Investigation of cardiac substructure automatic segmentation methods on synthetically generated 4D cone-beam CT images

Mark Gardner[1], Robert N Finnegan[2,3,4], Owen Dillon[1], Vicky Chin[1,4,5], Tess Reynolds[1], Paul J Keall[1]

1- Image X Institute, University of Sydney, Sydney, Australia

2- Institute of Medical Physics, School of Physics, University of Sydney, Sydney, NSW, Australia

3- Northern Sydney Cancer Centre, Royal North Shore Hospital, Sydney, New South Wales, Australia

4- Ingham Institute for Applied Medical Research, Sydney, Australia

5- Liverpool and Macarthur Cancer Therapy Centres, Department of Radiation Oncology, Sydney, Australia

Correspondence: Mark Gardner, Image X Institute, The University of Sydney, Eveleigh, New South Wales 2015, Australia. Email: mark.gardner@sydney.edu.au

Abstract

Background: STereotactic Arrhythmia Radioablation (STAR) is a novel non-invasive method for treating arrythmias in which external beam radiation is directed towards subregions of the heart. Challenges for accurate STAR targeting include small target volumes and relatively large patient motion, which can lead to radiation related patient toxicities. 4D Cone-beam CT (CBCT) images are used for stereotactic lung treatments to account for respiration-related patient motion. 4D-CBCT imaging could similarly be used to account for respiration-related patient motion in STAR, however the poor contrast of heart tissue in CBCT makes identifying cardiac substructures in 4D-CBCT images challenging. If cardiac structures can be identified in pre-treatment 4D-CBCT images, then the location of the target volume can be more accurately identified for different phases of the respiration cycle, leading to more accurate targeting and a reduction in patient toxicities.

Purpose: The aim of this simulation study is to investigate the accuracy of different cardiac substructure segmentation methods for 4D-CBCT images.

Methods: Repeat 4D-CT scans from 13 lung cancer patients were obtained from The Cancer Imaging Archive. Synthetic 4D-CBCT images for each patient were simulated by forward projecting and reconstructing each respiration phase of a chosen “testing” 4D-CT scan. 18 cardiac structures were segmented from each respiration phase image in the testing 4D-CT using the previously validated *platipy* toolkit. The *platipy* segmentations from the testing 4D-CT were defined as the ground truth segmentations for the synthetic 4D-CBCT images. Five different 4D-CBCT cardiac segmentation methods were investigated: 3D Rigid Alignment, 4D Rigid Alignment, Direct CBCT Segmentation, Contour Transformation and Synthetic CT Segmentation methods. For all methods except the Direct CBCT segmentation method, a separate 4D-CT (Planning CT) was used to assist in generating 4D-

CBCT segmentations. Segmentation performance was measured using the Dice Similarity Coefficient (DSC), Hausdorff Distance (HD), Mean Surface Distance (MSD), and Volume Ratio (VR) metrics.

Results: The mean±standard deviation DSC for all cardiac substructures for the 3D Rigid Alignment, 4D Rigid Alignment, Direct CBCT Segmentation, Contour Transformation and Synthetic CT Segmentation methods were 0.48±0.29, 0.52±0.29, 0.37±0.32, 0.53±0.29, 0.57±0.28 respectively. Similarly, the HD values were 10.9±3.6 mm, 9.9±2.6 mm, 17.3±5.3 mm, 9.9±2.8 mm, 9.3±3.0 mm, the MSD values were 2.9±0.6 mm, 2.9±0.6 mm, 6.3±2.5 mm, 2.5±0.6 mm, 2.4±0.8 mm and the VR Values were 0.81±0.12, 0.78±0.14, 1.10±0.47, 0.72±0.15, 0.98±0.44 respectively. Of the five methods investigated the Synthetic CT segmentation method generated the most accurate segmentations for all calculated segmentation metrics.

Conclusions: This simulation study investigates the accuracy of different cardiac substructure segmentation methods for 4D-CBCT images. Accurate 4D-CBCT cardiac segmentation will provide more accurate information on the location of cardiac anatomy during STAR treatments which can lead to safer and more effective STAR. As the data and segmentation methods used in this paper are all open source, this paper provides a useful benchmarking tool to evaluate other CBCT cardiac segmentation methods.

Introduction

STereotactic Arrhythmia Radioablation (STAR) is a promising new treatment for arrhythmias including Ventricular Tachycardia (VT), in which external beam radiation therapy is used to deliver radiation to the VT site. The advantage of STAR over the traditional catheter-based ablations is that STAR is non-invasive, faster and a more cost effective treatment[1]. Initial clinical results have shown promising evidence that STAR can reduce the number of VT episodes for VT patients who have not responded to pharmaceutical or catheter ablation treatments[2–4].

One of the major concerns for STAR is the possibility of side effects of radiation toxicity. Robinson *et al*. found that 2 out of the 21 patients treated with STAR had grade 3 cardiac side effects that were found to have been potentially due to radiation toxicity[2]. Similarly, Qian et al. reported that 3 out of 6 patients treated using STAR had adverse events possibly due to the treatment[4]. Accurate targeting of treatment volumes in the heart is made difficult by the combination of small target volumes and the large intrafraction motion, which is a combination of cardiac and respiration motion with respiration motion making up a significant component of this motion[5–8].

Accurately tracking the target volume during STAR treatments could lead to a reduction in planning target volume (PTV) margins, leading to less healthy tissue being irradiated, which could reduce the chance of radiation toxicity occurrence. 4D cone-beam CT (4D-CBCT) imaging, in which CBCT images are reconstructed for different phases of respiration, has been used for stereotactic lung tumor treatments which are similarly affected by respiratory motion. Indeed ESTRO ACROP guidelines recommend daily 4D-CBCT imaging for lung SBRT to compensate for patient respiration motion[9]. One could infer that these recommendations should also apply to STAR given the central location affected by respiratory motion and surrounded by critical structures. If the cardiac substructure locations can be accurately identified in 4D-CBCT images acquired immediately prior to patient treatment, then the location of the cardiac substructures for each respiration phase can potentially be used to increase the targeting accuracy for STAR.

In this simulation study we investigate methods for segmenting the cardiac substructures in synthetic CBCT images. To the authors' knowledge this is the first paper comparing methods for segmenting cardiac substructures in CBCT images, albeit restricted to synthetic 4DCBCT images.

Method

The methods used to develop the cardiac substructure CBCT segmentation methods are shown in Figure 1. The data used for this study came from a dataset that acquired multiple repeat 4D-CT scans for each patient. For each patient two 4D-CT scans were used, with one 4D-CT scan designated as the "planning 4D-CT", and the other 4D-CT scan the "testing 4D-CT". The planning 4D-CT was used by several of the CBCT cardiac segmentation methods to assist in generating the cardiac substructure segmentations. The testing CT scans were used to generate the ground truth cardiac substructure segmentations as well as the simulated 4D-CBCT images. The use of the simulated 4D-CBCT scans generated from 4D-CT scans meant that the 4D-CT generated ground truth segmentations were aligned with the 4D-CBCT images. The five CBCT cardiac segmentation methods that were evaluated were 3D Rigid Alignment, 4D Rigid Alignment, Direct CBCT segmentation, Contour Transformation and the Synthetic CT Segmentation methods. The 3D and 4D rigid alignment methods represent the methods currently used clinically for STAR treatments[10], and the Direct CBCT segmentation, Contour Transformation and Synthetic CT Segmentation methods are novel CBCT segmentation methods.

