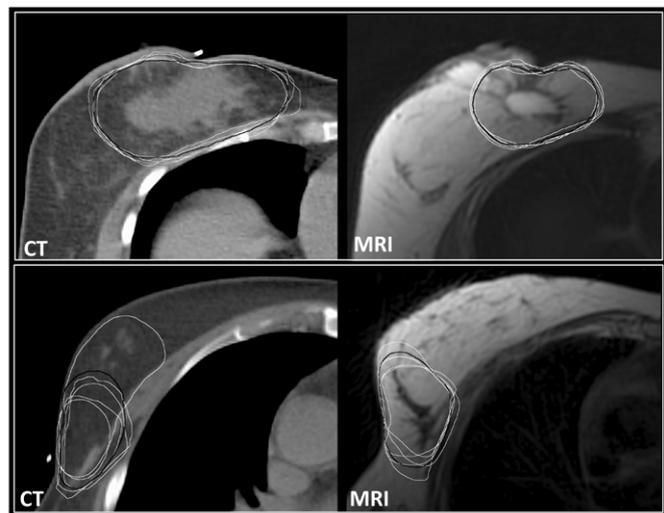
TABLE 1. Results for inter-observer variation in contouring. The difference in mean generalized conformity index (CI$_{gen}$) between the CT and MRI based contouring was statistically significant (p = 0.007)

Case	CT			MRI		
	Mean V [cm3] (SD)	CoV [%]	CIgen	Mean V [cm3] (SD)	CoV [%]	CIgen
1	198 (10)	5	0.85	106 (4)	4	0.87
2	241 (37.5)	16	0.69	256 (28)	11	0.77
3	108 (20.7)	19	0.56	125 (10)	8	0.75
4	75 (18.1)	24	0.64	92 (12)	13	0.69
5	175 (47.9)	27	0.46	217 (68)	31	0.63
6	140 (10.1)	7	0.76	125 (14)	11	0.76
7	103(39.7)	39	0.48	64 (2)	4	0.66
8	180 (22.5)	12	0.79	158 (14)	9	0.8
9	135 (19.8)	15	0.77	126 (20)	16	0.71
10	204 (44.3)	22	0.68	215 (19)	9	0.76
11	99 (14.4)	15	0.66	135 (11)	8	0.79
12	195 (27.3)	14	0.69	210 (29)	14	0.74
MEAN (SD)	154 (26)	18	0.67 (0.12)	152 (19)	12	0.74 (0.07)

CoV = Coefficient of Variance; CTV$_{LC}$ = Clinical Target Volume of Lumpectomy Cavity; SD = Standard Deviation

Discussion

Results of the present study rejected our null hypothesis: MRI, when compared with CT, led to (1) reduced IOV and (2) improved accuracy for CTV$_{LC}$ contouring in patients without markers in the tumour bed. Keeping in mind the cost and complexity of utilizing MRI for radiotherapy planning, our findings justify its use in selected cases.

CT-based delineation of the LC is prone to IOV, even among experienced radiation oncologists.[7,10,11,31-34] In various tumour sites, MRI has been shown to reduce contouring uncertainties when compared with CT.[35-41] Based on these findings, MRI is becoming increasingly implemented for contouring and is the recommended gold standard in some malignancies.[42] However, there are many studies that failed to demonstrate improved contouring with the use of MRI for various tumour sites.[43-47] As far as breast cancer is concerned, several authors investigated the impact of adding MRI to CT for delineation of lumpectomy cavity with negative or inconclusive outcome.[11,16,20-23] Den Hartogh *et al.* found that addition of postoperative MRI to CT guided delineation marginally increased the target volumes and failed to reduce the IOV.[22] Similarly, Kirby *et al.* reported that addition of MRI to CT resulted in tumour bed volumes that were discordant with those based on CT and

TABLE 2. Accuracy index (AI) and inter-delineation distances (IDD), based on the expert consensus (EC) delineation as the reference. The differences in AI and IDD between CT and MRI were statistically significant (p < 0.05)

