# Atlas-based segmentation technique incorporating inter-observer delineation uncertainty for whole breast

**L R Bell[1,2], J A Dowling[3], E M Pogson[1,2,5], P Metcalfe[1,2], L Holloway[1,2,4,5]**
[1]Centre for Medical Radiation Physics, University of Wollongong, Australia
[2]Liverpool & Macarthur Cancer Therapy Centres & Ingham Institute, Liverpool, Australia
[3]Australian e-Health Research Centre, CSIRO, Queensland, Australia
[4]SWSCS, University of New South Wales, Australia
[5]Institute of Medical Physics, University of Sydney, Australia

E-mail: lb998@uowmail.edu.au

**Abstract.** Accurate, efficient auto-segmentation methods are essential for the clinical efficacy of adaptive radiotherapy delivered with highly conformal techniques. Current atlas based auto-segmentation techniques are adequate in this respect, however fail to account for inter-observer variation. An atlas-based segmentation method that incorporates inter-observer variation is proposed. This method is validated for a whole breast radiotherapy cohort containing 28 CT datasets with CTVs delineated by eight observers. To optimise atlas accuracy, the cohort was divided into categories by mean body mass index and laterality, with atlas' generated for each in a leave-one-out approach. Observer CTVs were merged and thresholded to generate an auto-segmentation model representing both inter-observer and inter-patient differences. For each category, the atlas was registered to the left-out dataset to enable propagation of the auto-segmentation from atlas space. Auto-segmentation time was recorded. The segmentation was compared to the gold-standard contour using the dice similarity coefficient (DSC) and mean absolute surface distance (MASD). Comparison with the smallest and largest CTV was also made. This atlas-based auto-segmentation method incorporating inter-observer variation was shown to be efficient (<4min) and accurate for whole breast radiotherapy, with good agreement (DSC>0.7, MASD <9.3mm) between the auto-segmented contours and CTV volumes.

## 1. Introduction

Internal anatomical changes during the course of radiotherapy limit treatment accuracy, as current treatments are planned on static images obtained prior to delivery [1]. Adaptive radiotherapy is an increasingly investigated approach that aims to combat these anatomical changes by re-imaging and re-planning at multiple time points throughout the treatment course. However, this process has large workflow implications due to the time required to achieve this whilst maintaining plan quality.

Target delineation is the bottleneck of the planning process taking on average 18.6 min per patient for breast radiotherapy [2], and housing one of the largest uncertainties affecting modern radiotherapy accuracy [3]. Segmentation techniques, boasting significant decreases in delineation times as compared to manual contouring, have been proposed as the solution to this limitation with several techniques implemented with impressive results [4]. Atlas-based segmentation techniques in particular are suited for adaptive radiotherapy purposes as they are derived from real patient anatomy and verified clinician contours. However, current techniques generally fail to incorporate inter-observer variation, with single

clinician contours frequently taken to be the gold standard delineation. With the largest uncertainty in modern radiotherapy treatment occurring during the target delineation process, it is essential to account for delineation uncertainty when devising an adaptive radiotherapy process.

This atlas based segmentation technique improves on the accuracy of existing techniques by incorporating inter-observer variation into the segmentation process, whilst maintaining efficacy.

## 2. Method

This segmentation technique, although developed and validated for a whole breast, is applicable to any target volume definition in which the contours are defined based on anatomy. The general process involves generating atlas, registering the atlas to a new patient to obtain good correspondence between equivalent regions, and mapping the atlas segmentation to the new image space.

### *2.1. Technique development*

The atlas generation method, outlined in figure 1, has been utilised previously in medical imaging applications [5]. An initial dataset representative of the patient cohort was chosen with remaining datasets in the cohort registered to the initial dataset using rigid and affine transformation [6]. The results were averaged and used as the subject for the next iteration in which rigid, affine registration and diffeomorphic demons non-rigid registration [7] was used. Three iterations were performed to produce the final atlas.