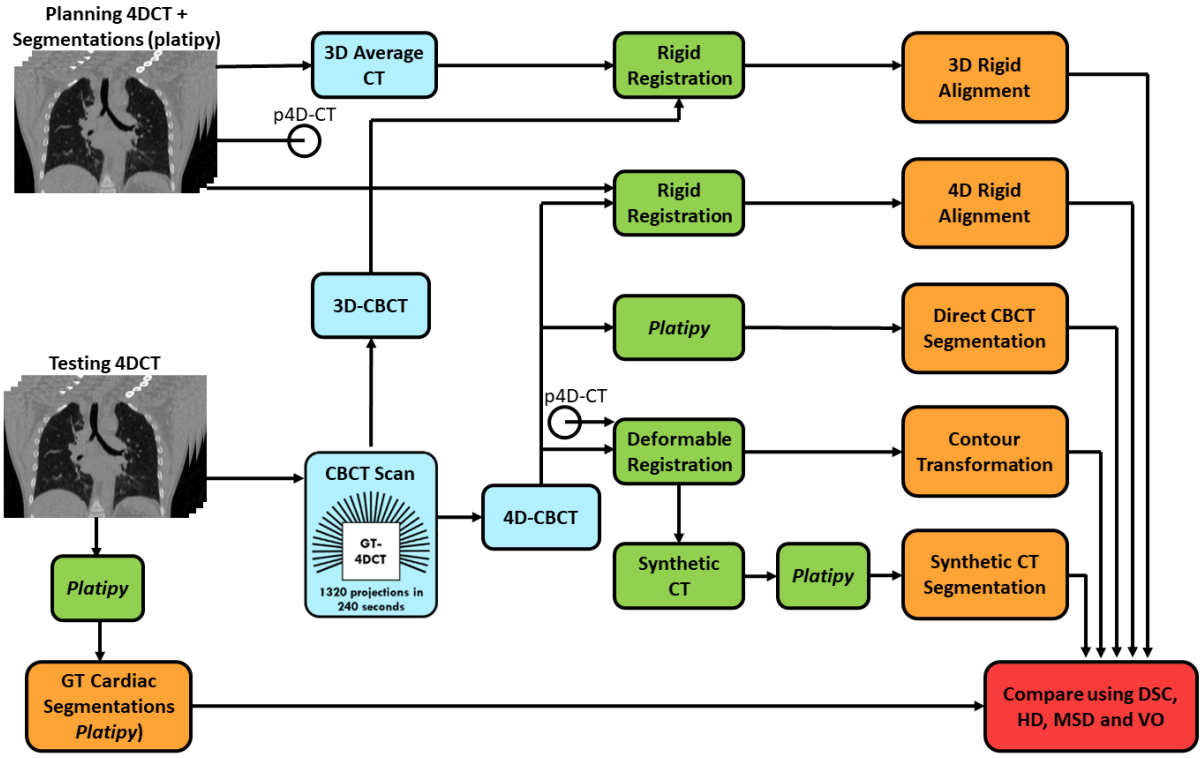


Figure 1 - A flowchart to show how the data were derived and how these data were used to generate the different segmentations for each patient. Note that CBCT images were simulated from real patient 4DCT images treated as the ground truth. The five CBCT cardiac segmentation methods that were evaluated were 3D Rigid Alignment, 4D Rigid Alignment, Direct CBCT segmentation, Contour Transformation and the Synthetic CT Segmentation methods. p4D-CT – planning 4D-CT, CBCT – cone-beam CT.

Patient data

The data were derived from the publicly available 4D-Lung dataset in The Cancer Imaging Archive (TCIA)[11–13] From the 4D-Lung dataset, repeat 4D-CT scans were acquired from 13 lung cancer patients, which were designated as the planning and testing 4D-CTs. The duration between the acquisition of the planning and testing 4D-CTs was 6-30 days (median 14 days). As previous analysis has shown considerable variability in lung anatomy for CBCT images acquired from different treatment fractions[14], we anticipate similar variation in anatomy between the repeat 4D-CT scans. The testing 4D-CT was used to create synthetic 4D-CBCT images for each patient, as described in a

previous publication[15,16]. Briefly, to create the simulated 4D-CBCT images, each image of each respiration phase in the testing 4D-CT were forward projected to create digitally reconstructed radiographs (DRRs) based on conventional CBCT geometry. 4D-CBCT images were created by reconstructing the DRRs for the different respiration phases. This method for creating synthetic 4D-CBCT images provides for each patient a set of 4D-CT and 4D-CBCT images with identical anatomy. To simulate a conventional 4D-CBCT scan, 132 simulated projections were created for each 4D-CT volume (1320 projections in total) patient around a 200° arc, which simulated a conventional 4D-CBCT acquisition time of 4 minutes. The voxel size for the CT was 0.98 x 0.98 x 3 mm^3 and the voxel size for each of the reconstructed CBCT volumes was 1 × 1 × 1 mm^3. The voxel values of the reconstructed 4D-CBCT volumes were converted from attenuation coefficients to HU values using the formula:

$$Vol_{HU} = \frac{1000 \times Vol_{Att}}{\alpha} - 1000$$

where α is an estimation of the attenuation value for water, which was set to 0.013 mm^{-1} such that the HU values for water, lung and bone were approximately 0, -500 and 1000 HU[17].

Segmentation Method

The cardiac substructure segmentations of the testing, planning and synthetic 4D-CT images, as well as the CBCT images in the direct CBCT segmentation method, were generated using the *platipy* cardiac segmentation tool (https://pyplati.github.io/platipy/)[18]. This method was selected as it is a fully automated and open-source tool that has previously been validated against manually created segmentations by clinicians on CT scans using a dataset from multiple institutions[19,20]. Briefly, this method uses a trained nnUnet method to obtain the segmentation of the whole heart[21]. Then, multiple atlas images were deformably registered to the CT image being segmented to get the cardiac chamber and great vessel segmentations with the whole heart segmentation used to assist in constraining the registration[22]. Finally, the segmentations for the coronary arteries, cardiac valves and conduction nodes are created geometrically based on published contouring guidelines[23–25].

For all segmentation methods, the following cardiac substructure segmentations were generated: whole heart, left and right ventricles, left and right atria, ascending aorta, pulmonary artery, superior vena cava, aortic, pulmonary, mitral and tricuspid valves, left anterior descending coronary artery, left circumflex artery, left and right coronary arteries, atrioventricular conduction node, sinoatrial conduction node. The ground truth CBCT segmentations were obtained by applying the *platipy* cardiac substructure segmentation method to each image of the patient testing 4D-CT volumes.

An automatic cardiac structure segmentation method was used to generate the ground truth segmentation methods instead of using manual segmentation contours to provide consistent and repeatable segmentations. While manual contouring is used clinically, inter-observer variability can lead to significant variation in the created contours[26], particularly for smaller substructures like the coronary arteries, cardiac valves and conduction nodes[20,27,28]. Auto-contouring methods such as *platipy* have been recommended as a method to reduce inter-observer variability[29].

All registrations were completed using the *elastix* toolbox [30]. Processing was completed on a linux desktop computer with an Intel Xeon Gold 6226R processor (2.9 GHz) with 128 GB of RAM and an NVIDIA RTX A6000 graphics processing unit (GPU). Code for implementing the Direct CBCT Segmentation, Contour Transformation and Synthetic CT Segmentation methods are available via a public GitHub repository (https://github.com/Image-X-Institute/CBCTCardiacSegmentation).

3D Rigid Alignment

The 3D rigid alignment method is the most common clinical method for positioning the patient prior to STAR treatments[10]. First, the 3D average-planning CT was rigidly registered to the 3D-CBCT, to align the general anatomy. Following this registration the whole heart segmentation from the 3D average-CT was extended by 20 mm to create an extended heart mask. Then, a 3 degree-of-freedom (DOF) translation registration between the 3D-CBCT and the shifted 3D-CT was performed using the extended heart mask on both the 3D-CBCT and shifted 3D-CT volumes to align the heart of the 3D-CT to the heart of the 3D-CBCT.

4D Rigid Alignment

The 4D rigid alignment method is similar to the 3D rigid alignment method, except it involves a 4D-CT to 4D-CBCT registration instead of a 3D-CT to 3D-CBCT registration. That is, every respiration phase image of the planning 4D-CT is rigidly registered to the corresponding 4D-CBCT respiration phase image. This method has been used clinically to treat a small number of patients[31].

Direct CBCT Segmentation

For the direct CBCT segmentation method the *platipy* segmentation method was applied to each CBCT volume.

Contour Transformation

The Contour Transformation method involved using a B-spline deformable image registration method to register the planning 4D-CT images to the 4D-CBCT images. An extended heart mask was created by padding the whole heart segmentation from the planning CT by 20 mm and this extended heart mask was used to mask the input images during the deformable registration project to optimize accurate registration around the heart. The output of the deformable registration was a transformation from every respiration phase image of the planning 4D-CT to every corresponding 4D-CBCT respiration phase image. This CT to CBCT transformation was used to transform the segmentations from every respiration phase of the planning 4D-CT to align with every corresponding 4D-CBCT respiration phase image.

Synthetic CT Segmentation

Similarly to the Contour Transformation method, in the Synthetic CT Segmentation method a B-spline deformable image registration method was used to get the transformation from the CT to the CBCT images. However, for the Synthetic CT Segmentation method, this transformation was applied to every respiration phase of the planning 4D-CT to generate a synthetic 4D-CT from the 4D-CBCT images in the form of deformed planning CT images. The *platipy* cardiac substructure segmentation method was then applied to each respiration phase of the generated synthetic 4D-CT. The advantage of this method over the Contour Transformation method is that deformable image registration may not always accurately preserve the shape of small substructure segmentations.

