January 2020

Incorporating Cardiac Substructures Into Radiation Therapy For Improved Cardiac Sparing

Eric Daniel Morris
Wayne State University

INCORPORATING CARDIAC SUBSTRUCTURES INTO RADIATION THERAPY FOR IMPROVED CARDIAC SPARING

by

ERIC DANIEL MORRIS

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2020

MAJOR: MEDICAL PHYSICS

Approved By:

Advisor Date

DEDICATION

To my family for their never-ending encouragement.

To the many friends who have always pushed me to do better and for their constant support.

Most importantly, to my wonderful wife Rebecca for loving me, teaching me, and sacrificing for me constantly throughout this journey.

ACKNOWEDGEMENTS

I would like to thank my dissertation committee: Dr. Indrin J. Chetty, Dr. Ewart Haacke, Dr. Ming Dong, Dr. Jacob Burmeister, and in particular my advisor Dr. Carri Glide-Hurst for their continued support, never ending encouragement, and constructive criticism throughout this process. I would also like to thank the rest of our research group: Dr. Siamak Nejad-Davarani, Dr. Ryan Price, Dr. Josh Kim, Dr. Hoda Sharifi, and Claudia Miller for the many insightful discussions and for sharing their expertise. Lastly, I would like to thank Dr. Ahmed Ghanem, Dr. Simeng Zhu, and Kate Aldridge for their generous contributions to this work. Henry Ford Health System holds research agreements with Philips Healthcare. Research sponsored in part by an HFHS Internal Mentored Grant (C. Glide-Hurst), NIH R01CA204189, and NIH R01EB016079. A portion of the data acquisition costs were supported by the Breast Cancer Research Foundation.

PREFACE

Note to the reader:

Chapters 3 and 4 presented in this dissertation have been previously published in peer reviewed journals and each of the two parts of Chapters 5 and 6 are currently being prepared for journal submission. As these manuscripts have been previously published or currently being prepared for submission, they were originally intended to serve as solitary documents. However, additional information has been added to assemble these manuscripts here.

Chapter 3 was originally published as “Cardiac Substructure Segmentation and Dosimetry Using a Novel Hybrid Magnetic Resonance and Computed Tomography Cardiac Atlas” in 2018 in the International Journal of Radiation Oncology Biology Physics.

Chapter 4 was originally published as “Cardiac Substructure Segmentation with Deep Learning for Improved Cardiac Sparing” in 2019 in the Journal of Medical Physics. This manuscript received Editor’s choice and was selected for the journal’s cover art.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3D.....Three-dimensional

4DCT.....Four-dimensional CT

A-P ..... Anterior-Posterior

AA ..... Ascending aorta

b-SSFP.....Balanced steady state free precession

CBCT ..... Cone beam computed tomography

CNN ..... Convolutional neural network

CRF.....Conditional random fields

CT .....Computed tomography

CT-SIM.....Computed tomography simulation

CTCA ..... Computed tomography coronary angiography

DIBH.....Deep inspiration breath hold

DICOM .....Digital imaging and communications in medicine

DIR ..... Deformable image registration

DL.....Deep learning

DNN ..... Deep neural network

DSC..... Dice similarity coefficient

DVH.....Dose volume histogram

ECG ..... Electrocardiogram

EE ..... End-exhalation

EI..... End-inhalation

FBCT..... Free breathing CT

GAN ..... Generative adversarial network

GPU ..... Graphics processing unit

GT ..... Ground truth

HD ..... Hausdorff distance

HU ..... Hounsfield unit

IGRT ..... Image guided radiation therapy

IMRT ..... Intensity modulated radiation therapy

IVC ..... Inferior vena cava

L-R ..... Left-Right

LA ..... Left atrium

LADA ..... Left anterior descending artery

LMCA ..... Left main coronary artery

LV ..... Left ventricle

LV-V5 ..... Left ventricular volume receiving 5 Gy

MA ..... Multi-atlas

MDA ..... Mean distance to agreement

MHD ..... Mean heart dose

MR ..... Magnetic resonance

MR-linac ..... Magnetic resonance linear accelerator

MR-SIM ..... Magnetic resonance simulation

MRgRT ..... Magnetic resonance-guided radiation therapy

MRI ..... Magnetic resonance imaging

MV ..... Majority vote

MV10 ..... MV with 10 atlas matches

NSCLC ..... Non-small cell lung cancer

OAR ..... Organ at risk

PA ..... Pulmonary artery

PRV ..... Planning organ at risk volume

PTV ..... Planning target volume

PV ..... Pulmonary veins

QUANTEC ..... Qualitative analysis of normal tissue effects in the clinic

RA ..... Right atrium

RCA ..... Right coronary artery

RICT ..... Radiation induced cardiac toxicity

RT ..... Radiation therapy

RTOG ..... Radiation Therapy Oncology Group

RTP ..... Radiotherapy treatment planning

RV ..... Right ventricle

S-I ..... Superior-Inferior

SBRT ..... Stereotactic body radiation therapy

SD ..... Standard deviation

ST10 ..... STAPLE method with 10 atlas matches

STAPLE ..... Simultaneous Truth and Performance Level Estimation

SVC ..... Superior vena cava

VMAT ..... Volumetric modulated arc therapy

CHAPTER 1 "CLINICAL MOTIVATION AND PROBLEM STATEMENT"

Cardiotoxicity from Radiation Therapy

Radiation therapy (RT) is a beneficial treatment option for approximately half of all cancer patients and is recognized as a crucial component of treating cancer throughout the world[1]. Over a quarter of all diagnosed cancers involve the thoracic region where there is neighboring cardiac normal tissue, and these cancers remain the most common cancer-related cause of death among American men and women[2]. Weather cancers in the thoracic region are localized or locally advanced, applications of RT have allowed for curative and palliative treatment options[3]. However, RT can lead to secondary effects due to the neighboring cardiac normal tissues within the irradiation field and cause cardiac toxicity. Cardiac toxicity is a potentially devastating complication of cancer treatment and occurs throughout, shortly after, or even many years after treatment[4]. Increased risks of radiation-induced cardiac toxicities (RICTs) including acute (e.g. pericarditis) and late (e.g. congestive heart failure, coronary artery disease, and myocardial infarction) have been linked to dose from RT for many thoracic cancers including Hodgkin's lymphoma[5], esophageal[6], late stage lung[7], and breast[8].

In Hodgkin's Lymphoma patients who have received RT, cardiovascular disease is the most common cause of death including coronary artery disease, valvular heart disease, congestive heart failure, pericardial disease, and sudden death[9]. Advanced stage lung cancer survivors, who undergo some of the highest doses of RT to the heart, exhibit the worst comorbidities across all cancers, with congestive heart failure being prevalent[10]. Patients with centrally located lung tumors have also experienced cardiac failure and pericarditis after stereotactic body therapy (SBRT)[4,11,12]. In a study by Hardy

et al., it was found that ischemic heart disease was more common in patients with tumors of the left lung as compared to the right after RT treatment[7]. Furthermore, a statistically significantly larger risk for radiation-associated coronary damage was also found in left-sided early stage breast cancer patients, as compared to right-sided[8]. Several studies have also shown that perfusion defects have been linked to excess cardiac dose[13,14]. Marks *et al.* found that 40% of left-sided breast cancer patients had perfusion defects from RT within just two years of their RT treatment[13]. If not addressed, the aforementioned secondary cardiac effects from these thoracic cancer treatments may lead to ischemic heart disease and even heart failure[15].

RICT is more acute than previously expected, beginning only a few years after RT and with elevated risk persisting for nearly 20 years[16]. Moreover, echocardiograms from thoracic RT treatments have revealed real time changes due to the radiation[14]. With life expectancy in cancer survivors steadily improving (i.e. patients living to see the long-term cardiac effects of their treatment), it becomes of paramount importance to mitigate RICT while still optimizing cancer outcomes.

Current State-of-the-Art for Cardiac Risk Assessment

When an RT plan is being developed for a patient with cancer, the current standard of care is to only delineate and consider the entire heart as a single organ and use simple metrics like mean heart dose (MHD) and dose/volume relationships to evaluate cardiac risks. Importantly, these whole-heart dose metrics do not provide any information about where the dose is being distributed. The Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC)[17] report provides radiation dose tolerance recommendations for organs at risk (OARs) via efforts from numerous investigators. Here, Gagliardi *et al.*

considers dose to the heart as a single organ and recommends that less than 10% of the heart receive greater than 25 Gy with the endpoint of long term cardiac mortality[18]. Contemporary cooperative trials use similar volumetric and MHD endpoints[19-21]. In a large population-based case control study with greater than 2,000 women undergoing breast cancer RT[16], Darby *et al.* found that cardiac damage was correlated with heart-absorbed dose, with a 7.4% increase of ischemic heart disease risk per one Gray of dose received[16]. It was also found that the MHD from left-sided breast cancer treatment was 5.4 Gy (range, < 0.1 to 28.6 Gy)[22], which suggests a ~40% increase in relative heart disease risk.

As the outcome remains poor for patients with locally advanced non-small cell lung cancer (NSCLC), for example a 5-year overall survival less than 20%[23], there have been copious efforts[24] to increase loco-regional control (i.e. objective tumor response plus freedom from local progression)[25] in lung cancer. Namely, dose escalation, or increasing the total dose prescribed through the course of RT, has been used to try and increase loco-regional control, although this has come at a cost. In a recent dose escalation trial for locally advanced NSCLC (Radiation Therapy Oncology Group (RTOG) report 0617), heart volumes receiving ≥ 5 and ≥ 30 Gy were independent predictors of overall survival[26] and a patients' quality of life[27]. Thus, efforts to reduce and better characterize radiation dose in the heart, particularly in NSCLC where dose escalation is being implemented, are advantageous.

Importance of Cardiac Substructures

The heart is complex and dose to sensitive substructures (e.g., coronary arteries, ventricles, atria, great vessels, etc.) contained within the heart have been strongly linked

to RICT[16,28,29]. Radiation dose to the left anterior descending coronary artery (LADA) has been linked to an increased risk of radiation-induced cardiac morbidity[30], myocardial infarction[31], and development of coronary artery calcifications[32]. In a study by Kataria *et al.*, a dependence between RICT and the maximum dose received by the LADA was found[33]. They recommended that the maximum dose to the LADA be implemented for OAR avoidance, rather than the mean dose, as it may be more analogous to a serial structure[33]. Additionally, Hahn *et al.* showed that there were dose-volume indices for the coronary arteries (i.e. V5 and V20) that were more accurate than the MHD in predicting ischemic heart disease risk[34].