Case	CT						MRI					
	IDD [mm]			AI			IDD [mm]			AI		
	Mean	SD	CoV [%]	Mean	SD	CoV [%]	Mean	(SD)	CoV [%]	Mean	SD	CoV [%]
1	2.1	0.8	38	0.88	0.02	3	1.8	0.58	32	0.88	0.01	2
2	3.6	1.9	53	0.78	0.1	12	3.2	1.9	59	0.83	0.02	3
3	3.9	2.6	67	0.67	0.07	10	3.2	1.8	56	0.78	0.02	2
4	2.3	1	43	0.75	0.06	8	2.3	0.9	39	0.77	0.05	6
5	6	4	67	0.61	0.15	24	4.5	2.1	47	0.73	0.14	20
6	2.6	1.1	42	0.82	0.02	2	2.6	1	38	0.83	0.02	2
7	6	4.1	68	0.64	0.09	14	3.6	1.8	50	0.76	0.05	7
8	2.4	1.2	50	0.85	0.02	2	2.2	1	45	0.86	0.04	4
9	3	2.9	97	0.81	0.04	5	2.9	1.8	62	0.80	0.05	6
10	4.5	3.5	78	0.77	0.07	9	3.3	2	61	0.84	0.02	2
11	3.3	2	61	0.74	0.09	12	2.5	1.3	52	0.83	0.02	3
12	3.4	2.3	68	0.77	0.06	8	3.6	1.7	47	0.80	0.02	2
MEAN	3.6	2.3	61	0.76	0.07	9	3	1.5	49	0.81	0.04	5

CoV = Coefficient of Variance; CTV$_{LC}$ = Clinical Target Volume of Lumpectomy Cavity; SD = Standard Deviation

clips alone. With the use of MRI, the tumour bed volume increased in 28 out of 30 cases included, resulting in a median CTV increase of 10.3% (-33.6%–80.9%).[20] Mast *et al.* compared CT- and MRI-based delineations of breast and LC by four observers in 10 patients. The mean CI for the LC was 0.52 for CT and 0.48 for CT combined with MRI (p = 0.33).[23] In another similar study, the inter-observer agreement was even lower. While MRI and CT enabled similar visualization of the LC, MRI resulted in lower generalized CI (0.32 +/- 0.25) when compared with CT (0.52 +/- 0.21).[21]

The rationale to use MRI in our study was to improve contouring consistency for cases without surgical clips in the tumour bed. Mean CVS on MRI (3.88 +/- 0.99) was significantly superior to CVS on CT (3.05 +/- 1.07) (p = 0.001). CVS was improved in 92% and was accompanied by an increase of CI_{gen} and AI in 83% cases. For both modalities, we found an increase of CI_{gen} and AI with increasing CVS (Figure 2). Therefore, inter-observer concordance depended directly on the ability to visualize lumpectomy cavity, which was superior on MRI. Of note, in all of the reports which failed to show benefit of MRI, clips were placed at the edges of LC.[11,16,20-23] In a study by Giezen *et al.*, four observers (2 radiologists and 2 radiation oncologists) obtained a mean CVS of 2.8 +/- 1.7 for MRI and 2.9 +/- 1.7 for CT. In contrast to our findings, Giezen *et al.* demonstrated superiority of CT over MRI for contouring, especially at low CVS.[21] With increasing CVS values, both modalities performed better and the CI_{gen} from MRI approached that from CT. The lack of added value of MRI in this and other published studies[20-23] could be attributed to better visibility of the clips on CT, introducing a bias in its favour, as acknowledged by the authors.[21] This effect becomes especially important at low CVS values. Our positive findings could be attributed also to the fact that MRI was performed as simulation procedure, replicating the CT planning supine position as much as achievable.