Evaluation Metrics

Quantitative metrics were used to compare the different CBCT segmentation methods to the ground truth cardiac structure segmentations. As the cardiac structures are a range of shapes and sizes a variety of different evaluation metrics were used. These metrics include Dice similarity coefficient (DSC), the mean surface distance (MSD), the (maximum) Hausdorff distance (HD), and volume ratio (VR) (computed as estimated/ground truth). DSC was used as a it is a common metric for comparing segmentations, and thus will be useful for comparing the results to other similar datasets. However, DSC is a poor metric for comparing the segmentation accuracy of different sized segmentations, and so MSD and HD were also calculated as they provide a more comparable measurement of overall

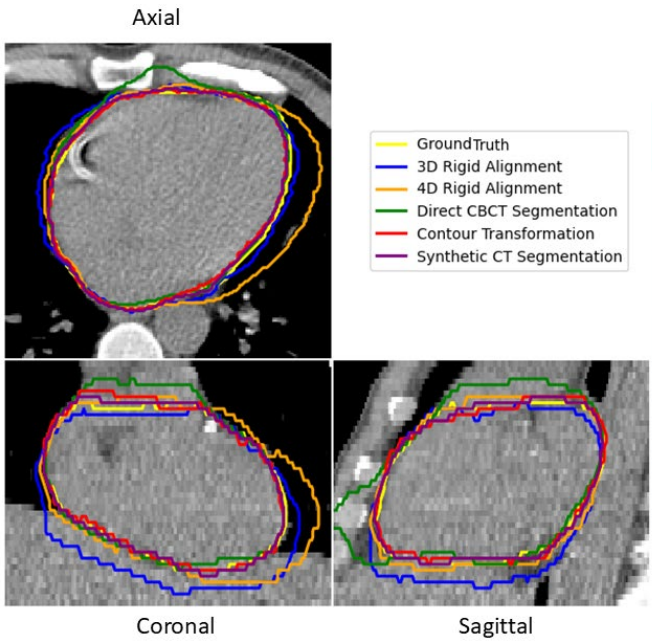
surface deviation (MSD) and maximum segmentation deviation (HD). VR was also used as under or over estimating the volume of segmentations can affect dose volume histogram calculations, as well as metrics such as mean and maximum dose[32]. The term "accuracy" is used to compare the segmentations from the different CBCT segmentation methods to the ground truth segmentations generated by *platipy*. The segmentations generated by the *platipy* auto-segmentation method have associated uncertainties and are not necessarily the "true" position of each cardiac substructure [19,20,33], however as mentioned previously the advantage of auto-contouring methods over manual contours is to provide consistency and overcome issues related to inter-observer variability.

The different segmentation methods were compared with each other for all respiration phases, cardiac structures and patients. Statistical analysis was conducted using a one-way repeated measurements ANOVA with significance $\alpha=0.05$ with Dunnett corrections applied for multiple comparisons. Statistical analysis was conducted using Graphpad Prism (Graphpad Software, USA).

Results

Qualitative results

An example of the whole heart segmentations generated by the different cardiac CBCT segmentation methods is shown in Figure 2. This example had the largest MSD value for the 3D Rigid Alignment whole heart segmentation. This figure shows how the Contour Transformation and Synthetic CT Segmentation methods can better compensate for the changing anatomy than the 3D and 4D Rigid Alignment methods. Additionally, the Direct CBCT Segmentation method produces a segmentation that mostly aligns with the ground truth heart segmentation, except for the most anterior part of the heart where the accuracy of the Direct CBCT Segmentation is poor. An example of every cardiac substructure segmentation for all CBCT segmentation methods for this patient is shown in the supplementary materials.



	DSC	HD (mm)	MSD (mm)	VR
3D Rigid Alignment	0.85	18.8	5.2	1.25
4D Rigid Alignment	0.87	20.4	3.8	1.25
Direct CBCT Segmentation	0.90	29.4	3.1	1.14
Contour Transformation	0.95	12.0	1.2	1.00
Synthetic CT Segmentation	0.96	6.0	0.9	1.01

Figure 2 – (Left) An example of the whole heart segmentations for the different cardiac CBCT segmentations methods overlayed on a CBCT image for patient 1 for the exhale respiration phase. (Right) Quantitative metrics for the heart segmentations shown for each segmentation method. The Contour Transformation and Synthetic CT Segmentation methods generated the most accurate Heart segmentation, while the anterior border of the Direct CBCT Segmentation method was not consistently accurate.

Quantitative results

The results showing the accuracy of the different CBCT segmentation methods for all patients using the evaluation metrics are summarized in Table 1. The results for each cardiac structure are shown in Figure 3-Figure 6.

*Table 1 – The Dice similarity coefficient (DSC), Hausdorff Distance (HD), mean surface distance (MSD) and volume ratio (VR) for all patients for all structures for the different cardiac CBCT segmentation methods. The different CBCT segmentation methods are also compared. 3D = 3D Rigid Alignment, 4D = Rigid Alignment, dCBCT = Direct CBCT Segmentation, CTrans = Contour Transformation, sCT = Synthetic CT Segmentation. Metrics are DSC-Dice Similarity Coefficient, HD-Hausdorff Distance, MSD-Mean Surface Distance, VR-Volume Ratio. Statistical significance is shown for one-way ANOVA results: * indicates significance p≤0.05. ** indicates significance p≤0.01 *** indicates significance p≤0.001*

	Mean (± 95% Confidence Intervals)	Statistically different to 3D	Statistically different to 4D	Statistically different to dCBCT	Statistically different to CTrans	Statistically different to sCT
Dice similarity coefficient [0,1]						
3D Rigid Alignment	0.48 (0.47, 0.49)		**	***	***	***
4D Rigid Alignment	0.52 (0.50, 0.53)	**		***	**	**
Direct CBCT Segmentation	0.37 (0.36, 0.38)	***	***		***	***
Contour Transformation	0.53 (0.52, 0.55)	***	**	***		p=0.081
Synthetic CT Segmentation	0.57 (0.56, 0.58)	***	**	***	p=0.081	
Hausdorff Distance (mm)						
3D Rigid Alignment	10.9 (9.1, 12.7)		**	***	p=0.057	***
4D Rigid Alignment	9.9 (8.6, 11.2)	**		***	p=0.995	p=0.097
Direct CBCT Segmentation	17.3 (14.7, 20.0)	***	***		***	***
Contour Transformation	9.9 (8.5, 11.4)	p=0.057	p=0.995	***		p=0.36
Synthetic CT Segmentation	9.3 (7.8, 10.8)	***	p=0.097	***	p=0.36	
Mean surface distance (mm)						
3D Rigid Alignment	2.9 (2.5, 3.2)		p=0.887	***	***	p=0.072
4D Rigid Alignment	2.9 (2.6, 3.2)	p=0.887		***	***	p=0.012*
Direct CBCT Segmentation	6.3 (5.1, 7.5)	***	***		***	***
Contour Transformation	2.5 (2.2, 2.8)	***	***	***		p=0.79
Synthetic CT Segmentation	2.4 (2.0, 2.7)	p=0.072	p=0.012*	***	p=0.79	
Volume Ratio (estimated/ground truth volume)						
3D Rigid Alignment	0.81 (0.75, 0.86)		p=0.022*	p=0.053	***	p=0.379
4D Rigid	0.78 (0.72,	p=0.022*		p=0.027*	***	p=0.251

Alignment	0.85)					
Direct CBCT Segmentation	1.10 (0.87, 1.34)	p=0.053	p=0.027*		p=0.010**	p=0.074
Contour Transformation	0.72 (0.65, 0.80)	***	***	p=0.010**		p=0.104
Synthetic CT Segmentation	0.98 (0.77, 1.20)	p=0.379	p=0.251	p=0.074	p=0.104	

Table 1 shows that the Synthetic CT segmentations had higher DSC values than all other segmentation methods, all significantly higher except the Contour Transformation method. Similarly, Figure 3 shows the DSC values for the Synthetic CT Segmentation method were higher than the other segmentation methods for all cardiac structures, apart from the aortic valve and the Left Coronary Artery. Table 1 and Figure 3 also show that the Contour Transformation segmentations had significantly higher DSC values than the 3D & 4D Rigid alignment methods and the direct CBCT segmentation method. The DSC values in Table 1 are relatively low (≈0.5) which Figure 3 shows is due to the lower DSC values for smaller cardiac substructures for all segmentation methods when compared to the cardiac chamber and whole heart segmentations. Similar results have been observed when reporting on segmentations for different cardiac substructure sizes[19]. These lower DSC values are due to structures such as the coronary arteries being much smaller than the whole heart and cardiac chambers, and for the same physical shift, the DSC will be lower for smaller volumes. Figure 4 and Figure 5 do not show a similar trend in decreasing HD and MSD values for the smaller cardiac substructures, showing that, for these metrics, appears to be no major decrease in segmentation performance for the smaller cardiac substructures.