The volume of the left ventricle (LV) receiving 5 Gy (LV-V5) has been shown to be more predictive of acute cardiac events than the MHD[28]. Further, radiation damage to the LV has been strongly associated with future acute coronary events[28]. Similarly, higher doses at the base of the heart, near the great vessels (ascending aorta (AA), superior vena cava (SVC), and pulmonary artery (PA)) are directly related to worse patient survival[35]. Lastly, the radiation induced affect in the heart will depend on where the dose is delivered. For example, a study by Wang *et al.* assessed 112 NSCLC patients with a 8.8 year median follow up[36]. This study found that pericardial events were strongly correlated with atrial dose, whereas, ischemic events were strongly correlated with ventricular dose[36].

There have been numerous multi-institutional studies aiming at assessing the relationship between cardiac substructures and radiation dose. In a multi-institutional cohort of nearly 800 SBRT lung cancer patients, doses to the left atrium (maximum dose), SVC, atria, and vessels were significantly associated with non-cancer death[37]. In

Hodgkin's Lymphoma patients, the relative risk of death from cardiac diseases was substantially decreased using subcarinal blocking[38] while pericarditis was reduced when left ventricular and subcarinal areas were shielded[39]. Importantly, recent sub-analysis of a cooperative group trial for NSCLC (RTOG 0617), showed that doses to the atria, pericardium, and ventricles were more strongly associated with survival than standard dose/volume heart metrics[40-42]. Overall, mounting evidence suggests that local doses to sensitive regions within the heart are strongly associated with RICT, yet challenges exist for routine dose evaluation.

Current Challenges with Assessing Cardiac Substructures

While cardiac substructures have been shown to have importance for cardiac toxicities, these structures are not visible on standard computed tomography (CT) simulations (CT-SIMs) and thus are not typically considered in the treatment planning process[18]. Magnetic resonance imaging (MRI), on the other hand, drastically improves the visibility of the cardiac substructures[43,44] as shown in Figure 1. There are several reasons, as stated by Dweck *et al.*, that MRI is the modality of choice for visualizing cardiac substructures over CT[44]. Namely, it offers superior soft-tissue contrast, it is not affected by the calcium blooming that hampers CT, and it does not involve exposure to ionizing radiation[44]. However, most cancer patients do not undergo MRI due to high costs, lack of insurance reimbursement, and accessibility barriers. Furthermore, most radiation oncologists have limited experience delineating cardiac anatomy and manual contouring can take several hours per patient[45-48].

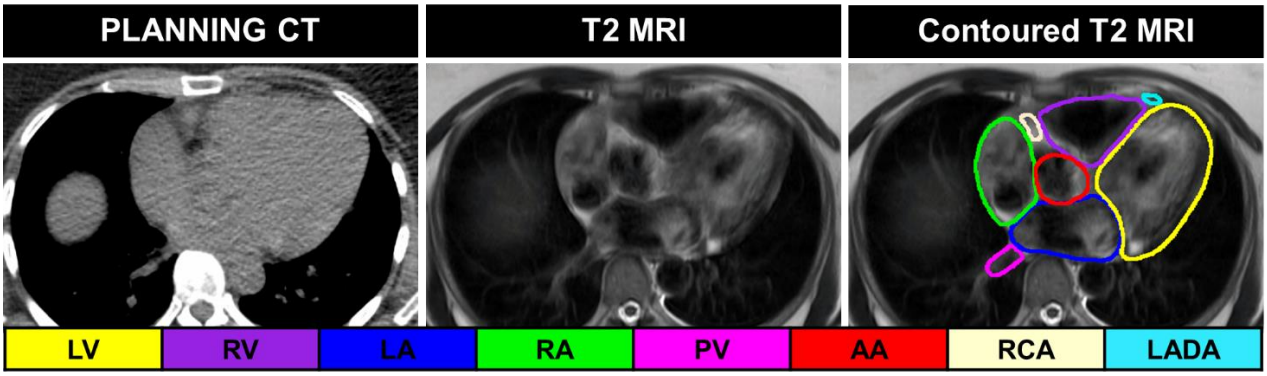


Figure 1: Left: Axial planning CT, Middle: Axial T2 MRI, Right: Contoured axial T2 MRI. Delineated substructures are outlined across the bottom with abbreviations defined in the text.

While overall whole-heart displacement has been measured during conditions of respiration[49], little is known about the motion of most cardiac substructures. When considering motion due to respiration, cardiac substructures have been shown to displace greater than 1.5 cm in the dominant direction of respiration (superior-inferior axis)[33,50]. In some clinical settings, left-breast cancer patients are treated in breath-hold conditions, yielding mostly cardiac motion influences. In clinical RT, cardiac motion is not typically managed as dose volume parameters for the whole heart are not significantly influenced by motion from the cardiac cycle[51]. Cardiac motion management technology is also not currently implemented into clinical linear accelerators. However, individual cardiac substructures may move differently than the entire heart and each other. This was observed in a study by Wang *et al.* who showed that the displacement of the LADA during deep inspiration breath-hold (DIBH) varied substantially when compared to whole heart displacement with maximal extents of the LADA over 7 mm (2.5 ± 1.4 mm average excursion in DIBH)[31]. The coronary arteries and the ventricles have been reported to be the most mobile regions of the heart during the cardiac cycle, displacing 3-8 mm between

end-diastolic and end-systolic phases[52]. Thus, planning organ at risk volume (PRV) margins of 3-4 mm have been suggested for these specific substructures[33,53]. Given the potential for varied sensitivity of cardiac substructures and their independent motion trajectory from the rest of the heart and each other, it is important to give them further consideration as the dose received by specific substructures may differ significantly from the dose to the entire heart.

The Challenges of Cardiac Substructure Automatic Segmentation: Problem Statement

Obtaining paired clinical MRI and CT data for the purposes of cardiac evaluation presents a challenge as their acquisition is not standard of care. The MRIs used in this study are acquired under breath-hold conditions and are not electrocardiogram (ECG) gated, which means that they do not provide temporal data across the cardiac cycle. Thus, respiratory motion is assumed to be negligible during the scan. However, due to extended scan times and heart rate, at least 30 cardiac cycles are captured during the scan. Therefore, the heart and its substructures on the T2-weighted and TrueFISP scans are represented by their average position over the course of the scan. Additionally, cardiac substructure variations in position arise from inter-fraction setup uncertainty which will be quantified in this work.

Introducing cardiac substructures into treatment planning optimization causes an additional challenge: not only is there a desire to lower the radiation dose to cardiac substructures, but also the doses to other OARs from treatments near the heart, such as the spinal cord, lungs, esophagus etc. need to be conserved and within clinical tolerances. Additionally, in adding treatment planning objectives, it needs to be ensured that target coverage is not compromised. This is further complicated by anatomical

variations among patients. Simultaneously meeting current clinical tolerances for the treatment target and the OARs, as well as the ability to spare sensitive cardiac substructures might yield a drastic improvement in patient care.

This work, when taken together, will develop an image processing pipeline to segment cardiac substructures to better quantify potential opportunities for enhanced cardiac sparing in radiation therapy planning (RTP), which will be accomplished via the following specific aims:

Specific Aims

1) Substructure segmentation using a novel atlas method using volumetric T2 MRI rigidly registered to CT-simulation,
2) Further improve substructure segmentation efficiency and accuracy using deep learning,
3) Quantify intra-fraction motion due to respiration, and inter-fraction setup uncertainties,
4) Translate the described technologies to MR-linear accelerator (MR-linac) and treatment planning comparisons.

CHAPTER 2 "IMAGING AND SEGMENTATION TECHNIQUES"

Summary of Computed Tomography and Magnetic Resonance Imaging

Imaging with CT has long been the standard-of-care for RTP due to its exceptional geometric accuracy and spatial fidelity[54]. When a CT image is acquired, each voxel (i.e. three-dimensional pixel) is assigned a numerical value called a CT number based on the reading from the CT detector array. Thus, the CT number is proportional to the attenuation coefficient (μ) at a particular voxel. The attenuation coefficient at each voxel is then compared to that of water (μ_{water}), as shown by the following equation, for conversion to Hounsfield Units (HUs).

$$HU = 1000 * \frac{\mu - \mu_{water}}{\mu_{water}} \quad (1)$$

The HU is a measure of the radiodensity within a CT and is directly related to the measured attenuation. Grayscale values are then assigned as a function of HU for display purposes. Generating a CT image of a phantom with known electron density values can then be used to convert from HU to relative electron density that is then inputted into the treatment planning system for subsequent dose calculation. Thus, an additional benefit of CT imaging of paramount importance is that it allows for a direct conversion from measured attenuation to an object's electron density to enable accurate dose calculation[55].

As the intensity in a CT image represents the x-ray attenuation at a certain point, image intensities are mostly homogeneous among areas of soft-tissue. The limited range of electron densities restricts the overall image contrast, thereby making it challenging to differentiate between regions inside the heart, as shown on the left side of Figure 1. MRI

is often used as an adjunct imaging modality to CT, as it allows for increased soft-tissue contrast (Figure 1, right), specifically in the setting of cardiac imaging[56].

An MRI is obtained by measuring the net magnetization of the hydrogen atoms that exist within tissues[57]. When a specimen is exposed to a strong enough magnetic field, the nuclear spins of the hydrogen atom will be aligned either in the direction of the applied magnetic field, or directly opposed to it[57]. The majority of atoms are in direct alignment with the magnetic field as it is a lower energy state, which causes a net magnetization (i.e. longitudinal magnetization)[57]. In order to measure a signal in MRI, there must be a transverse magnetization present[57]. This transverse component is formed by using a transmit coil to apply an external radiofrequency field (RF) at the same frequency as the Larmor frequency until a peak transverse component is obtained for a given sequence at which point that RF field is turned off[57]. The rotating magnetization leads to a current change in the receive coil allowing for detection of a signal[57]. By the appropriate use of magnet field gradients right after the RF pulse is applied, that signal can be spatially encoded so that a Fourier transform can be applied to the signal to create an image usually in either 2D or 3D[57].