To our knowledge, there are only two publications in addition to our present study which demonstrated added value of MRI for delineation of post-lumpectomy tumour bed.[24,25] In the study by Jolicoeur *et al.*, there were no surgical clips implanted at time of lumpectomy. Three observers delineated the post-lumpectomy tumour bed in 70 patients. Highly significant IOV was demonstrated for CT based contouring of the tumour bed (p < 0.0001), while agreement was high for the MRI-based approach. The volumes of MRI based contours were 30–40% smaller than the CT-derived

volumes. In another study with three observers and 36 cases, mean CVS for the LC was 3.3 and 4.3 for CT and T2 MRI, respectively (p < 0.0001). Better CVS was reflected in superior inter-observer consistency and volumetric agreement of contours. The authors stated that surgical clips were occasionally, but not routinely placed by the referring surgeons.[24]

Based on our results, addition of MRI to CT could be justified as a good alternative to CT alone for selected patients in whom the placement of surgical clips in the tumour bed was omitted. But despite concerns regarding their reliability as a surrogate for tumour bed[17,18,48], placement of clips followed by CT-based contouring of LC should be currently considered as the gold standard.[16] This approach has been shown to improve the accuracy of LC contouring, reduce the overall boost volume and help prevent geographical miss and underdosage of the LC.[13-16,49-52] But the technique of placement and the number of inserted markers differs between institutions and surgeons and is even omitted in some cases. Kirwan *et al.* recently reported on a retrospective study of 196 cases, assessing the compliance with recommendations for clip insertion. Although recommended by the clinical guidelines, the clip insertion was omitted in 56% of cases while additional 7% of patients had only two or fewer clips inserted. Ten of 31 referring surgeons routinely omitted clips and the omission rate was significantly higher for centres with low (≤ 1 patient) when compared with high (≥ 14 patients) rate of recruitment to IMRT clinical trials (67% *vs.* 27%, respectively; p < 0.001).[19] These results emphasize the need for good collaboration between radiation oncologists and surgeons and standardization of clip placement.[9] Auditing of clip insertion has been suggested as one of the key performance indicators for quality control of breast cancer surgery.[19]

Based on their study which demonstrated reduction of IOV when adding MRI to CT, Jolicoeur *et al.* proposed that the use of CT-MRI fusion may obviate the need for surgical clips altogether.[25] However, while reduction of IOV indicates increased contouring agreement, it doesn't necessarily imply improved accuracy. To assess the accuracy, individual delineations would in theory need to be compared with the ground truth or correct delineation. In the absence of the histopathological proof, the ground truth is an elusive concept. Different approaches, including simultaneous truth and performance level estimation (STAPLE), expert consensus (EC) or their combination have been used as surrogates for correct delineation.[27,53]

In our current study, the concept of EC delineation was applied. Keeping in mind the limitations of the "ground truth" definition, our results indicate that adding MRI to CT improves contouring accuracy in cases without surgical clips in the LC cavity.

Comparison of our results with findings of other studies is challenging due to the variable conditions under which contouring was performed and the diversity of methods used for IOV assessment. The impact of variables such as experience and specialty of observers, use of guidelines, type of surgery, etc. should be kept in mind when comparing reports.[16] As far as the methods for IOV assessment are concerned, CI is one of the most commonly used quantifiers. In general, CI is a measure of overlap between analyzed volumes, but there is a diversity of formalisms used in the literature which cannot be directly compared. The generalized CI (CI_{gen}) formalism is independent of the number of delineations, enabling the comparisons between studies with different number of observers.[29] Regardless of the CI formalism used, the impact of contouring variation on CI is inversely proportional to the size of the analyzed volume. Therefore, same absolute deviation between analyzed contours will result in lower CI for small volumes (*i.e.* tumour bed) when compared with larger volumes (*i.e.* tumour bed with a margin). The effect of margins on CI is particularly relevant in breast cancer, where the contours are typically cropped to exclude the skin and chest wall, improving the apparent conformity between observers.