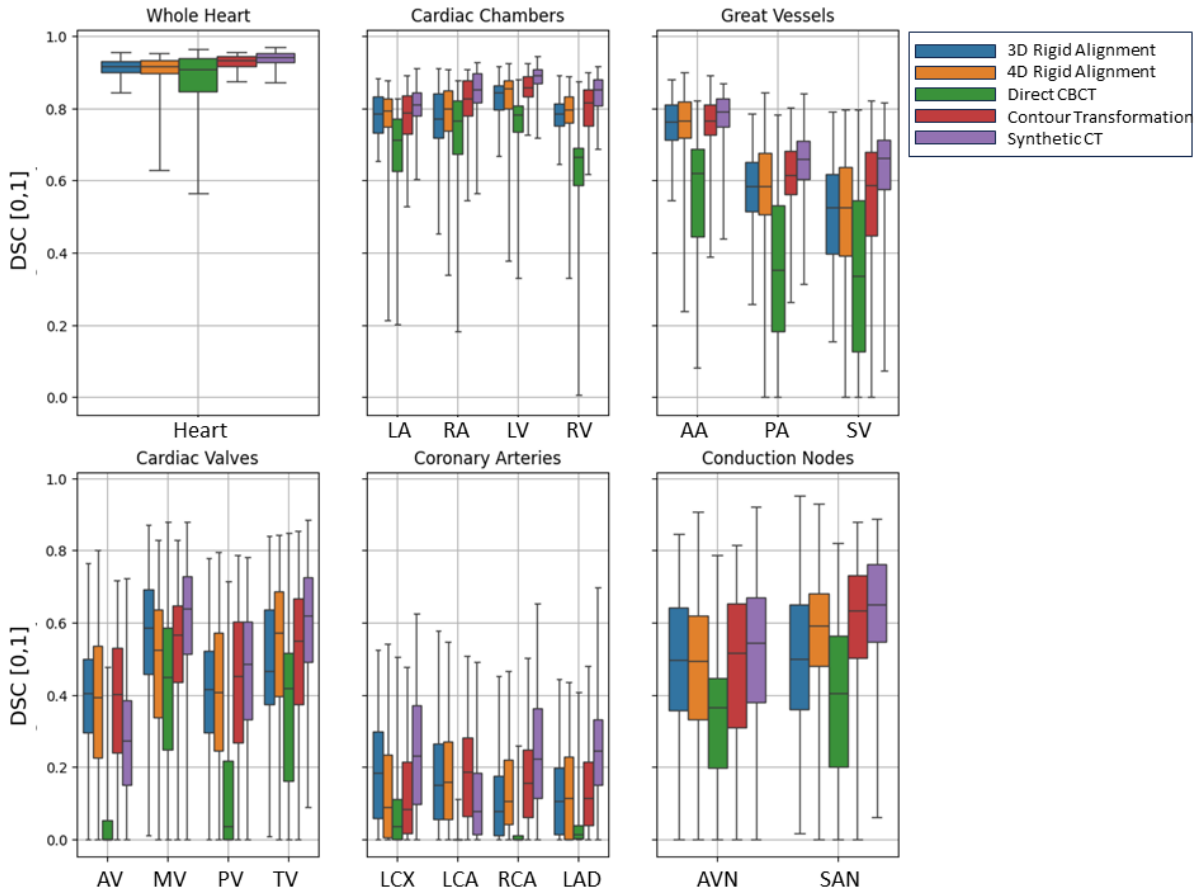


Figure 3 – The Dice Similarity Coefficient (DSC) comparing the different CBCT segmentation methods to the ground truth segmentations for each substructure segmentation for every patient. LA-Left Atrium, RA-Right Atrium, LV-Left Ventricle, RV-Right Ventricle, AA-Ascending Aorta, PA-Pulmonary Artery, SV-Superior Vena Cava, AV-Aortic Valve, MV-Mitral Valve, PV-Pulmonic Valve, TV-Tricuspid Valve, LCX-left circumflex artery, LCA-Left Coronary Artery, RCA-Right Coronary Artery, LAD-left anterior descending coronary artery, AVN-Atrial Ventricular Conduction Node, SAN-Sino Atrial Conduction Node

Table 1 shows that the 4D Rigid Alignment, Contour Transformation and Synthetic CT segmentations had the lowest HD values for all cardiac structures. Figure 4 shows the accuracy of the segmentations from the different CBCT segmentation methods was reasonably consistent for the different cardiac structures, except for the pulmonary artery which had larger HD values than the other structures. This increased HD for the pulmonary artery is due the unique shape of the pulmonary artery segmentation, leading to slight shifts in the pulmonary artery segmentation causing large increases in HD.

For the MSD metrics, Table 1 shows that the Contour Transformation and Synthetic CT segmentations had the lowest MSD when compared to the other cardiac segmentation methods. Table 1 also shows that the 3D and 4D rigid alignment methods produced segmentations with similar MSD values, while the direct CBCT segmentation method had the largest MSD. Figure 5 shows that segmentations generated by the Contour Transformation and Synthetic CT methods have similar MSD values for the different cardiac structures.

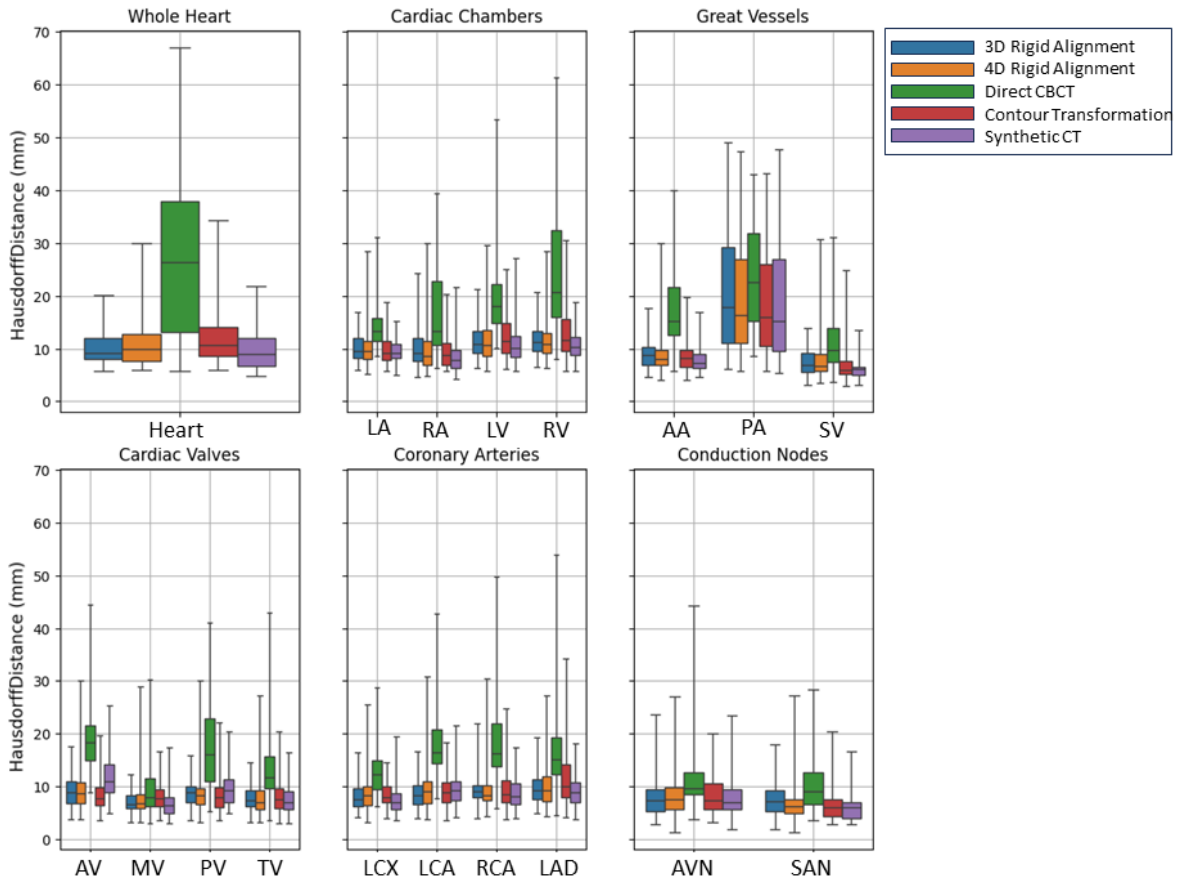


Figure 4 - The Hausdorff Distance (HD) comparing the different CBCT segmentation methods to the ground truth segmentations for each substructure segmentation for every patient. LA-Left Atrium, RA-Right Atrium, LV-Left Ventricle, RV-Right Ventricle, AA-Ascending Aorta, PA-Pulmonary Artery, SV-Superior Vena Cava, AV-Aortic Valve, MV-Mitral Valve, PV-Pulmonic Valve, TV-Tricuspid Valve, LCX-left circumflex artery, LCA-Left Coronary Artery, RCA-Right Coronary Artery, LAD-left anterior descending coronary artery, AVN-Atrial Ventricular Conduction Node, SAN-Sino Atrial Conduction Node