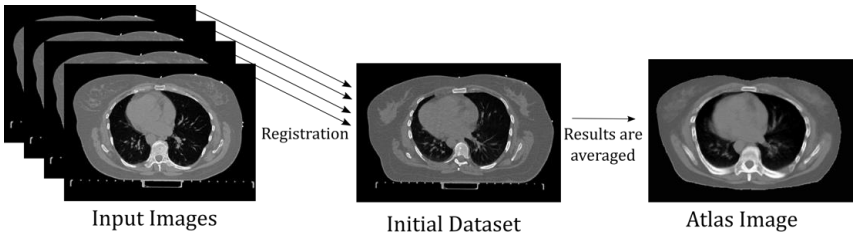


**Figure 1.** Method for atlas generation.

All CTVs except those of the left-out dataset were mapped to atlas space using the existing deformation fields, and merged. The resulting contouring probability model was thresholded to 50% to obtain an auto-segmented CTV contour.

### *2.2. Validation*

The technique was validated on a whole breast radiotherapy cohort containing datasets from 28 patients (14 right, 14 left), collected for a previous study [8]. Non-contrast computed tomography (CT) scans were used with whole breast clinical target volumes (CTV) delineated by eight independent observers. Observers followed a delineation protocol and were blind to other observer contours.

Since atlas accuracy correlates with similar anatomy of input patients, the cohort was divided into categories, grouping similar patients. Body mass index (BMI) was used as a surrogate for similar patient anatomy, with small/large patients classed as </> 29 BMI. Datasets were further divided according to laterality. An atlas was generated from the 6-9 datasets in each category using a leave-one-out approach. The atlas was registered using rigid and affine transformation followed by non-rigid registration, to the 'left out' dataset. The segmentation was mapped to this dataset from atlas space, and clipped such that it did not extend past the patient surface.

A gold standard, consensus contour was generated from the observer CTVs on the left-out dataset in each category using the Simultaneous and Truth Performance Level Estimation (STAPLE) algorithm [9]. To assess similarity, the gold standard contour was compared to the segmentation using the dice similarity coefficient (DSC) [10] and mean absolute surface distance (MASD) [11]. A comparison between the smallest and largest CTV with the segmentation was also made to determine the range in accuracy.

## 3. Results

Average atlases with corresponding whole breast CTV segmentations were successfully generated for each category as shown in figure 2 with corresponding probability maps.

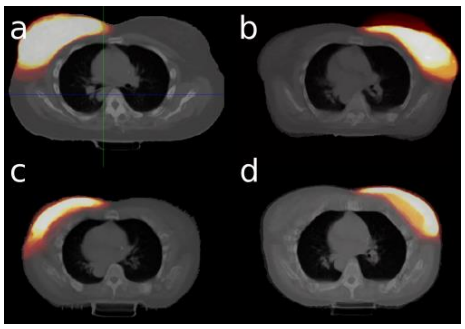


**Figure 2.** Probability maps of observer contours for a) large right, b) large left, c) small left, d) small right atlas'. Heat map indicates regions of highest probability.

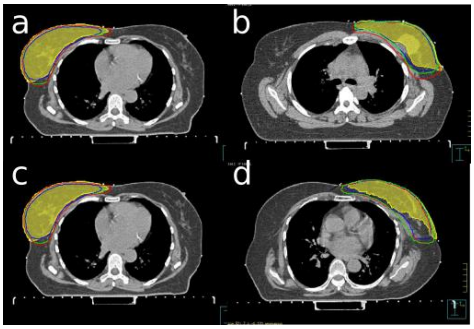


**Figure 3.** Segmentation (yellow) for a) large-right, b) large-left, c) small-left, d) small-right datasets. STAPLE (green), smallest (blue) and largest (red) CTVs are shown for comparison.