In our study, mean CI_{gen} of 0.67 (+/- 0.12) and 0.74 (+/- 0.07) was obtained for CT and MRI-based contouring of CTV_{LC}, respectively. Major *et al.* studied the impact of contouring guidelines on consistency of LC and planning target volume (PTV) contouring for multi-catheter partial breast irradiation. When contouring was performed on pre-implant scans by experienced observers and according to the guidelines (similar conditions as in our study), they obtained a CI_{gen} of 0.59 and 0.73 for LC and PTV, respectively. The margins for PTV were similar to our margins for CTV_{LC}, making the resulting volume sizes comparable between the two studies. Of note, CI_{gen} for PTV, obtained by CT and clip-based contouring[54] was similar to our CI_{gen} for CTV_{LC}, obtained by MRI in patients without clips. The lower CI_{gen} for LC when compared with PTV[54] reflects the sensitivity of CI to the volume size, as described above. Majority of other published studies reported on contouring uncertainties for tumour bed, with a CI ranging from 0.32–0.52.[21-23] Our results compare favourably with the existent literature. This can

be attributed to strict compliance with contouring guidelines, participation of experienced observers and use of high quality imaging.

Low number of observers and cases that were entered in analysis can be considered as the main limitations of our study. Considering the need for specific expertise in breast radiotherapy, experience in interpretation of MRI and relative rarity of cases without clips in LC, higher number of observers and cases is challenging to obtain outside a multi-institutional setting. This challenge is reflected in the limited number of observers and cases in studies, published by several authors before us.[20-23] Multi-centre collaborative projects may represent the optimal approach to overcome this limitation and shed more light on the subject of contouring uncertainties in general.

Conclusions

In breast cancer patients without clips in the tumour bed after breast conserving surgery, MRI improved the visualization of lumpectomy cavity when compared with CT. Consequently, inter-observer agreement and accuracy of contouring of lumpectomy cavity clinical target volume were improved. Placement of surgical clips, followed by CT-based contouring is the gold standard for contouring of the boost volume for postoperative irradiation in breast cancer. However, in patients without clips, addition of MRI to CT simulator images should be considered to improve delineation accuracy. Further studies with higher number of observers and cases are required to confirm our findings.

Declarations

The study protocol was reviewed and given ethical approval by the Institutional Medical Corporation Medical Research Centre. Datasets generated and analysed during study are not publicly available due to patient confidentiality but are available from corresponding author on reasonable request and after institutional approval. This study was not funded.

Acknowledgements

References

1. Fisher B, Anderson S, Bryant J, Margolese RG, Deutsch M, Fisher ER, et al. Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. *N Engl J Med* 2002; **347:** 1233-41. doi: 10.1056/NEJMoa022152

2. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials. *Lancet* 2011; **378:** 1707-16. doi: 10.1016/S0140-6736(11)61629-2

3. Bartelink H, Maingon P, Poortmans P, Weltens C, Fourquet A, Jager J, et al. Whole-breast irradiation with or without a boost for patients treated with breast-conserving surgery for early breast cancer: 20-year follow-up of a randomised phase 3 trial. *Lancet Oncol* 2015; **16:** 47-56. doi: 10.1016/S1470-2045(14)71156-8

4. Strnad V, Ott OJ, Hildebrandt G, Kauer-Dorner D, Knauerhase H, Major T, et al. 5-year results of accelerated partial breast irradiation using sole interstitial multicatheter brachytherapy versus whole-breast irradiation with boost after breast-conserving surgery for low-risk invasive and in-situ carcinoma of the female breast: a randomised, phase 3, non-inferiority trial. *Lancet* 2016; **387:** 229-38. doi: 10.1016/S0140-6736(15)00471-7

5. Vaidya JS, Wenz F, Bulsara M, Tobias JS, Joseph DJ, Keshtgar M, et al. Risk-adapted targeted intraoperative radiotherapy versus whole-breast radiotherapy for breast cancer: 5-year results for local control and overall survival from the TARGIT-A randomised trial. *Lancet* 2014; **383:** 603-13. doi: 10.1016/S0140-6736(13)61950-9