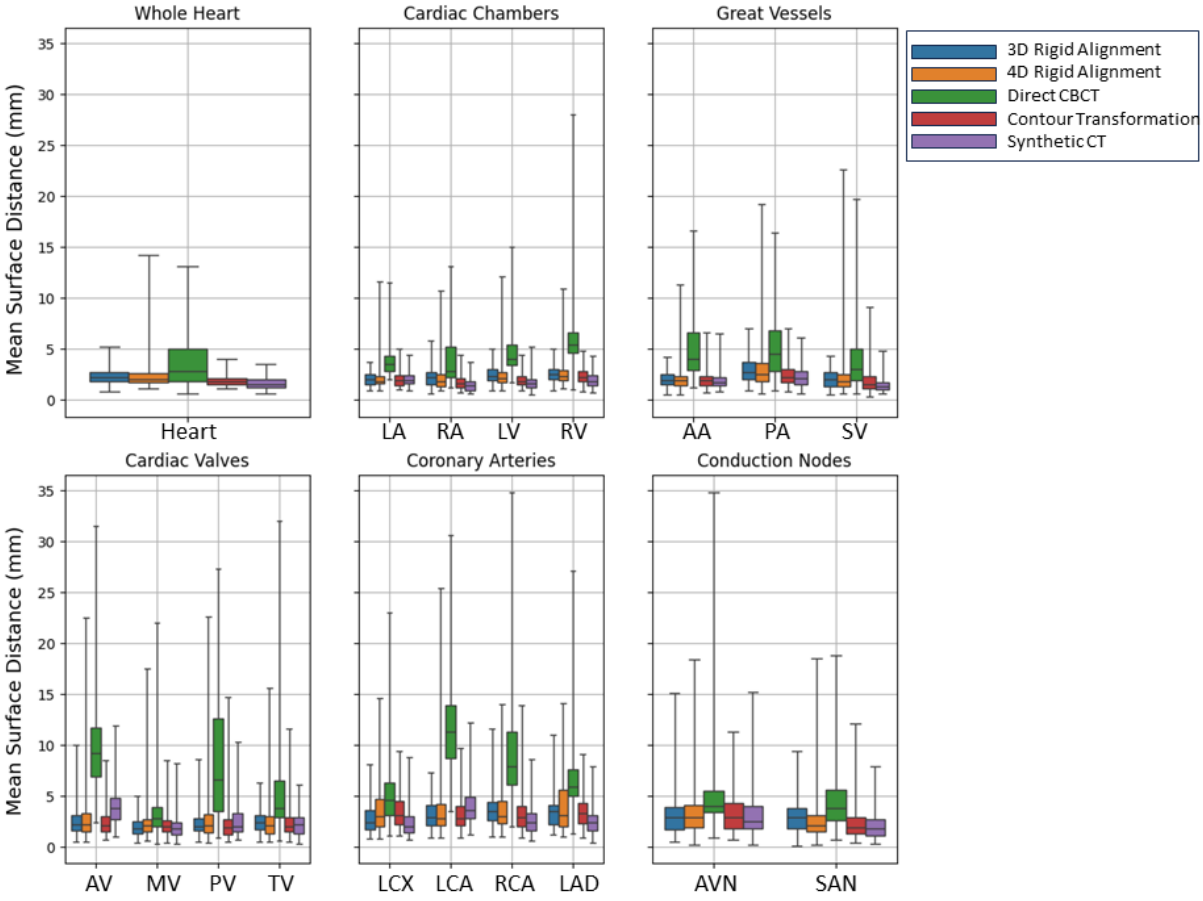


Figure 5 - The mean surface distance (MSD) comparing the different CBCT segmentation methods to the ground truth segmentations for each substructure segmentation for every patient LA-Left Atrium, RA-Right Atrium, LV-Left Ventricle, RV-Right Ventricle, AA-Ascending Aorta, PA-Pulmonary Artery, SV-Superior Vena Cava, AV-Aortic Valve, MV-Mitral Valve, PV-Pulmonic Valve, TV-Tricuspid Valve, LCX-left circumflex artery, LCA-Left Coronary Artery, RCA-Right Coronary Artery, LAD-left anterior descending coronary artery, AVN-Atrial Ventricular Conduction Node, SAN-Sino Atrial Conduction Node

The VR results for the different segmentation methods are shown in Table 1. Table 1 shows that in general the 3D Rigid Alignment, 4D Rigid Alignment and Contour Transformation methods produced volumes smaller than the ground truth segmentations. Figure 6 shows that in general the direct CBCT and Synthetic CT segmentation methods produced segmentations that had a more accurate volume, except for the aortic and pulmonic valves. These results suggest that applying the geometric-based segmentation rules that are used by *platipy* enable consistently sized segmentations that follow clinical guidelines for the smaller cardiac substructures.

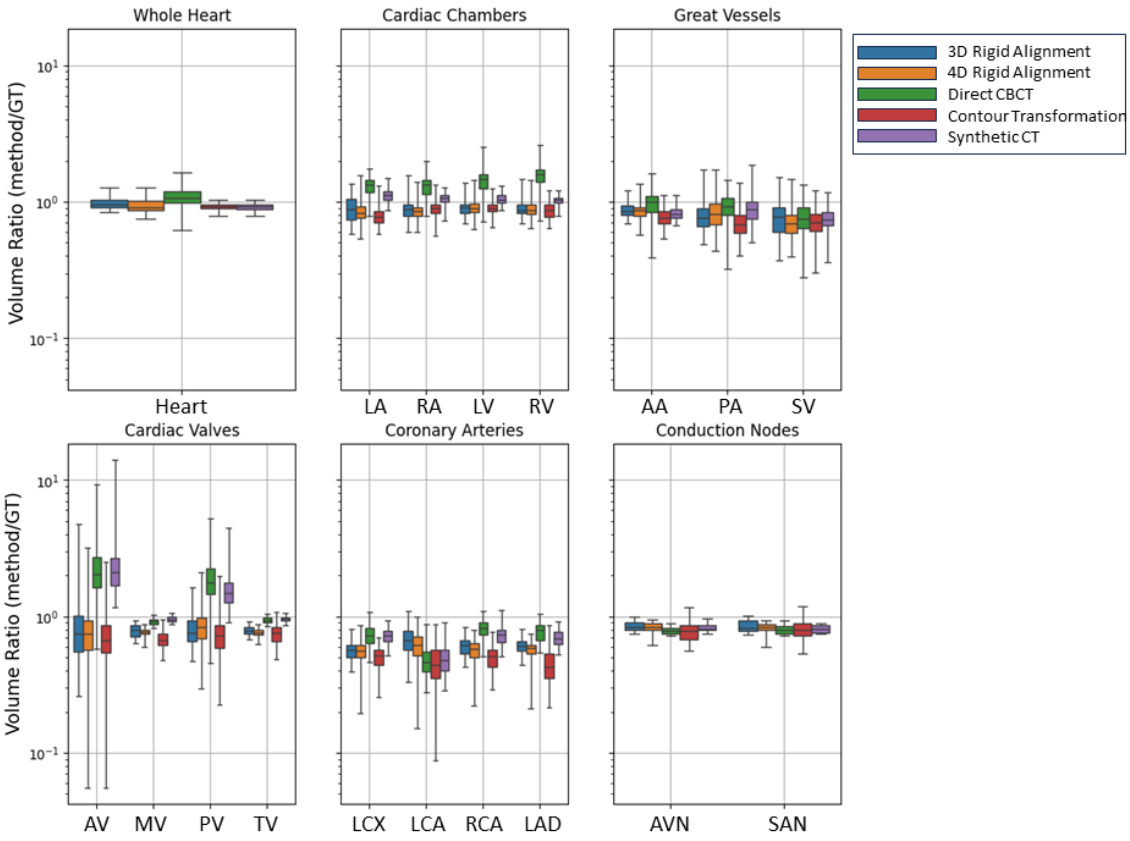


Figure 6 - The Volume ratio comparing the different CBCT segmentation methods to the ground truth segmentations for each substructure segmentation for every patient. LA-Left Atrium, RA-Right Atrium, LV-Left Ventricle, RV-Right Ventricle, AA-Ascending Aorta, PA-Pulmonary Artery, SV-Superior Vena Cava, AV-Aortic Valve, MV-Mitral Valve, PV-Pulmonic Valve, TV-Tricuspid Valve, LCX-left circumflex artery, LCA-Left Coronary Artery, RCA-Right Coronary Artery, LAD-left anterior descending coronary artery, AVN-Atrial Ventricular Conduction Node, SAN-Sino Atrial Conduction Node