In short, MRI can be manipulated and tuned by adjusting various image sequence parameters[58], and it provides volumetric and multi-planar imaging at a broad range of slice thicknesses[58]. The tissue dependence of the previously mentioned time constants allows for MRI to provide the superior soft-tissue contrast when compared to CT[56] and leads to improved target and OAR visualization[59,60] as discussed in the next section.

Rationale for Magnetic Resonance Imaging Implementation

The segmentation of tumors on CT images is impeded due to low contrast and ambiguous boundaries[61] and can present large uncertainties in RTP for various cancer types[62-66]. For example, in a nasopharyngeal carcinoma study completed by Emami *et al.*, the use of CT imaging alone for tumor delineation failed to include the entire extent of the target[67]. This was made even more apparent by the increase in target volume on the MRI of 74%[67]. In a similar study of over 250 patients, Chung *et al.* found that using MRI allowed for the detection of intracranial tumor infiltration in over 40% of patients, whereas the CT scan had negative findings[68]. In regards to pancreatic cancer, a recent study by Gurney-Champion *et al.* showed that the availability of MRI images for target delineation significantly reduced inter-observer variability in the majority of patient cases when comparing to CT alone[69]. Regarding breast cancer radiotherapy, Hartogh *et al.* found a 4% increase ($P < 0.001$) in inter-observer agreement when using MRI for breast tumor gross target volume delineation over CT[70]. Moreover, they found that for two out of 14 patients the entire tumor was missed (i.e. dense fibroglandular tissue or macrocalcifications segmented instead of lesion) when using CT alone[70]. Lastly, the co-registration of MRI with CT allowed for a decrease in the local standard deviation of the gross target volume from 4.4 to 3.3 mm[71].

MRI is also valuable in the delineation of OARs. When delineating the brachial plexus, Kong *et al.* discussed the necessity of incorporating MRI since the use of CT alone presents challenges[72]. Bainbridge *et al.* summarized various studies on OAR delineation in thoracic radiotherapy[73]. They found that even though the use of a CT-based atlas improved contouring reproducibility in the heart and esophagus, delineation

consistency further improved with the integration of MRI[73]. When conducting OAR delineation in the abdomen, Wachter *et al.* found that defining the prostate apex on CT would have led to 6-13 mm of additional treatment outside of the tumor that was defined on MRI[74]. They recommend that MRI be used for delineation of OARs to avoid unnecessary radiation to the anus and penile structures[74]. Lastly, Khoo *et al.* evaluated OAR segmentation ability in independent observers of the prostate, rectum, bladder, and seminal vesicles and found that MRI provided an improvement to segmentation over CT for each studied structure[75].

One potential complicating factor in delineation accuracy occurs as a result of the susceptibility of both CT and MRI to motion artifacts from patient movement during the imaging session. When a patient breathes freely, the target will displace along the axis of respiratory motion and appear elongated. To mitigate these motion artifacts and increase reproducibility in patient position, breath-hold techniques are commonly incorporated. As breath hold scans require the full cooperation of the patient, scan times characteristically range from 10-25 seconds[76]. For this reason, sequences such as fast gradient echo (e.g. TrueFISP) and turbo-spin echo are frequently utilized in MRI for thoracic and abdominal regions where respiratory motion is considerable[76,77]. Four-dimensional CT (4DCT) or four-dimensional MRI are commonly used techniques for patients that are physically unable to undergo breath hold imaging. In 4DCT for example, a scan is acquired in free breathing over numerous respiratory cycles as a large number of projections is required for each breathing phase to provide an adequate signal to noise ratio[78]. During the scan, the respiratory waveform is also recorded. Images are then binned by phase or by

amplitude to generate multiple three-dimensional (3D) datasets at different stages of breathing[79].

Cardiac Imaging

The heart can be imaged through several techniques, including but not limited to, radionuclide cardiac imaging, echocardiography, cardiac CT, and cardiac MR[80]. Cardiac MR is advantageous because not only does it have superb soft-tissue contrast for structure analysis, it also allows for the analysis of myocardial perfusion and function[80]. In RT however, there is no standard MR imaging sequence for segmenting the substructures of the heart as the concept of applying these structures to treatment planning in radiotherapy is an emerging area of interest.

Whether cardiac segmentation is completed on MR images that are T1-weighted, T2-weighted, or weighted as a combination such as T2/T1 (i.e. TrueFISP), the standard of care for cardiac imaging is to suppress the blood (i.e. force it to be black on the image) during the acquisition[81]. This causes an increase in contrast between the rapidly moving blood and the cardiac muscle for improved visualization. Additionally, increased water (i.e. edema), as well as infarction appear bright on T2-weighted images. T2-weighted cardiac MR can also be used to differentiate acute coronary syndrome from non-acute coronary syndrome, as well as if an infarction occurred recently[82-84]. Lastly, T2-weighted cardiac MR allows for the distinction of the high risk location for both non-reperfused and reperfused myocardial infarction[85-87]. For these reasons, volumetric T2-weighted images are often included in protocols for cardiac MRI. The ViewRay MRIdian MR-linac (ViewRay, Mountain View, CA) utilizes a balanced steady-state free precession (b-SSFP) (i.e. TrueFISP) sequence. In the TrueFISP sequence, the signal is balanced and is

directly related to the T2/T1 value of the tissue. Throughout this body of work, both volumetric T2 images, as well as TrueFISP images were utilized.

The vast majority of cardiac imaging is acquired at a comfortable expiration[88]. Although a patient may be able to hold their breath longer at end-inspiration, diaphragm position between inspirations is much more variable than it is at a repeated expiration[88]. In the absence of respiratory motion, the heart can still displace 3-8 mm between end-systolic (i.e. contraction to pump blood) and end-diastolic (i.e. relaxation after contraction) phases, with the coronary arteries and the ventricles being the most mobile regions[52]. Initially established for coronary artery visualization[89], ECG gating may be used to acquire images of the heart at a certain point in its cardiac cycle, through coupling the correlated ECG pulse with the MRI data[53]. ECG gating is often triggered by the R-wave as it is the strongest signal in the ECG pulse and represents the depolarization of the ventricular myocardium which activates the pumping contraction[90]. ECG gating may occur prospectively where imaging is triggered at predefined points in the cardiac cycle through the R-wave timing[91]. It may also occur retrospectively where images and ECG waves are collected over multiple cardiac cycles and rebuilt into specific intervals at a later time (e.g. 5% intervals between R-waves for a 20-phase ECG gated dataset)[53]. ECG gating allows for cine cardiac motion studies where structural extent is examined. Protocols for diagnostic cardiac imaging may include contrast-enhanced, off axis planes, and cine sequences in order to evaluate cardiac anatomy and function. The latter two series are not often acquired in the axial plane, leading to challenges integrating them into treatment planning due to inaccurate co-localization and subsequent registration.

Additionally, coronary artery segmentations may be improved through the use of high resolution (e.g. 0.78 x 0.78 x 1.6 mm^3) CT coronary angiography (CTCA) that uses contrast-enhancement[92] to drastically increase the visualization of the coronary artery lumen and wall[93]. However, as acquiring CTCA predominantly occurs for evaluation of vascular disease[94] (i.e. not always considered standard of care in RTP), and MRI can provide improved visualization without increased radiation dose or contrast, volumetric MRIs are utilized in this work.

Image Segmentation Techniques

The segmentation of an image can be defined as the splitting an image into two or more meaningful regions. More specifically, it is a process where each pixel in an image is assigned a label, and pixels with similar labels may be linked such that a visual or logical property is realized[95]. These groupings of pixels with the same label are called delineations, or segmentations. Once RT images are acquired, tumors and OARs are delineated, often by a physician, to enable consideration in the treatment planning process. Conducting segmentation manually can present numerous problems including being extremely time consuming and vulnerable to window and level settings[47]. Moreover, manual segmentations can introduce inter- and intra-observer variability[96].

Implementing automatic segmentation methods can drastically decrease the required time it takes to generate clinically usable delineations. Methods for automatic segmentation can be broadly split up into supervised and unsupervised. In unsupervised image segmentation, only the image itself is considered. Thus, unsupervised segmentation techniques utilize image intensity and gradient analysis, which perform well when boundaries in the image are distinctly defined[97]. On the contrary, supervised

segmentation techniques integrate prior knowledge about the image[98]. This prior knowledge is usually in the form of other similarly annotated images that can inform the current segmentation task (i.e. training samples). Based on a survey of the current literature at present, six major categories of image segmentation methodologies have been identified: (1) manual delineations, (2) image thresholding techniques, (3) graph-based approaches, (4) atlas-based approaches, (5) machine learning methods, and (6) deep learning methods. The chief characteristics, limitations, and some examples of each segmentation category are outlined in Table 1.

Segmentation Method	Characteristics/Benefits	Examples	Limitations
Manual	-Visual inspection and interpretation for manual delineation[96] -Simple and straight-forward[96]	-Physician	-Time consuming[96] -Vulnerable to window/level setting[96] -Sensitive to inter and intra-observer variability[96]
Threshold (Binarization)	-Thresholds are selected based on image histogram or manually selected seed pixel[99] -Simple implementation and highly efficient[96]	-Edge detection -Seed growing	-Only two classes are generated[100] -Difficult manual decision[96] -Highly sensitive to heterogeneities[96], motion artifact[96], and image noise[99]
Graph Based	-Image pixels are expressed as nodes on a graph[99] -Can be used to enhance thresholding techniques[97]	-Graph cut -Maximum-flow -Morphological Watersheds	-More complex implementation than thresholding[99] -Can result in cutting small sets of isolated nodes in a graph[101] -Ideally suited for obtaining a rough segmentation of an image's principle regions[97]
Atlas Based	-Form of supervised learning where a predefined library of images informs new segmentations[102]	-Probabilistic atlases -Statistical shape models	-Number of selected atlases will affect result[103] -Long time to generate result[104] -Can depend on deformable image registration framework[104]
Machine Learning	-A nonlinear classifier where a model is trained and tuned[97] -Representations are built from pre-specified filters and are not learned from the image itself[97]	-Markov Random Fields -Conditional Random Fields -Random Forest -Support Vector Machine	-More complex structure than above techniques[97] -High computational cost due to employing iterative schemes[97] -May have redundant features that cause overfitting[97] -Filter bank needs to be designed specifically for task[97]

Deep Learning	-A model is trained and tuned but features are learned and guided by the training data and are not pre-specified[97]	-Artificial neural networks -Convolutional neural networks -Recurrent neural networks	-Most complex structure[97] -Possibility of overfitting if not enough variation in training data[105] -High computational cost[97] -Long training times[105]

Table 1: Summary of characteristics, limitations and some common examples of various broad segmentation techniques.