6. Offersen BV, Overgaard M, Kroman N, Overgaard J. Accelerated partial breast irradiation as part of breast conserving therapy of early breast carcinoma: a systematic review. *Radiother Oncol* 2009; **90:** 1-13. doi: 10.1016/j.radonc.2008.08.005

7. Segedin B, Petric P. Uncertainties in target volume delineation in radiotherapy – are they relevant and what can we do about them? *Radiol Oncol* 2016; **50:** 254-62. doi: 10.1515/raon-2016-0023

8. Strnad V, Hannoun-Levi JM, Guinot JL, Lössl K, Kauer-Dorner D, Resch A, et al. Recommendations from GEC ESTRO Breast Cancer Working Group (I): Target definition and target delineation for accelerated or boost Partial Breast Irradiation using multicatheter interstitial brachytherapy after breast conserving closed cavity surgery. *Radiother Oncol* 2015; **115:** 342-8. doi: 10.1016/j.radonc.2015.06.010

9. Major T, Gutiérrez C, Guix B, van Limbergen E, Strnad V, Polgar C., on behalf of Breast Cancer Working Group of GEC ESTRO. Recommendations from GEC ESTRO Breast Cancer Working Group (II): Target definition and target delineation for accelerated or boost partial breast irradiation using multicatheter interstitial brachytherapy after breast conserving open cavity surgery. *Radiother Oncol* 2016; **118:** 199-204. doi: 10.1016/j.radonc.2015.12.006

10. Boersma LJ, Janssen T, Elkhuizen PHM, Poortmans P, van der Sangen M, Scholten AN, et al. Reducing interobserver variation of boost-CTV delineation in breast conserving radiation therapy using a pre-operative CT and delineation guidelines. *Radiother Oncol* 2012; **103:** 178-82. doi: 10.1016/j.radonc.2011.12.021

11. Vinod SK, Min M, Holloway LC. A review of interventions to reduce interobserver variability in volume delineation in radiation oncology. *J Med Imaging Radiat Oncol* 2016; **60:** 393-406. doi: 10.1111/1754-9485.12462

12. Park CK, Pritz J, Zhang GG, Forster KM, Harris EE. Validating fiducial markers for image-guided radiation therapy for accelerated partial breast irradiation in early-stage breast cancer. *Int J Radiat Oncolo Biol Phys* 2012; 82: 425-31. doi: 10.1016/j.ijrobp.2011.07.027

13. Shaikh T, Chen T, Khan A, Yue NJ, Kearney T, Cohler A, et al. Improvement in interobserver accuracy in delineation of the lumpectomy cavity using fiducial markers. *Int J Radiat Oncol Biol Phys* 2010; **78:** 1127-34. doi: 10.1016/j.ijrobp.2009.09.025

14. Coles CE, Wilson CB, Cumming J, Benson JR, Forouhi P, Wilkinson JS, et al. Titanium clip placement to allow accurate tumour bed localisation following breast conserving surgery: audit on behalf of the IMPORT Trial Management Group. *Eur J Surg Oncol* 2009; **35:** 578-82. doi: 10.1016/j.ejso.2008.09.005

15. Weed DW, Yan D, Martinez AA, Vicini FA, Wilkinson TJ, Wong J. The validity of surgical clips as a radiographic surrogate for the lumpectomy cavity in image-guided accelerated partial breast irradiation. *Int J Radiat Oncol Biolo Phys* 2004; **60:** 484-92. doi: 10.1016/j.ijrobp.2004.03.012

16. Yang JT, Tao R, Elkhuizen PHM, van Vliet Vroegideweij C, Li G, Powell SN. Tumor bed delineation for external beam accelerated partial breast irradiation: A systematic review. *Radiother Oncol* 2013; **108:** 181-9. doi: 10.1016/j.radonc.2013.05.028

17. Yang Z, Chen J, Hu W, Pan Z, Cai G, Yu X, et al. Planning the breast boost: how accurately do surgical clips represent the CT seroma? *Radiother Oncol* 2010; **97:** 530-4. doi: 10.1016/j.radonc.2010.09.007