In a clinical context the advantage of the Contour Transformation method over the Synthetic CT Segmentation method is that the most time-consuming step in the Contour Transformation method is the deformable registration. As deformable registration has been proposed for adaptive planning[34] it is clinically feasible to include a deformable registration step into the patient positioning procedure, especially if it leads to an increase in treatment accuracy. However, for the Synthetic CT segmentation method, once the synthetic CT has been generated, new cardiac substructure segmentations are required to be generated. Although in the future the cardiac substructure segmentation method can be optimized to reduce the processing time, currently it requires significantly more time (≈12mins) than a deformable registration (≈0.37 mins).

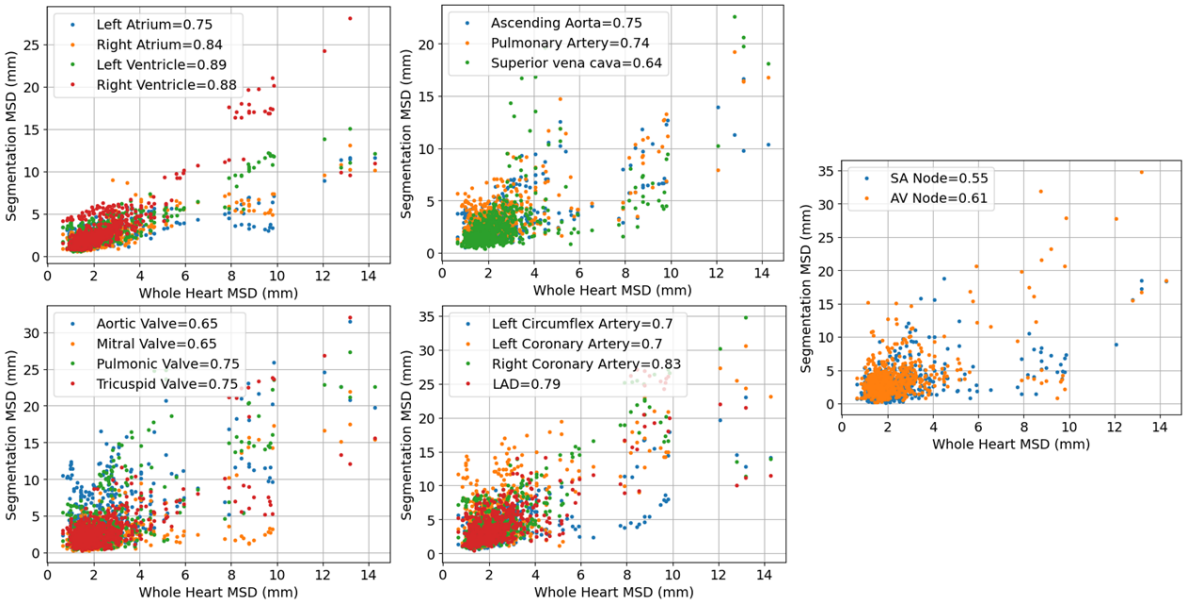


Figure 7 - Relationship between the whole heart segmentation MSD and the MSD values for each substructure segmentations. Correlation values between the whole heart MSD and each segmentations MSD value are displayed in the figure legends. MSD – mean surface distance.

The relationship between the MSD of each whole heart segmentation, and the MSD values for the other cardiac substructure segmentations are shown in Figure 7. In all of the cardiac CBCT segmentation methods described in this paper, the whole heart segmentation is generated first, which is then used to generate segmentations for the smaller cardiac substructures. Figure 7 shows the effect that an inaccurate whole heart segmentation has on the accuracy of the substructure segmentations with the correlation values indicating a strong dependency in substructure segmentation accuracy on whole heart segmentation accuracy. The substructure with the strongest correlation to the whole heart MSD was the right ventricle segmentation.

Discussion

This paper described the investigation of cardiac substructure automatic segmentation methods on 4D cone-beam CT images. The CBCT segmentation methods were compared for multiple patients and for different respiration phases. Segmentation accuracy was compared using the DSC, MSD, HD, and VR metrics. The results showed that using deformable image registration methods (such as the Contour Transformation and Synthetic CT segmentation methods), the accuracy of localizing different cardiac substructures in patient 4D-CBCT images could be increased when compared to estimating cardiac substructure locations in CBCT images via rigid registration.

The results showed that the Synthetic CT Segmentation method was able to consistently produce the most accurate cardiac substructure segmentations in 4D-CBCT images. The use of the deformable registration step in the Synthetic CT segmentation method compensated for the change in patient anatomy between the planning 4D-CT and the simulated 4D-CBCT images. Additionally, when compared with the Contour Transformation method, using *platipy* to generate the cardiac substructure segmentations in the Synthetic CT segmentation method ensured that volumes of the smaller substructures (i.e., cardiac valves and coronary arteries) were more accurately defined using the geometric rules as defined by contouring guidelines.

The *platipy* CT segmentation software has previously been validated on breast and lung CT scans[19,20,35]. In the paper describing the development of the *platipy* algorithm the mean whole heart

MSD and DSC values for of 2.16 mm and 0.93 respectively for 30 patients[19]. Similarly, Olloni *et al.* evaluated the accuracy of *platipy* on 12 breast cancer CT scans, and the MSD and DSC values for the whole heart segmentation were 1.5 mm and 0.95[20]. In comparison, for the Contour Transformation method the whole heart MSD and DSC values were 1.91 mm and 0.93 respectively and for the Synthetic CT Segmentation method the whole heart MSD and DSC values were 1.68 mm and 0.94 respectively. This result shows that the accuracy of CBCT segmentations generated by the Contour Transformation and Synthetic CT Segmentation methods is similar to the accuracy of the segmentations generated from CT images.

Balgobind et al. compared the manual contouring accuracy of 31 cardiac and surrounding structures from 20 radiation oncology centers on CT images using the DSC, median distance to agreement (MDA) and HD metrics [28]. For the larger structures such as the whole heart and cardiac chambers, the median DSC values for the Contour Transformation and Synthetic CT segmentation methods were similar to the DSC values for the manual contours (0.93 and 0.87 respectively) but the MDA and HD values were smaller (<0.5mm for MDA and <4.5mm for HD). Similarly Milo et al. compared manual contouring accuracy of 15 observers from 4 radiation oncology centers. For the whole heart and cardiac chamber segmentations the median DSC values were approximately 0.9 and 0.8 respectively which is similar to the reported DSC values for the Contour Transformation and Synthetic CT segmentation methods reported in this paper [27]. As discussed previously both the comparisons by Balgobind et al and Milo et al. showed that the large inter-observer variability caused larger errors in the manual contours for smaller cardiac substructures such as the coronary arteries and heart valves [27,28].

The Contour Transformation and Synthetic CT segmentation methods are highly dependent on the accuracy of the deformable registration method used, which for this paper was a multi-resolution B-spline based deformable registration method, with a mutual information loss function. Previous investigations have found that the B-spline registration method to be an accurate and reliable deformable image registration method[36,37]. Areas of poor image quality[38] or areas of the same pixel value[39] result in poorer registration accuracy, so if another registration method is used then it will also need to be robust to these sources of error. While the algorithm to implement the B-spline registration used in this paper to generate the synthetic CT images was implemented using open-source software, this approach to generating synthetic CT images is similar to the commercially available *VelocityAI* software (Varian Medical Systems, Palo Alto, CA). While there may be some small differences between the performance of the *Elastix* software and *VelocityAI*[40], we can expect similar synthetic CT accuracy to the results in this paper when clinically implemented using the clinical software.