Image thresholding techniques, such as edge detection and seed growing methods, are easy to implement and are highly efficient. Thresholding an image involves selecting one or more points on an image's histogram in order to bifurcate the image into distinct regions. These methods can also involve the user selecting an initial seed pixel and a value, for example τ. All pixels adjoining to the initial seed pixel with intensities $\pm \tau$ are included in the segmentation. This process is repeated with all included pixels until a border is generated (i.e. pixel values greater than τ or less than $-\tau$). These methods present a difficult decision to the user and can be limited by heterogeneities, motion artifacts, or noise in the image.

Unlike unsupervised methods, such as image thresholding and graph-based techniques, supervised methods employ image delineations that have already been generated on similar datasets. Several atlas-based methods have been used in various applications of automatic segmentation and have been described in the literature[106-111]. The main differentiation between these atlas-based studies and others is the registration algorithm for mapping image coordinates onto the atlas[102]. However, as inter-subject registrations may yield considerable differences in shape, a trait that all aforementioned atlas-based methods have in common is that the transformations used are non-rigid[102]. The process of automatic segmentation with an atlas-based method is shown in Figure

2.

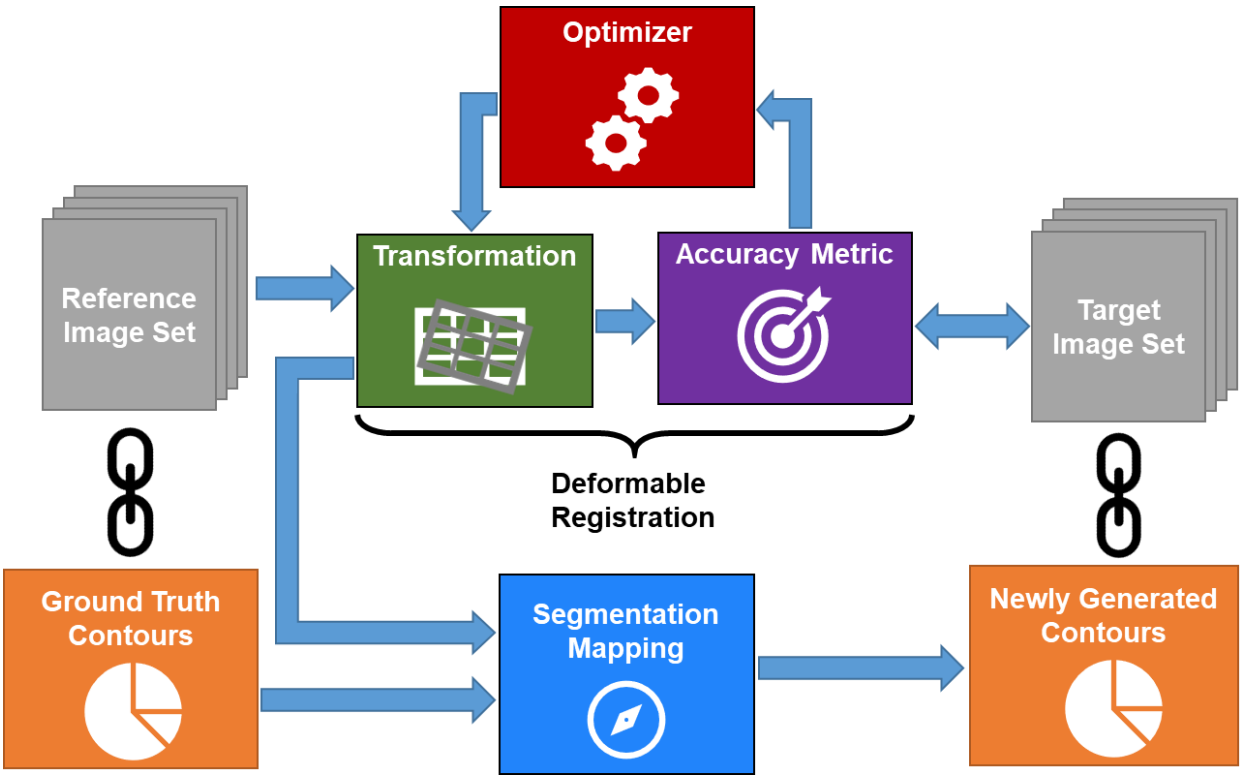


Figure 2: Process of generating new segmentations on a target image set via atlas-based automatic segmentation.

The reference image set used for atlas-based automatic segmentation can be a single image or a library of multifarious pre-segmented datasets similar to the image set to be segmented, each with an associated set of ground truth contours[102]. A transformation using deformable image registration is then applied to the reference image data set and is assessed via a metric of accuracy (e.g. mutual information-based algorithms) to measure agreement with the target image data set. An iterative optimization process can then be used to refine the deformable image registration to improve the accuracy between the reference and target data set. The final transformation matrix (i.e. deformation vector field) can then be applied to the original ground truth contours to generate new segmentations on the target image data set (segmentation mapping in Figure 2). To

summarize, consider an atlas *A*, a reference image set *R,* and a target image set *T*, where each point in *R* has a corresponding equivalent in *T*. This correspondence is represented by the coordinate transformation *F,* that maps *R* onto *T*[102]. Thus, for a location *(i,j)* in *T*, the corresponding location in domain *R* is *F(i,j)*. Through the association of *R* with *A*, the label at any location in *T* can be found using:[102]

$$(i,j) \mapsto A(F(i,j)) \tag{2}$$

The transformation *F* represents the deformable image registration (DIR) that characterizes an n-dimensional vector *v*[97]. Consequently, atlas based segmentation methods can be time consuming (converging on an optimal *v*) and highly dependent on DIR framework[104].

Artificial intelligence is a technique that enables a machine to mimic the behavior of a human. Machine learning is a subset of artificial intelligence, as shown by Figure 3 which enables machines to achieve artificial intelligence through algorithms and statistical techniques trained with data. The training process informs the decisions made by the machine learning framework which allows for improvement as experience is gained[112].

Supervised machine learning methods for the automatic segmentation of images involves training and tuning a predictive model. Machine learning methods are able to improve upon the previously mentioned segmentation techniques by incorporating past labeled data. Machine learning utilizes numerous statistical tools to explore and analyze this previously labeled data with image representations being built from pre-specified filters tuned to a specific segmentation task. Although machine learning techniques are more efficient with image samples and have a less complicated structure, they are often not as accurate when compared to deep learning techniques for automatic image

segmentation[97].

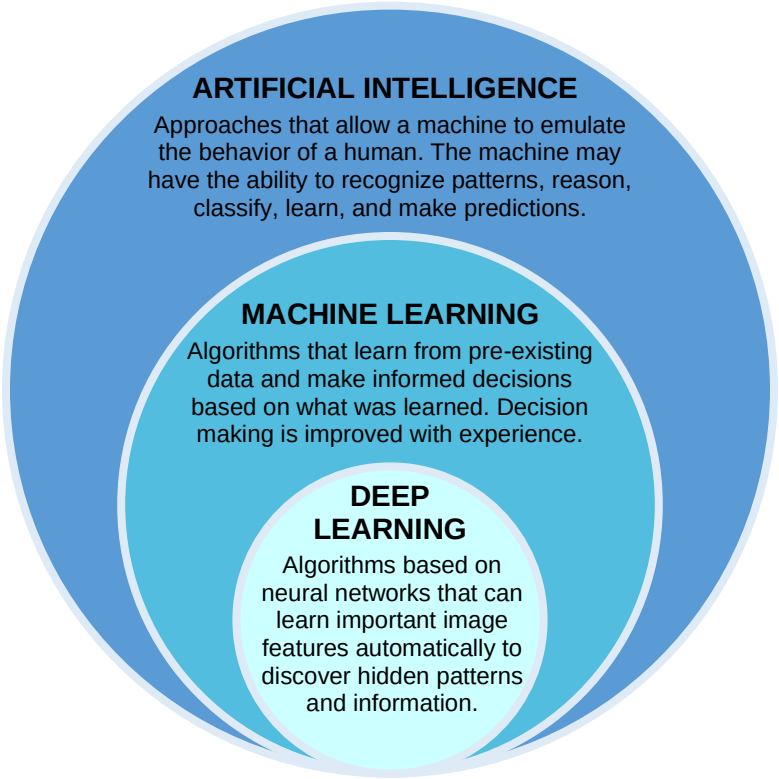


Figure 3: Relationship between and descriptions of artificial intelligence, machine learning, and deep learning.