18. Goldberg H, Prosnitz RG, Olson JA, Marks LB. Definition of postlumpectomy tumor bed for radiotherapy boost field planning: CT versus surgical clips. *Int J Radiat Oncol Biol Phys* 2005; **63:** 209-13. doi: 10.1016/j.ijrobp.2005.01.044

19. Kirwan CC, Al Sarakbi W, Loncaster J, Chan HY, Thompson Am, Wishart GC. Tumour bed clip localisation for targeted breast radiotherapy: compliance is proportional to trial-related research activity: tumour bed clip localisation in breast radiotherapy. *Eur J Surg Oncol* 2014; **40:** 158-62. doi: 10.1016/j.ejso.2013.11.016

20. Kirby AM, Yarnold JR, Evans PM, Morgan VA, Schmidt MA, Scurr ED, et al. Tumor bed delineation for partial breast and breast boost radiotherapy planned in the prone position: what does MRI add to X-ray CT localization of titanium clips placed in the excision cavity wall? *Int J Radiat Oncol Biol Phys* 2009; **74:** 1276-82. doi: 10.1016/j.ijrobp.2009.02.028

21. Giezen M, Kouwenhoven E, Scholten AN, Coerkam EG, Heijenbrok M, Jansen WPA, et al. MRI- versus CT-based volume delineation of lumpectomy cavity in supine position in breast-conserving therapy: an exploratory study. *Int J Radiat Oncol Biol Phys* 2012; **82:** 1332-40. doi: 10.1016/j.ijrobp.2011.05.008

22. Den Hartogh MD, Phillipens MEP, van Dam IE, Kleynen CE, Tersteeg RJ, Kotte ANTJ, et al. Post-lumpectomy CT-guided tumor bed delineation for breast boost and partial breast irradiation: Can additional pre- and postoperative imaging reduce interobserver variability? *Oncol Lett* 2015; 10: 2795-801. doi: 10.3892/ol.2015.3697

23. Mast M, Coerkamp E, Heijenbrok M, Scholten A, Jansen W, Kouwenhoven E, et al. Target volume delineation in breast conserving radiotherapy: are co-registered CT and MR images of added value? *Radiat Oncol* 2014; **9:** 65. doi: 10.1186/1748-717X-9-65

24. Jacobson G, Zamba G, Betts V, Muruganandham M, Buechler-Price J. Image-based treatment planning of the post-lumpectomy breast utilizing CT and 3TMRI. *Int J Breast Cancer* 2011; **2011:** 1-5. doi: 10.4061/2011/246265

25. Jolicoeur M, Racine ML, Trop I, Hathout L, Nguyen D, Derashodian T, et al. Localization of the surgical bed using supine magnetic resonance and computed tomography scan fusion for planification of breast interstitial brachytherapy. *Radiother Oncol* 2011; **100:** 480-4. doi: 10.1016/j.radonc.2011.08.024

26. Petersen RP, Truong PT, Kader HA, Berthelet E, Lee JC, Hilts ML, et al. Target volume delineation for partial breast radiotherapy planning: clinical characteristics associated with low interobserver concordance. *Int J Radiat Oncol Biol Phys* 2007; **69:** 41-8. doi: 10.1016/j.ijrobp.2007.01.070

27. Petrič P, Hudej R, Rogelj P, Blas M, Tanderup K, Fidarova E, et al. Uncertainties of target volume delineation in MRI guided adaptive brachytherapy of cervix cancer: A multi-institutional study. *Radiother Oncol* 2013; **107:** 6-12. doi: 10.1016/j.radonc.2013.01.014

28. Rogelj P, Hudej R, Petrič P. Distance deviation measure of contouring variability. *Radiol Oncol* 2013; 47(1): 86-96. doi: 10.2478/raon-2013-0005.