The results in Figure 6 show that the different CBCT segmentation methods underestimated the total volume of the different cardiac substructures. This underestimation of the cardiac substructure volumes is particularly evident for the smaller substructures such as the cardiac valves and conduction nodes, which are defined as a constant size, and so the VR metrics for these geometric-based substructures for all methods (apart from the transformation methods) should be equal. This may be due to the different voxel sizes between the CT and CBCT images, as the slice thickness for the CT scans was 3 mm, but the voxel spacing in the SI direction was 1 mm. The quantitative VR metrics of the different segmentation methods may hence be improved if the voxel size of the CT volumes was the same as the CBCT volumes being compared.

A limitation of this study is that synthetic 4D-CBCT images were generated from patient 4D-CT images, rather than direct 4D-CBCT patient measurements. This step was necessary as a ground truth for the cardiac substructure segmentation is required, and given the relatively poor quality of CBCT

images, it can be difficult to identify cardiac substructures in 4D-CBCT images, even for experienced clinicians. Since for the Contour Transformation and Synthetic CT Segmentation methods a deformable registration is used to generate the cardiac substructure segmentations, this same method can't be used to also generate the ground truth segmentations by deformably registering real patient 4D-CT images to real 4D-CBCT images. These synthetic CBCT images are idealized reconstructions and would have better image quality that real CBCT images. Similarly, factors such as the conversion from attenuation coefficients to HU values, the lack of intra-bin motion for the different respiration phases, and the absence of noise sources such as scatter and beam hardening can reduce the quality of 4D-CBCT reconstructions which would potentially reduce the accuracy of the cardiac segmentation methods described in this paper. Future work will look at validating these CBCT segmentation methods on real patient CBCT images.

There have been other cardiac segmentation methods developed, including methods for segmenting the heart into 17 segments according to the American Heart Association 17-segment model[41]. For the Contour Transformation and Synthetic CT Segmentation methods described in this paper, the methods are not dependent on the specific cardiac substructures or segmentation model that were shown in the results of this paper. Hence any method for segmenting the heart in a CT scan should be able to be applied with a similar level of accuracy to segmenting CBCT scans using either the Contour Transformation or Synthetic CT scan method, assuming the cardiac segmentation method used has also been robustly validated.

Figure 4 and Figure 5 show that using the HD and MSD metrics respectively, the Direct CBCT segmentation method had significantly higher HD and MSD values for the overall heart segmentation than the 3D rigid alignment method, which suggests that the poor performance of the direct CBCT segmentation could be due to the deep learning segmentation of the whole heart not being accurate. For this reason we produce Figure 7, investigating specifically the correlation between whole heart segmentation MSD and the corresponding MSD for smaller substructures. We observe that 64% to 89% of the variation in MSD error for smaller substructures can be predicted from the MSD error in the whole heart segmentation, emphasizing how critical whole heart segmentation is for all other substructures.

Whilst the deep learning-based method was not able to segment cardiac substructures consistently and accurately on the simulated CBCT images, future work could investigate developing and training new segmentation networks which are able to segment cardiac substructures directly from the CBCT image. These deep learning-based CBCT segmentation networks would need to be trained using CBCT images with varying levels of noise, with features that can overcome artefacts such as streaking artefacts.

There have been several studies that have found that for STAR, both respiration and cardiac motion contribute significantly to the overall target volume motion[5–8]. ECG and respiration gating has been proposed for treatments using MRI-Linacs, which can extend the treatment time[42]. Akdag *et al.* investigated the effect of cardiac gating on treatment accuracy and on the treatment duty cycle[43], as well as combining ECG-gating with Multi-leaf collimator tracking for managing respiration motion[44] and found that it although it would increase the time, it was clinically feasible. However, MRI-Linacs are expensive, and as such are not as common clinically, when compared to conventional linacs. Successful implementation of cardiac substructure segmentation in planning 4D-CT and pre-treatment 4D-CBCT images would allow for management of respiratory motion, which would allow for cardiac-gated STAR treatment without the need for an MRI-Linac.

Conclusion

This paper described the investigation of cardiac substructure automatic segmentation methods on synthetic 4D cone-beam CT images. The results showed that the developed cardiac CBCT segmentation methods were more accurate than 3D and 4D rigid alignment of the contours from the planning CT. By simulating 4D-CBCT pre-treatment images, and accurately segmenting the cardiac substructures in these 4D-CBCT images, the target volume can be more clearly identified prior to STAR treatments. More accurate targeting in cardiac CBCT treatments will lead to more effective treatments and lower radiation-related treatment toxicities. As the data and segmentation methods used in this paper are all open source, this paper provides a useful benchmarking tool to evaluate other CBCT cardiac segmentation methods.

Acknowledgements

The authors wish to thank Helen Ball, Ricky O'Brien, Chen Cheng and Hunor Kertész for reviewing the manuscript and to Helen Ball for providing guidance in the statistical analysis of the results.

References

1. Wei C, Boeck M, Qian PC, et al. Cost of cardiac stereotactic body radioablation therapy versus catheter ablation for treatment of ventricular tachycardia. *PACE - Pacing Clin Electrophysiol*. 2022;45(9):1124-1131. doi:10.1111/pace.14512

2. Robinson CGCG, Samson PPP, Moore KMSKMS, et al. Phase I/II Trial of Electrophysiology-Guided Noninvasive Cardiac Radioablation for Ventricular Tachycardia. *Circulation*. 2019;139(3):313-321. doi:10.1161/CIRCULATIONAHA.118.038261

3. Hašková J, Wichterle D, Kautzner J, et al. Efficacy and Safety of Stereotactic Radiotherapy in Patients With Recurrent Ventricular Tachycardias: The Czech Experience. *JACC Clin Electrophysiol*. Published online February 21, 2024. doi:10.1016/J.JACEP.2023.12.002

4. Qian PC, Quadros K, Aguilar M, et al. Substrate Modification Using Stereotactic Radioablation to Treat Refractory Ventricular Tachycardia in Patients With Ischemic Cardiomyopathy. *JACC Clin Electrophysiol*. 2022;8(1):49-58. doi:10.1016/J.JACEP.2021.06.016

5. Lydiard S, Pontré B, Lowe BS, Ball H, Sasso G, Keall P. Cardiac radioablation for atrial fibrillation: Target motion characterization and treatment delivery considerations. *Med Phys*. 2021;48(3):931-941. doi:10.1002/MP.14661

6. Prusator MT, Samson P, Cammin J, et al. Evaluation of Motion Compensation Methods for Noninvasive Cardiac Radioablation of Ventricular Tachycardia. *Int J Radiat Oncol Biol Phys*. 2021;111(4):1023-1032. doi:10.1016/j.ijrobp.2021.06.035

7. Stevens RRF, Hazelaar C, Fast MF, et al. Stereotactic Arrhythmia Radioablation (STAR): Assessment of cardiac and respiratory heart motion in ventricular tachycardia patients - A STOPSTORM.eu consortium review. *Radiother Oncol*. 2023;188. doi:10.1016/j.radonc.2023.109844

8. Harms J, Schreibmann E, Mccall NSNS, Lloyd MSMS, Higgins KAKA, Castillo R. Cardiac motion and its dosimetric impact during radioablation for refractory ventricular tachycardia. *J Appl Clin Med Phys*. 2023;24(6):e13925. doi:10.1002/ACM2.13925

9. Guckenberger M, Andratschke N, Dieckmann K, et al. ESTRO ACROP consensus guideline on implementation and practice of stereotactic body radiotherapy for peripherally located early stage non-small cell lung cancer. *Radiother Oncol*. 2017;124(1):11-17.

doi:10.1016/J.RADONC.2017.05.012

10. Lydiard S, Blanck O, Hugo G, O'Brien R, Keall P. A Review of Cardiac Radioablation (CR) for Arrhythmias: Procedures, Technology, and Future Opportunities. *Int J Radiat Oncol*. 2021;109(3):783-800. doi:10.1016/J.IJROBP.2020.10.036

11. Roman NO, Shepherd W, Mukhopadhyay N, Hugo GD, Weiss E. Interfractional Positional Variability of Fiducial Markers and Primary Tumors in Locally Advanced Non-Small-Cell Lung Cancer During Audiovisual Biofeedback Radiotherapy. *Int J Radiat Oncol*. 2012;83(5):1566-1572. doi:10.1016/J.IJROBP.2011.10.051

12. Balik S, Weiss E, Jan N, et al. Evaluation of 4-dimensional Computed Tomography to 4-dimensional Cone-Beam Computed Tomography Deformable Image Registration for Lung Cancer Adaptive Radiation Therapy. *Int J Radiat Oncol*. 2013;86(2):372-379. doi:10.1016/J.IJROBP.2012.12.023