Conditional random fields (CRF) are an example of a machine learning technique described in Table 1. The purpose of CRF are to create a relationship between the observed evidence (i.e. the image) and the hidden "true states" (i.e. probability map) in order to take the context of the image into account. The probability map (*x* pixels in Figure 4) has encoded neighborhood information that can be based on several image features such as intensity, color, texture differences, edges, etc. The CRF seek to minimize the energy function $E(x,y)$ (Equation 5) that is built of energy potentials $\varphi(x,y)$ (i.e. unary potential, Equation 3) and $\psi(x,y)$ (i.e. pairwise potential, Equation 4). The unary potential encodes the local information about a given pixel and how likely it is to belong to a certain

class. The pairwise potential encodes the neighborhood information that considers how different a pixel's label is from that of its neighbor. Thus, when the energy function *E* is minimized, the pairwise term encourages smooth annotations and the unary term causes annotations to become more localized (i.e. reducing holes in segmentations or far remote island annotations).

$$Unary = \sum_i \varphi(x_i, y_i) \quad (3)$$

$$Pairwise = \sum_{i,j} \psi(x_i, y_i) \quad (4)$$

$$E(x, y) = Unary + Pairwise \quad (5)$$

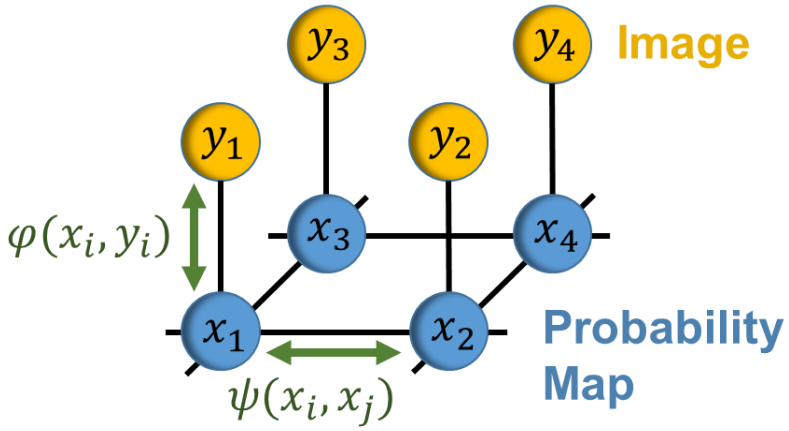


Figure 4: Relationship between pixels in the image to be segmented (y_i) and the pixels in the probability map (x_i)[113]

Deep learning is a subset of machine learning, as shown by Figure 3, which was originally designed to mimic the learning style of the human brain using neurons. Unlike machine learning where the "useful" features for the segmentation process must be decided by the user, with deep learning, the "useful" features are decided by the network without human intervention. If an image is considered as the input to a convolutional neural network (CNN), each pixel in the input image would then be fed to a neuron in the input layer of the network, represented by purple circles in Figure 5. Each channel (blue

lines in Figure 5) has a weight assigned to it, or w_i, and all neurons have a unique number associated with them called bias, or B_i. The biases and weights of a network are learnable parameters that can either be initialized at 0, random values, or predetermined values[114]. The bias is added to the weighted sum of all inputs reaching a particular neuron and is shown in the following equation.

$$H = B_1 + (n_1 w_1 + n_2 w_2 + \cdots + n_i w_i) \quad (6)$$

The value *H* is then applied to an activation function, with an example activation function (σ) shown on the right of Figure 5. Thus, the activation function then becomes,

$$Activation\ function = \sigma(H) \quad (7)$$

The result of the activation function determines if the neuron gets activated. For example, when using the sigmoid activation function shown in Figure 5, activation values greater than or equal to 0.5 activate the neuron. Every activated neuron passes on information to the following layers which continues in each layer until the second to last layer. The one neuron activated in the output layer corresponds to the decision of the network (i.e. the label of the pixel). This process occurs for all pixels in the training data and represents a forward pass of the data.

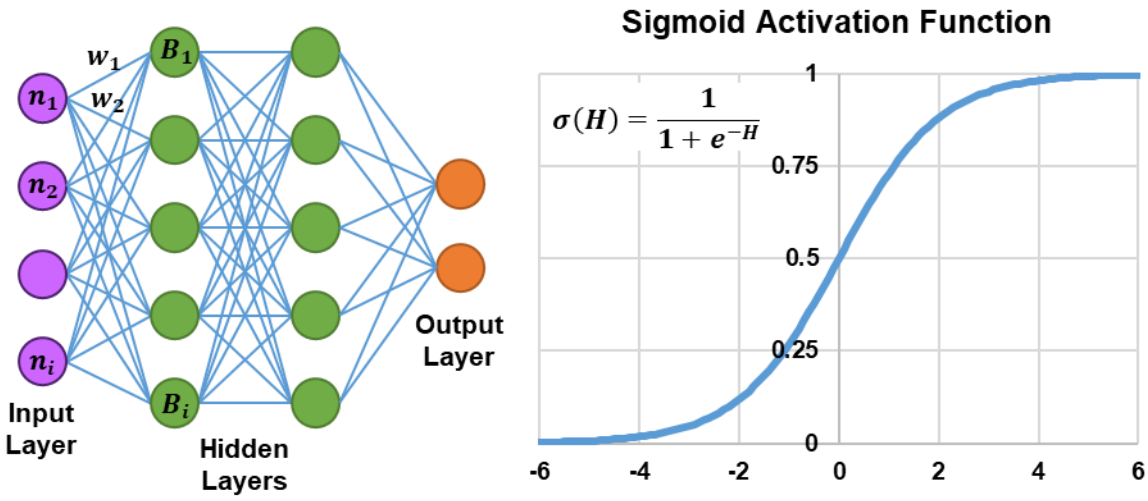


Figure 5: Left: An example of a neural network structure. Channels are represented as blue lines and neurons are represented as circles with the input layers shown in purple, the hidden layers in green, and the output layer in orange. Right: Equation and plot for the sigmoid activation function.

After completing a forward pass of the training data, a backward pass then occurs through a process called backpropagation. Backpropagation is used to adjust the weights and biases to produce a well-trained network. Due to the training data having ground truth, weights and biases may be adjusted based on the error at each neuron (i.e. the difference between the current predicted value and the known truth value). The cost of the network is the cumulative error across the entire training dataset. In total, the goal of the backpropagation process is to compute the partial derivatives of each weight and bias in the network (i.e. $\partial C/\partial w_n$ and $\partial C/\partial B_n$), so that the negative gradient of the cost function may be applied to minimize the overall cost. The neural network is trained over copious forward and backward passes of the training samples so that there is convergence in the adjustment of the weights and biases. The typical number of utilized epochs ranges from 50 to 1000[115] and training times can span from 20 minutes to 100 hours depending on the size of the dataset and the employed hardware[116]. One forward pass and one

backward pass (i.e., backpropagation process) of all the training samples is called an epoch[117].

Recently, CNNs using deep learning have been applied to the complicated task of segmenting cardiac images with promising results[118]. This is achieved through the ability of the deep learning neural network to automatically learn complex features from image data rapidly. Coupling the ability for rapid feature extraction with advanced computer hardware, such as graphical processing units (GPUs), has allowed for automatic segmentation time to be radically decreased. A deep neural network (DNN) is able to learn a mapping function between an image and a corresponding feature map (i.e. segmented ground-truth) by incorporating multiple hidden layers between the input and output layer (Figure 5). Deep learning applications may be limited by the size of datasets made available in the field of radiation oncology. However, these limitations may be circumvented through the utilization of image augmentation. The U-Net[119] is a DNN architecture that has shown great promise for generating accurate and rapid delineations for applications in RT[120].

The U-Net architecture was inspired by the original fully convolutional network from Long *et al.*[121] and was implemented initially by Ronneberger *et al.*[119] in 2015. The U-Net gets its name from its 'U-shaped' architecture, as shown in Figure 6. It was originally implemented to segment biomedical image data in an end-to-end setting utilizing 30 total annotated image sets. Starting from the top left of Figure 6, the input image (size 512 x 512) undergoes two convolutions to generate a multi-channel feature map (purple boxes in Figure 6) with 64 initial feature channels. Next, a non-linear activation function is applied, as discussed in Figure 5 above, followed by a maximum pooling operation. The

maximum pooling operation reduces the *x* and *y* size by a factor of two for the feature map, which acts on each channel separately, and then the number of feature channels is doubled. This process of successive convolutions and maximum pooling operations (i.e. contraction pathway) gradually increases the "what" while gradually decreasing the "where." In a standard classification neural network, such as the one shown in Figure 5, the network would stop at this point and all feature maps would be combined into a single output vector. However, the U-Net has an additional expansive path that creates a high-resolution segmentation map. The expansion pathway (Figure 6, right) replaces the maximum pooling operations with up-convolutions to increase the resolution of the feature maps. Then, in order to localize, features from the contraction pathway are joined with the up-convolved feature map through concatenation[119]. The U-Net is an end-to-end solution with remarkable potential to segment medical images, even when the amount of training data is scarce[122].

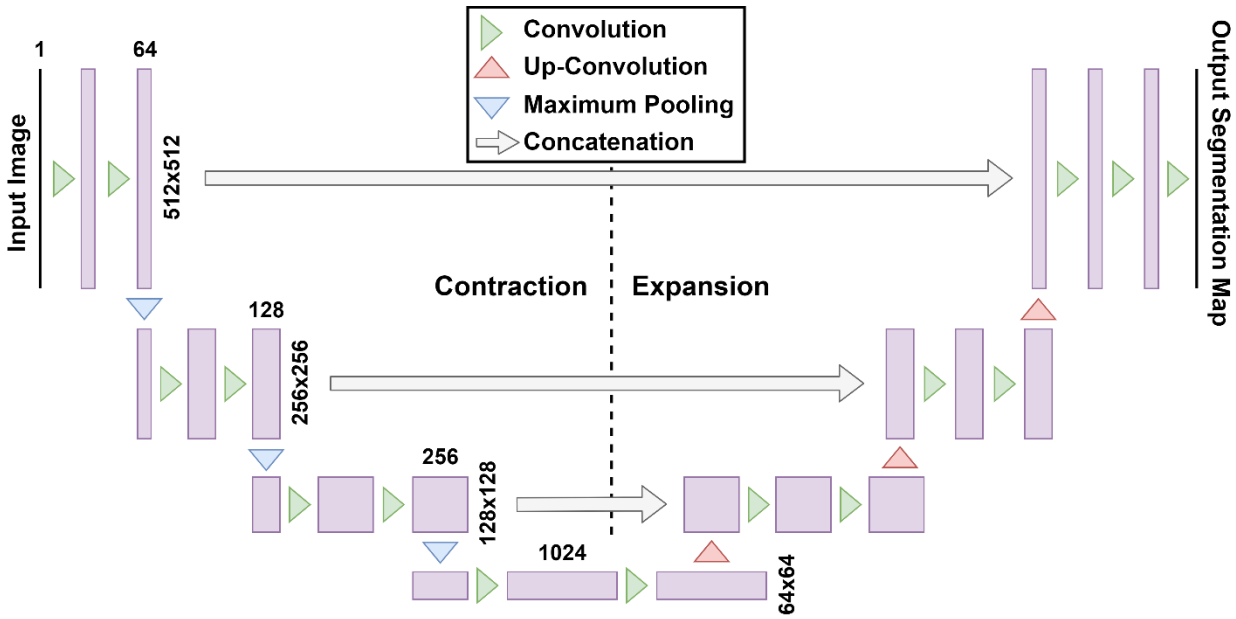


Figure 6: Architecture for original U-Net by Ronneberger *et al.*[119] with the contraction path shown on the left and the expansion path shown on the right. The original input

image has a size of 512 x 512. Feature maps are represented by purple rectangles with the number of feature maps on top of the rectangle.