29. Kouwenhoven E, Giezen M, Struikmans H. Measuring the similarity of target volume delineations independent of the number of observers. *Phys Med Biol* 2009; **54:** 2863-73. doi: 10.1088/0031-9155/54/9/018

30. Petric P, Hudej R, Rogelj P, Blas M, Segedin B, Logar HB, Dimopoulos JC. Comparison of 3D MRI with high sampling efficiency and 2D multiplanar MRI for contouring in cervix cancer brachytherapy. *Radiol Oncol* 2012; **46:** 242-51. doi: 10.2478/v10019-012-0023-1

31. Landis DM, Luo W, Song J, Bellon JR, Punglia RS, Wong JS, et al. Variability among breast radiation oncologists in delineation of the postsurgical lumpectomy cavity. *Int J Radiat Oncol Biol Phys* 2007; **67:** 1299-308. doi: 10.1016/j.ijrobp.2006.11.026

32. Van Mourik AM, Elkhuizen PH, Minkema D, Duppen JC; Dutch Young Boost Study Group, van-Vliet-Vroegindeweij C. Multiinstitutional study on target volume delineation variation in breast radiotherapy in the presence of guidelines. *Radiother Oncol* 2010; **94:** 286-91. doi: 10.1016/j.radonc.2010.01.009

33. Struikmans H, Warlam-Rodenhuis C, Stam T., Stapper G, Tersteeg RJ, Bol GH, Raaijmakers CP. Interobserver variability of clinical target volume delineation of glandular breast tissue and of boost volume in tangential breast irradiation. *Radiother Oncol* 2005; **76:** 293-9. doi: 10.1016/j.radonc.2005.03.029

34. Hurkmans CW, Admiraal M, van der Sangen M, Dijkmans I. Significance of breast boost volume changes during radiotherapy in relation to current clinical interobserver variations. *Radiother Oncol* 2009; **90:** 60-5. doi: 10.1016/j.radonc.2007.12.001

35. Rasch CRN, Steenbakers RJHM, Fitton I, Duppen JC, Nowak PJ, Pameijer FA, et al. Decreased 3D observer variation with matched CT-MRI for target delineation in Nasopharynx cancer. *Radiat Oncol* 2010; **5:** 21. doi: 10.1186/1748-717X-5-21

36. Viswanathan AN, Dimopoulos JCA, Kirisits C, Berger D, Pötter R. CT versus MRI-based contouring in cervical cancer brachytherapy: results of a prospective trial and preliminary guidelines for standardized Contours. *Int J Radiat Oncol Biol Phys* 2007; **68:** 491-8. doi: 10.1016/j.ijrobp.2006.12.021

37. De Brabandere M, Hoskin P, Haustermans K, Van den Heuvel F, Siebert FA. Prostate post-implant dosimetry: interobserver variability in seed localisation, contouring and fusion. *Radiother Oncol* 2012; **104:** 192-8. doi: 10.1016/j.radonc.2012.06.014

38. Villeirs GM, Van Vaerenbergh K, Vakaet L, Bral S, Claus F, De Neve WJ, et al. Interobserver delineation variation using CT versus combined CT + MRI in intensity-modulated radiotherapy for prostate cancer. *Strahlenther Onkol* 2005; **181:** 424-30. doi: 10.1007/s00066-005-1383-x

39. Cattaneo GM, Reni M, Rizzo G, Castellone P, Ceresoli GL, Cozzarini C, et al. Target delineation in post-operative radiotherapy of brain gliomas: interobserver variability and impact of image registration of MR(pre-operative) images on treatment planning CT scans. *Radiother Oncol* 2005; **75:** 217-23. doi: 10.1016/j.radonc.2005.03.012

40. Weiss E, Hess CF. Comment by E. Weiss and C. F. Hess on G. M. Villeirs et al. "Interobserver delineation variation using CT versus combined CT + MRI in intensity-modulated radiotherapy for prostate cancer" in: Strahlenther Onkol 2005; 181:424-30 (No. 7). *Strahlenther Onkol* 2005; **181:** 743-4. doi: 10.1007/s00066-005-9383-4