The whole breast CTV segmentation is visually compared to the consensus, smallest and largest contour in figure 3. The time required to segment, including registration of the atlas to the required patient and propagating contours, was on average 3 min and 43 seconds and didn't exceed 4 min and 49 seconds. The DSC and MASD values indicating the accuracy of the auto-segmentation in delineating the whole breast CTV and accounting for inter-observer variation is shown in table 2.

**Table 2.** Similarity metrics comparing whole breast segmentation with manual target volumes

| | DSC | | | | MASD (mm) | | | |
|---|---|---|---|---|---|---|---|---|
| | Large Left | Large Right | Small Left | Small Right | Large Left | Large Right | Small Right | Small Left |
| Consensus | 0.81 | 0.85 | 0.71 | 0.79 | 7.99 | 3.47 | 6.17 | 5.16 |
| Smallest CTV | 0.77 | 0.86 | 0.71 | 0.79 | 9.28 | 3.64 | 4.81 | 5.16 |
| Largest CTV | 0.81 | 0.8 | 0.71 | 0.79 | 7.50 | 4.71 | 6.13 | 5.07 |

## 4. Discussion

Coverage probability maps overlaid on average atlas' showed high observer agreement for the majority of the whole breast segmentation. Observer agreement was lowest in the anterior-medial and posterior-lateral directions as is supported in the literature [12].

Segmented contours generally compared well with the gold standard for each category. Insufficiencies in segmentation were observed close to the chest wall for large and small BMI left breast patients. The authors hypothesise that this is due to larger contouring variations due to the close proximity of the heart. Manual review of this key area by a clinician is recommended. Improvements in segmentation in these instances could be achieved through optimisation of the probability map thresholding value.

## 5. Conclusion

This atlas based segmentation technique incorporating interobserver variation has been shown to be accurate (DSC >0.7, MASD <1cm) and efficient (time < 5 min) as is necessary for the clinical efficacy

of adaptive radiotherapy. This successfully demonstrates an approach for presenting likely inter-observer variation determined from previous datasets on a current dataset. There is scope for further optimisation of segmentation correlation with gold standard contours as well as computation time. This method is a feasible solution to contouring inefficiencies for adaptive radiotherapy, with the potential for viable application to other anatomical sites.

## 6. Acknowledgments

The authors would like to acknowledge V.Batumalai and G.Delaney for the provision of data for this study. This research is funded by a Cancer Australia and National Breast Cancer Foundation (1033237), Liverpool Hospital Radiation Oncology trust fund, Australian Postgraduate Award and South Western Sydney Radiation Oncology Student Scholarship.

## 7. References

[1] Vial P *et al* 2008 *Med. Phys.* **35** 1267-77
[2] Reed V K *et al* 2009 *Int. J. Radiat. Oncol. Biol. Phys.* **73** 1493-500
[3] Njeh C F 2008 *J. Med. Phys.* **33** 136-40
[4] Dowling J A *et at* 2011 *Proc. Int. Workshop held in conjunction with MICCAI (Toronto, Canada, 22 September, 2011)* (Heidelberg: Springer) pp 10-21
[5] Dowling J A *et al* 2012 *Int. J. Radiat. Oncol. Biol. Phys.* **83** e5-11
[6] Ourselin S *et al* 2001 *Image Vision Comput.* **19** 25-31
[7] Vercauteren T *et al* 2009 *NeuroImage* **45** S61-72
[8] Batumalai V *et al* 2011 *Clin. Oncol.* **23** 108-13
[9] Warfield S K *et al* 2004 *IEEE T Med. Imaging* **23** 903-21
[10] Abreu *et al* 2006 *Proc. 12th Pacific Rim Int. Symp. on Dependable Computing (Riverside, California, 18-20 December, 2006*) (IEEE) pp 39-46
[11] Babalola K O *et al* 2009 *NeuroiImage* **47** 1435-47
[12] Pogson E M *et al* 2014 *J. Phys. Conf. Ser.* **489** 012057