Overall, there are several commercial algorithms for automatic segmentation already available, but their core mechanics fall within atlas-based measures instead of machine learning strategies[123]. These atlas-based measures are highly reliant on the amount of contrast in the target organ for segmentation[124]. On the contrary, recent deep learning techniques[125,126] are well poised for the task of accurate automatic segmentation with less reliance on organ contrast[127,128] as the algorithm is designed to acquire higher order features from raw data[125]. Overall, deep learning techniques offer the potential for rapid and accurate cardiac substructure segmentations over other currently available segmentation methods.

CHAPTER 3 "CARDIAC SUBSTRUCTURE SEGMENTATION AND DOSIMETRY USING A NOVEL HYBRID MAGNETIC RESONANCE COMPUTED TOMOGRAPHY CARDIAC ATLAS"

Introduction

Increased risk of cardiotoxicity, including coronary artery disease and cardiomyopathy, has been linked to RT for many cancer sites in close proximity to the heart[5-7]. Moreover, major coronary events have been linked to radiation dose to the heart with a ~7%/Gy increase in rate of ischemic heart disease[16]. Radiation dose from breast cancer treatments specifically have also been correlated to myocardial infarction, congestive heart failure, and cardiovascular death more than 10 years after RT[129]. Therefore, minimizing cardiac dose due to RT treatments is critical and has been given recent attention[130,131].

Currently, OAR dose limits for breast cancer RTP account for the entire heart volume. For example, according to a cooperative group trial for left-sided breast cancers, no more than 5% of the whole heart may exceed 20 Gy[19]. However, heart substructures are not routinely included in RTP as they are not visible on standard CT-SIM datasets and dose limits are not currently well established. Furthermore, dose to these cardiac substructures may have prognostic inferences. In left breast cancer patients, the LADA is often exposed to the highest treatment-related radiation dose[16] and increased dose has been linked to increased risk of late radiation induced cardiac morbidity[30]. Additionally, measurements of radiation dose to cardiac substructures, like the LV, may be useful in the prediction of future acute coronary events[28]. By localizing these sensitive substructures within the heart, we can then estimate the dose to these regions. This may further our understanding of their potential roles in radiation-related cardiotoxicity.

The current standard of care for radiation treatment planning is based on CT-SIM to enable electron density mapping for dose calculation. However, cardiac substructures are not easily discernible on standard CT-SIM datasets. Some single modality atlas methods have employed contrast-enhanced CT[45,132,133]. For coronary arteries such as the LADA, contrast-enhanced CT may aid in localization although the majority of the structure is not discernible[134]. At present, contrast-enhanced CT (either diagnostic or for CT-SIM) is not widely available for breast cancer patients, nor are they included in the current National Comprehensive Cancer Network breast cancer recommendations[135] or cooperative clinical trial group guidelines for delineation [136-138]. Thus, developing atlas solutions that can be applied to widespread, clinically available data, such as standard CT-SIMs, is advantageous.

MRI also improves the visibility of cardiac substructures[43,44] as illustrated in Figure 1. Thus, numerous automatic segmentation methods have been established utilizing MRI[139]. However, MRIs are not frequently acquired and integrated into routine RTP. Recently, a multi-scale patch method was used to generate a multi-modality atlas (e.g. cardiac MRI and contrast-enhanced CT) for automatic segmentation of 7 substructures[47]. The purpose of the current study is to develop and validate a novel hybrid MR/CT segmentation atlas with the overall goal of segmenting 12 sensitive cardiac substructures and the whole heart on standard, non-contrast, treatment planning CTs. The completion of this work offers potential for widespread implementation and may provide important information for dosimetric assessment for OAR sparing.

Methods

Image Acquisition

Thirty-one patients who underwent RT for left-sided whole-breast cancer were enrolled on an Institutional Review Board approved study to acquire cardiac MRI scans. T2-weighted acquisitions were performed on a 3T Philips Ingenia (Philips Medical Systems, Cleveland, OH) with images acquired at end-expiration (EE). Patients were positioned supine and imaged in a multi-coil configuration with a 32-channel dStream Torso coil (Philips Medical Systems, Cleveland, OH) and a 20-channel integrated posterior coil. Imaging parameters included an 8 mm slice thickness, in-plane resolution 0.7 x 0.7 mm^2, and an echo time of 81 ms. This two-dimensional acquisition involved a single breath hold at EE with an average total acquisition time across all patients of 22.1 ± 4.4 seconds (range: 15.1-31.0 seconds).

Non-contrast CT-SIM was performed on a Brilliance Big Bore CT Simulator (Philips Medical Systems, Cleveland, OH) with a 3 mm slice thickness. Eight patients underwent 4DCT while the other 23 patients underwent a CT-SIM under free-breathing conditions (FBCT) based on institutional practices. FBCT and 4DCT images were acquired with an in-plane resolution of 1.1x1.1 mm^2 – 1.4x1.4 mm^2, 120-140 kVp, and 275-344 mAs. All patients were imaged in the supine position and immobilized on a Posiboard (Civco, The Netherlands) with their arms above their head.

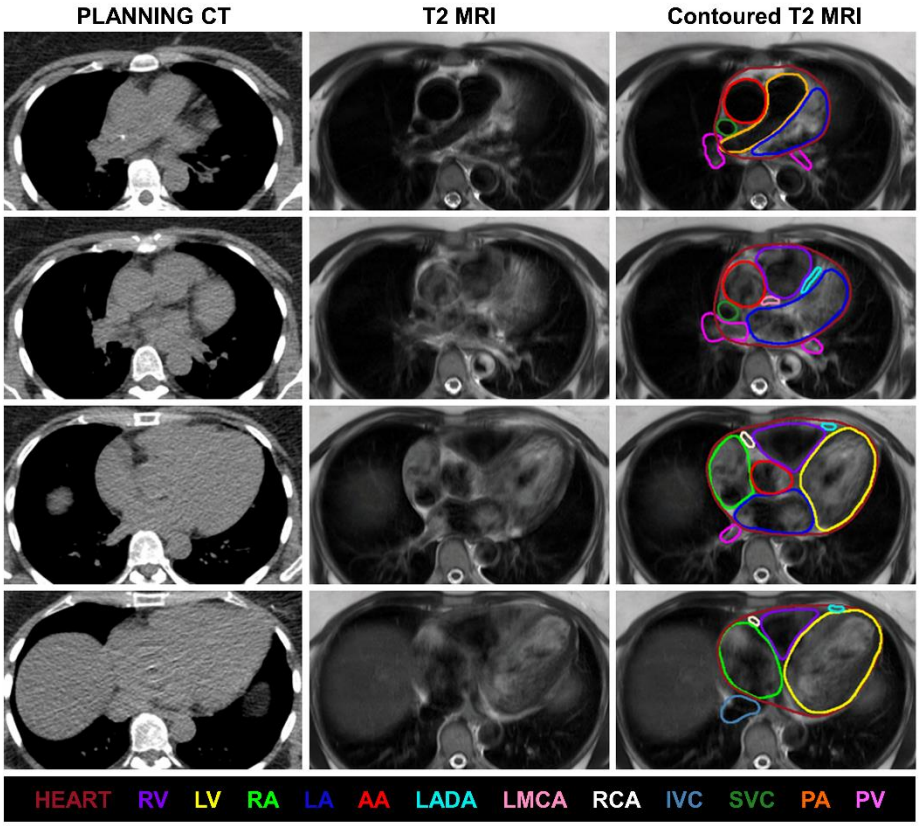
Image Registration

A local, cardiac-confined, rigid registration was performed between the non-contrast CT and axial T2-weighted MR in MIM (version 6.7.12, MIM Software Inc., Cleveland, OH). Rigid registrations were conducted using normalized mutual information

as the similarity metric, which has been shown to robustly align multi-modality images[140]. Visual inspection of the cardiac-confined rigid registration was performed by a radiation oncologist before completing manual segmentations. To compensate for respiratory motion, the 50% (EE) phase of the 4DCT was rigidly registered to the MRI for 8 subjects. Despite 23 patients being imaged with FBCT at arbitrary phases of the breathing cycle, visual inspection of the locally confined heart rigid registration by a physicist and radiation oncologist revealed that the registration quality was adequate for delineation purposes.

Contour Delineation

A radiation oncologist delineated 12 cardiac substructures and the whole heart, as outlined in Figure 7, with substructure selection based on CT and cardiac MRI auto-segmentation atlases [132,139,141]. Substructures were also selected based on their roles in major cardiac function and proximity to the radiation field.



Cardiac Substructure List
Heart
Left/Right Ventricle (LV/RV)
Left/Right Atrium (LA/RA)
Pulmonary Vein (PV)
Pulmonary Artery (PA)
Ascending Aorta (AA)
Superior Vena Cava (SVC)
Inferior Vena Cava (IVC)
Left Anterior Descending Artery (LADA)
Left Main Coronary Artery (LMCA)
Right Coronary Artery (RCA)

Figure 7: Left: Axial planning CT, axial T2-weighted MRI, and contoured axial T2-weighted MRI, shown at 4 different axial locations. Right: List of cardiac substructures assessed in this study.

A radiation oncologist followed a published cardiac atlas consensus contouring guideline[141] to delineate the substructures on the CT using a MR/CT rigid registration with a fixed window/level on CT (50/500 for large structures, 50/150 for cardiac vessels) and giving preference to MR anatomical information. However, as the epicardial border of the heart is visible on CT, the CT was used to generate the whole heart contour. Contours were verified by a radiologist with a cardiac subspecialty and 30 years of clinical experience.