41. Hegazy N, Pötter R, Kirisits C, Berger D, Federico M, Sturdza A, Nesvacil N. High-risk clinical target volume delineation in CT-guided cervical cancer brachytherapy: impact of information from FIGO stage with or without systematic inclusion of 3D documentation of clinical gynecological examination. *Acta Oncol* 2013; **52:** 1345-52. doi: 10.3109/0284186X.2013.813068

42. Haie-Meder C, Pötter R, Van Limbergen E, Briot E, De Brabandere M, Dimopoulos J, et al. Recommendations from Gynaecological (GYN) GEC-ESTRO Working Group (I): concepts and terms in 3D image based 3D treatment planning in cervix cancer brachytherapy with emphasis on MRI assessment of GTV and CTV. *Radiother Oncol* 2005; **74:** 235-45. doi: 10.1016/j.radonc.2004.12.015

43. Barkati M, Simard D, Taussky D, Delouya G. Magnetic resonance imaging for prostate bed radiotherapy planning: An inter- and intra-observer variability study. *J Med Imaging Radiat Oncol* 2016; **60:** 255-9. doi: 10.1111/1754-9485.12416

44. Weltens C, Menten J, Feron M, Bellon E, Demaerel P, et al. Interobserver variations in gross tumor volume delineation of brain tumors on computed tomography and impact of magnetic resonance imaging. *Radiother Oncol* 2001; **60:** 49-59.

45. Buijsen J, van den Bogaard J, van der Heide H, Engelsman S, van Stiphout R, Janssen M, et al. FDG-PET-CT reduces the interobserver variability in rectal tumor delineation. *Radiother Oncol* 2012; 12: 371-6. doi: 10.1016/j.radonc.2011.12.016

46. Daisne JF, Duprez T, Weynand B, Lonneux M, Hamoir M, Reychler H, Grégoire V. Tumor volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. *Radiology* 2004; **233:** 93-100. doi: 10.1148/radiol.2331030660

47. Khoo V, Adams EJ, Saran F, Bedford JL, Perks JR, et al. A comparison of clinical target volumes determined by CT and MRI for the radiotherapy planning of base of skull meningiomas. *Int J Radit Oncol Biol Phys* 2000; **46:** 1309-17.

48. Kass R, Kumar G, Klimberg VS, Kass L, Henry-Tillman R, Johnson A, et al. Clip migration in stereotactic biopsy. *Am J Surg* 2002; **184:** 325-31.

49. Benda RK, Yasuda G, Sethi A, Gabram SG, Hinerman RW, Mendenhall NP. Breast boost: are we missing a target? *Cancer* 2003; **97:** 905-9. doi: 10.1002/cncr.11142

50. Dzugashvili M, Pichenot C, Dunant A, Balleyquier C, Delaloge S, Mathieu MC, et al. Surgical clips assist in the visualization of the lumpectomy cavity in three-dimensional conformal accelerated partial-breast irradiation. *Int J Radiat Oncol Biol Phys* 2010; **76:** 1320-4. doi: 10.1016/j.ijrobp.2009.04.089

51. Kovner F, Agay R, Merimsky O, Stadler J, Klausner J, Inbar M. Clips and scar as the guidelines for breast radiation boost after lumpectomy. *Eur J Surg Oncol* 1999; **25:** 483-6. doi: 10.1053/ejso.1999.0683

52. Krawczyk JJ, Engel B. The importance of surgical clips for adequate tangential beam planning in breast conserving surgery and irradiation. *Int J Radiat Oncol Biol Phys* 1999; **43:** 347-50.

53. Warfield SK, Zou KH, Wells WM. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *IEEE Trans Med Imaging* 2004; **23:** 903-21. doi: 10.1109/TMI.2004.828354

54. Major T, Gutierrez C, Guix B, Emoke Mozsa, Hannoun-Levi JM, Lössl K, et al. Interobserver variations of target volume delineation in multicatheter partial breast radiotherapy after open cavity surgery. *Brachytherapy* 2015; **14:** 925-32. doi: 10.1016/j.brachy.2015.06.008