13. Hugo GD, Weiss E, Sleeman WC, et al. A longitudinal four-dimensional computed tomography and cone beam computed tomography dataset for image-guided radiation therapy research in lung cancer. *Med Phys*. 2017;44(2):762-771. doi:10.1002/MP.12059

14. Kipritidis J, Hugo G, Weiss E, Williamson J, Keall PJ. Measuring interfraction and intrafraction lung function changes during radiation therapy using four-dimensional cone beam CT ventilation imaging. *Med Phys*. 2015;42(3):1255-1267. doi:10.1118/1.4907991

15. Blake SJ, Dillon O, Byrne HL, O'Brien RT. Thoracic motion-compensated cone-beam computed tomography in under 20 seconds on a fast-rotating linac: A simulation study. *J Appl Clin Med Phys*. Published online January 21, 2023:e13909. doi:10.1002/ACM2.13909

16. Gardner M, Dillon O, Byrne H, Keall PJ, O'Brien R. Data-driven rapid 4D Cone-Beam CT reconstruction for new generation linacs. *Phys Med Biol*. Published online September 6, 2024. doi:10.1088/1361-6560/AD780A

17. Kamalian S, Lev MH, Gupta R. Computed tomography imaging and angiography – principles. *Handb Clin Neurol*. 2016;135:3-20. doi:10.1016/B978-0-444-53485-9.00001-5

18. Chlap P, Finnegan RN. PlatiPy: Processing Library and Analysis Toolkit for Medical Imaging in Python. *J Open Source Softw*. 2023;8(86):5374. doi:10.21105/JOSS.05374

19. Finnegan RN, Chin V, Chlap P, et al. Open-source, fully-automated hybrid cardiac substructure segmentation: development and optimisation. *Phys Eng Sci Med*. 2023;46(1):377-393. doi:10.1007/S13246-023-01231-W/FIGURES/9

20. Olloni A, Lorenzen EL, Jeppesen SS, et al. An open source auto-segmentation algorithm for delineating heart and substructures – Development and validation within a multicenter lung cancer cohort. *Radiother Oncol*. 2024;191:110065. Accessed January 8, 2024. http://www.thegreenjournal.com/article/S0167814023093726/fulltext

21. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods 2020 182*. 2020;18(2):203-211. doi:10.1038/s41592-020-01008-z

22. Finnegan R, Dowling J, Koh ES, et al. Feasibility of multi-atlas cardiac segmentation from thoracic planning CT in a probabilistic framework. *Phys Med Biol*. 2019;64(8):085006. doi:10.1088/1361-6560/AB0EA6

23. Duane F, Aznar MC, Bartlett F, et al. A cardiac contouring atlas for radiotherapy. *Radiother Oncol*. 2017;122(3):416-422. doi:10.1016/J.RADONC.2017.01.008

24. Socha J, Rygielska A, Uziębło-Życzkowska B, et al. Contouring cardiac substructures on average intensity projection 4D-CT for lung cancer radiotherapy: A proposal of a heart valve contouring atlas. *Radiother Oncol*. 2022;167:261-268. doi:10.1016/J.RADONC.2021.12.041

25. Loap P, Servois V, Dhonneur G, Kirov K, Fourquet A, Kirova Y. A Radiation Therapy Contouring Atlas for Cardiac Conduction Node Delineation. *Pract Radiat Oncol*. 2021;11(4):e434-e437. doi:10.1016/J.PRRO.2021.02.002

26. Finnegan RNR, Quinn A, Booth J, et al. Cardiac substructure delineation in radiation therapy – A state-of-the-art review. *J Med Imaging Radiat Oncol*. Published online May 17, 2024. doi:10.1111/1754-9485.13668

27. Milo MLH, Offersen BV, Bechmann T, et al. Delineation of whole heart and substructures in thoracic radiation therapy: National guidelines and contouring atlas by the Danish Multidisciplinary Cancer Groups. *Radiother Oncol*. 2020;150:121-127. doi:10.1016/J.RADONC.2020.06.015

28. Balgobind B V., Visser J, Grehn M, et al. Refining critical structure contouring in STereotactic Arrhythmia Radioablation (STAR): Benchmark results and consensus guidelines from the STOPSTORM.eu consortium. *Radiother Oncol*. 2023;189:109949. doi:10.1016/J.RADONC.2023.109949

29. Vinod SK, Min M, Jameson MG, Holloway LC. A review of interventions to reduce inter-observer variability in volume delineation in radiation oncology. *J Med Imaging Radiat Oncol*. 2016;60(3):393-406. doi:10.1111/1754-9485.12462

30. Klein S, Staring M, Murphy K, Viergever MA, Pluim JPW. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging*. 2010;29(1):196-205. doi:10.1109/TMI.2009.2035616

31. Scholz EP, Seidensaal K, Naumann P, André F, Katus HA, Debus J. Risen from the dead: Cardiac stereotactic ablative radiotherapy as last rescue in a patient with refractory ventricular fibrillation storm. *Hear Case Reports*. 2019;5(6):329-332. doi:10.1016/J.HRCR.2019.03.004

32. Milano MT, Marks LB, Constine LS. Survivorship and Late Effects. In: *Clinical Radiation Oncology, Third Edition*. W.B. Saunders; 2012:239-256. doi:10.1016/B978-1-4377-1637-5.00013-4

33. Chin V, Finnegan RN, Chlap P, et al. Dosimetric Impact of Delineation and Motion Uncertainties on the Heart and Substructures in Lung Cancer Radiotherapy. *Clin Oncol*. 2024;36(7):420-429. doi:10.1016/J.CLON.2024.04.002

34. Lim-Reinders S, Keller BM, Al-Ward S, Sahgal A, Kim A. Online Adaptive Radiation Therapy. *Int J Radiat Oncol*. 2017;99(4):994-1003. doi:10.1016/J.IJROBP.2017.04.023

35. Chin V, Finnegan RN, Chlap P, et al. Validation of a Fully Automated Hybrid Deep Learning Cardiac Substructure Segmentation Tool for Contouring and Dose Evaluation in Lung Cancer Radiotherapy. *Clin Oncol*. 2023;35(6):370-381. doi:10.1016/j.clon.2023.03.005

36. Kirby N, Chen J, Kim H, Morin O, Nie K, Pouliot J. An automated deformable image registration evaluation of confidence tool. *Phys Med Biol*. 2016;61(8):N203. doi:10.1088/0031-9155/61/8/N203

37. Mohamed ASR, Ruangskul MN, Awan MJ, et al. Quality assurance assessment of diagnostic and radiation therapy-simulation CT image registration for head and neck radiation therapy: anatomic region of interest-based comparison of rigid and deformable algorithms. *Radiology*. 2015;274(3):752-763. doi:10.1148/RADIOL.14132871

38. Veiga C, Lourenço AM, Mouinuddin S, et al. Toward adaptive radiotherapy for head and neck patients: Uncertainties in dose warping due to the choice of deformable registration algorithm. *Med Phys*. 2015;42(2):760-769. doi:10.1118/1.4905050

39. Singhrao K, Kirby N, Pouliot J. A three-dimensional head-and-neck phantom for validation of multimodality deformable image registration for adaptive radiotherapy. *Med Phys*. 2014;41(12):121709. doi:10.1118/1.4901523

40. Zhong H, Adams J, Glide-Hurst C, Zhang H, Li H, Chetty I. Development of a deformable dosimetric phantom to verify dose accumulation algorithms for adaptive radiotherapy. *J Med Phys*. 2016;41(2):106-114. doi:10.4103/0971-6203.181641

41. van der Ree MH, Visser J, Planken RN, et al. Standardizing the Cardiac Radioablation Targeting Workflow: Enabling Semi-Automated Angulation and Segmentation of the Heart According to the American Heart Association Segmented Model. *Adv Radiat Oncol*. 2022;7(4). doi:10.1016/j.adro.2022.100928

42. Mayinger M, Kovacs B, Tanadini-Lang S, et al. First magnetic resonance imaging-guided cardiac radioablation of sustained ventricular tachycardia. *Radiother Oncol*. 2020;152:203-207. doi:10.1016/j.radonc.2020.01.008

43. Akdag O, Borman PTS, Mandija S, et al. Experimental demonstration of real-time cardiac physiology-based radiotherapy gating for improved cardiac radioablation on an MR-linac. *Med Phys*. 2024;51(4):2354-2366. doi:10.1002/MP.17024

44. Akdag O, Borman PTS, Woodhead P, et al. First experimental exploration of real-time cardiorespiratory motion management for future stereotactic arrhythmia radioablation treatments on the MR-linac. *Phys Med Biol*. 2022;67(6). doi:10.1088/1361-6560/AC5717