Atlas Generation

Non-contrast CT ground truth delineations derived from MR/CT hybrid information for a subset of 20 patients were inputted into an intensity-based deformable registration atlas in MIM. To perform a sample size estimation, an initial test cohort of 5 subjects was evaluated. To achieve 80% power (medium effective size on a repeated measure ANOVA, alpha error of 0.05 to compare the difference between 8 means), 10 patient cases were required. However, because the Dice similarity coefficient (DSC) was not normally distributed and a non-parametric Friedman test was to be used for analysis, the sample size (n=10) was then divided by a correction factor of 0.955 (i.e. the asymptotic relative efficiency)[142], resulting in a required sample size of n=11 for the testing patient cohort. Thus, the deformable registration atlas was applied to 11 validation patients (i.e. test subjects).

To generate the hybrid segmentation atlas, a reference structure set was first generated from a predetermined patient with average anatomy to act as a template[143]. The template patient was selected based on a moderate habitus, minimal motion artifact, and standard heart geometry and anatomical position. A local, cardiac-confined, rigid registration was then performed between the template patient and the 20 subsequent patients, including the template patient. The process of applying the atlas to a test subject is outlined in Figure 8. One of the 11 test subjects was then selected and a mutual information-based algorithm[144] was used to locate the atlas subject(s) that were deemed the best matches to the test subject. A free-form DIR was then completed between each selected atlas subject and the test subject. The commercially available free-form intensity-based DIR algorithm has limitless degrees of freedom and utilizes adequate regularity

(i.e. penalty term weight) to ensure smooth deformation[145] and has been previously validated in CT/CT registrations yielding high segmentation accuracy[146]. Finally, the generated deformation vector field was used to propagate the ground truth segmentations from the best match to the test subject's CT.

To optimize atlas performance, three atlas approaches were evaluated (1) single-atlas method, and two multi-atlas segmentation approaches (2) majority vote (MV) and (3) Simultaneous Truth and Performance Level Estimation (STAPLE). The single-atlas method (Figure 8, top row) deforms contours from the single best matching atlas subject to the test subject whereas multi-atlas approaches (Figure 8, bottom row) use various best matching atlas subjects. In MV, after multiple contours for the same substructure are deformably propagated to the test subject, the most frequent contour at each voxel is established as the true segmentation[104]. The STAPLE method uses a probability map to create an estimate of the true segmentation from a collection of contours by using an expectation-maximization algorithm[147]. The resultant segmentation is then formed by optimally combining the existing contours through assigning weights based on sensitivity and specificity[147].

To further optimize segmentation, the number of multi-atlas matches was iterated (3, 5, 10, and 15) for MV and STAPLE methods for the 11 validation datasets. Once final contours were obtained, post-processing including contour smoothing and filling was performed[104]. Image processing time was logged in MIM and tabulated for each approach for a representative validation patient. Image processing was conducted on a 64-bit Microsoft Windows PC with a quad-core Intel® Xeon® CPU-E5-1630 v4 at 3.70GHz and 16GB of memory.

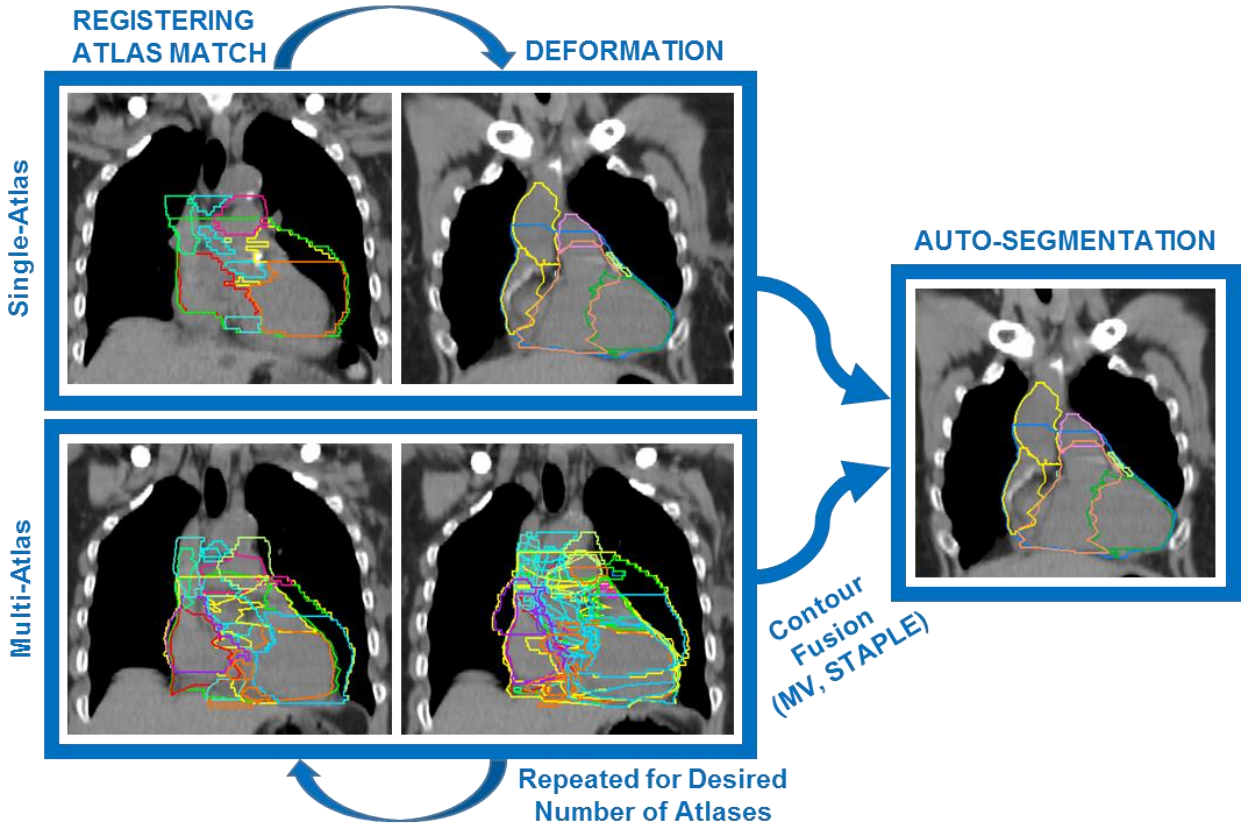


Figure 8: Process of generating an automatic segmentation with an atlas-based method using both a single-atlas (top row) and a multi-atlas (bottom row). Abbreviations: Majority vote (MV), Simultaneous Truth and Performance Level Estimation (STAPLE).

Atlas Validation

Atlas performance was assessed via DSC[45,132,139] mean distance to agreement (MDA)[148], and centroid displacement between propagated and ground truth delineations for the 11 test cases. The DSC is used to measure the spatial overlap between two structures (Equation 8) and is a value from 0, representing no overlap, to 1, representing perfect agreement.

$$\mathrm{DSC} = \frac{2|\mathrm{M} \cap \mathrm{N}|}{|\mathrm{M}| + |\mathrm{N}|} \tag{8}$$

Where M and N are the volumes of the manually delineated and propagated contour, respectively. MDA is used as a geometrical measure to assess the agreement

between two contours by averaging the per voxel shortest distance from each point on the test contour to the reference contour, with an increased agreement yielding a lower MDA[148].

To evaluate clinical acceptance of the auto-segmented contours, qualitative consensus scoring was completed for 12 cardiac substructures and whole heart contours on a subset of 5 of the 11 validation cases by 2 radiation oncologists and a radiologist with a cardiac subspecialty. Scores were assigned on a 5-point scale[149]: (1) clinically unacceptable, (2) major modifications required, (3) moderate modifications required, (4) minor modifications required, (5) clinically acceptable.

Dosimetric Assessment

A dosimetric assessment of left breast cancer patients was conducted to illustrate a potential clinical application of the validated atlas in the test cohort. For the 11 test subjects, the clinically approved and delivered treatment primary whole breast (tangential fields, 6-18 MV, 42.7-45.0 Gy) and boost (3D planned with 6-15 MV photons or 12 MeV electrons to 10.0-16.2 Gy) plans were exported from the Eclipse Treatment Planning System (Version 11.0, Varian Medical Systems, Palo Alto, CA) and into MIM for direct dose summation. Dosimetric evaluation of the cardiac substructures included measurements of the minimum, mean, and maximum dose to each substructure. The MHD, LV-V5, and left anterior descending artery mean and maximum doses ($LADA_{mean}$, $LADA_{max}$) are highlighted as they have been shown to be predictive of acute cardiac events[16,28] and are important indicators for ischemic heart disease[30,150].

Statistical Analysis

Data are presented as mean ± standard deviation (SD). Statistical analysis of DSC and MDA between atlas methods was performed using the Friedman test with a Wilcoxon signed ranks test for post hoc pairwise comparisons and were Bonferroni corrected. Statistical analyses between ground truth and auto-segmented volumes and doses were performed using a two-tailed Wilcoxon signed ranks tests with $P < 0.05$ considered statistically significant. All analyses were performed using SPSS version 25.0 (SPSS, Chicago, IL, USA).

Results

Contour Generation

The average time for the manual delineation of the 12 cardiac substructures and the whole heart was ~3 hours per patient. For a representative test subject, the atlas auto-segmentations took between 1-10 minutes depending on the selected amount of atlas matches. When applying the STAPLE method with 10 atlas matches to this same representative patient, the radiation oncologist required ~30 minutes of additional time to edit segmentations for all 12 substructures and the whole heart for clinical implementation.

Atlas Performance Evaluation

Figure 9 outlines atlas performance for single-atlas, MV, and STAPLE methods with 3, 5, 10, and 15 atlas matches for MV and STAPLE. Figure 9 (left) shows that median DSC values across all structures were between 0.71 and 0.80 for the single-atlas method and STAPLE method with 10 atlas matches (ST10), respectively.

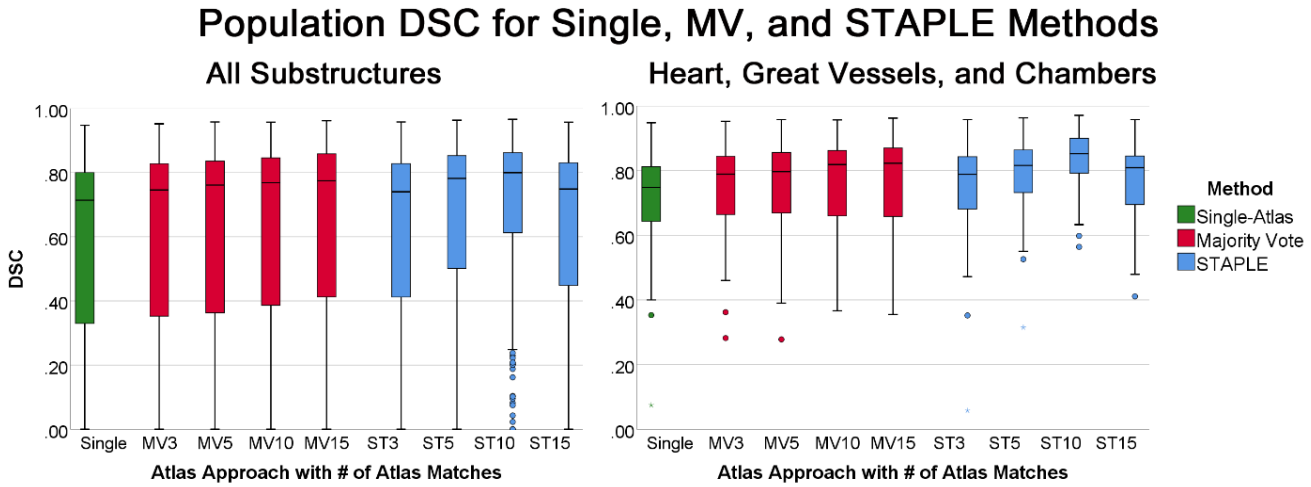


Figure 9: Validation patient Dice Similarity Coefficient (DSC) results over all substructures (Left) and all high performing substructures (i.e. heart, cardiac chambers, and great vessels) (Right). Boxplots and line indicate the interquartile range and median, respectively. Whiskers indicate the minimum and maximum, with data points > 1.5 times the interquartile range and > 3 times the interquartile range marked by circles and stars, respectively.

In general, atlas approaches performed similarly, yielding mean DSCs > 0.75 for 7/13 cardiac structures (heart, chambers, AA, and PA) over the 11 validation patients. Table 3 outlines DSC results for the single atlas method and select high performing multi-atlas methods. Across the 11 test subjects, all coronary artery segmentations had DSC values < 0.42. The right side of Figure 9 summarizes the 11 validation patient mean DSC results after exclusion of the coronary arteries, where median DSC values range from 0.75 to 0.85 for the single-atlas and ST10 method, respectively. In comparing atlas methods, ST10 generated the highest mean DSC and lowest MDA values for all high performing substructures (i.e. heart, chambers, and great vessels). The post hoc pairwise comparisons, shown in Table 2, revealed that ST10 outperformed the single-atlas for all substructures except the coronary arteries for both MDA and DSC ($P < 0.05$).

	Method 1/2	Heart	AA	LA	LV	RA	RV	PA	PV	SVC	IVC
DSC	Single/MV10	0.95	0.78	0.01*	0.63	0.23	0.01*	<0.01*	1.00	1.00	1.00
	Single/ST10	<0.01*	<0.01*	<0.01*	<0.01*	<0.01*	<0.01*	<0.01*	<0.01*	<0.01*	<0.01*
	ST3/ST10	0.03*	0.01*	<0.01*	0.01*	0.01*	<0.01*	<0.01*	<0.01*	0.02*	0.11
	ST5/ST10	1.00	1.00	1.00	1.00	1.00	0.02*	0.05*	1.00	0.46	1.00
	MV10/ST10	0.26	0.10	1.00	0.01*	0.26	1.00	1.00	<0.01*	<0.01*	<0.01*
MDA	Single/MV10	1.00	1.00	1.00	<0.01*	1.00	0.46	0.95	<0.01*	0.09	1.00
	Single/ST10	<0.01*	<0.01*	<0.01*	<0.01*	<0.01*	<0.01*	<0.01*	<0.01*	0.01*	<0.01*
	ST3/ST10	0.04*	0.02*	<0.01*	0.02*	0.04*	<0.01*	<0.01*	<0.01*	<0.01*	0.29
	ST5/ST10	1.00	1.00	1.00	0.70	1.00	<0.01*	0.14	1.00	1.00	1.00
	MV10/ST10	<0.01*	0.14	<0.01*	1.00	0.02*	0.18	0.01*	1.00	1.00	0.14

Table 2: Statistical pairwise comparisons of mean distance to agreement (MDA) and Dice similarity coefficient (DSC) between select atlas methods for each high performing substructure. Statistical differences are marked by a star and represent method 2 having a statistically higher median than method 1 ($P < 0.05$). Atlas methods are defined as single-atlas (Single), Majority Vote (MV), and Simultaneous Truth and Performance Level Estimation (ST) followed by the number of atlas matches.

However, the single atlas method's DSCs and MDAs performed similarly to STAPLE and Majority Vote when fewer than 5 atlas matches were used ($P > 0.05$, results not shown). Regarding MDA, ST10 outperformed ST3 and ST15 for > 8 high performing substructures. Additionally, ST10 outperformed ST5 for the RV and PA DSCs at the expense of ~5 minutes processing time. Thus, all further analyses were conducted using ST10.

Per Structure Mean ± SD	DSC from Atlas Application			Consensus Score
	Single	MV10	ST10	ST10
Heart	0.92 ± 0.03	0.94 ± 0.01	0.95 ± 0.01	4.2 ± 0.4
Left Ventricle	0.83 ± 0.04	0.88 ± 0.01	0.91 ± 0.01	4.6 ± 0.5
Right Atrium	0.80 ± 0.05	0.84 ± 0.04	0.87 ± 0.03	4.2 ± 0.4
Left Atrium	0.77 ± 0.03	0.84 ± 0.03	0.86 ± 0.03	3.8 ± 0.4
Pulmonary Artery	0.74 ± 0.07	0.81 ± 0.03	0.84 ± 0.03	4.0 ± 0.0
Ascending Aorta	0.73 ± 0.09	0.79 ± 0.07	0.84 ± 0.03	4.4 ± 0.5
Right Ventricle	0.71 ± 0.06	0.80 ± 0.05	0.83 ± 0.03	4.2 ± 0.4
Superior VC	0.67 ± 0.09	0.66 ± 0.08	0.80 ± 0.04	4.0 ± 0.7
Inferior VC	0.46 ± 0.23	0.55 ± 0.11	0.70 ± 0.07	4.0 ± 0.7
Pulmonary Vein	0.47 ± 0.13	0.50 ± 0.06	0.64 ± 0.06	3.2 ± 0.4
Average over Heart, Chambers, and Great Vessels	**0.71 ± 0.08**	**0.76 ± 0.05**	**0.82 ± 0.03**	**4.1 ± 0.5**
LAD Artery	0.15 ± 0.14	0.04 ± 0.04	0.27 ± 0.09	1.8 ± 0.8
RT Cor. Art.	0.14 ± 0.10	0.03 ± 0.05	0.22 ± 0.10	2.4 ± 0.9
LT Main Cor. Art.	0.05 ± 0.08	0.00 ± 0.00	0.12 ± 0.12	1.4 ± 0.5

Table 3: Mean and standard deviation (SD) results for select atlas methods showing the Dice Similarity Coefficient (DSC) per substructure and across all high performing substructures for the validation population (heart, chambers, and great vessels). Consensus scores from physician grading of the ST10 method are also shown. Abbreviations defined in text.

Segmentation Results for ST10

Figure 10 summarizes the mean MDA and DSC results between manually delineated ground truth contours and ST10 atlas generated contours over the 11 test cases. Over all 12 substructures and the whole heart, 9 had an MDA < 2.1 mm and a mean DSC > 0.70, suggesting excellent atlas performance. The coronary arteries performed the worst (mean DSC < 0.3 and MDA between 3.1-4.2 mm).

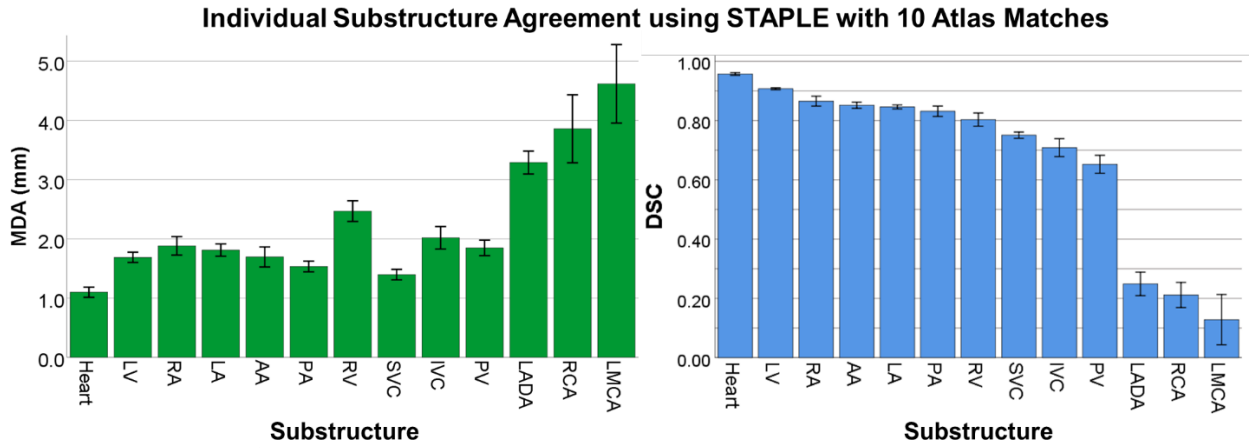


Figure 10: Mean distance to agreement (MDA) (Left) and Dice similarity coefficient (DSC) (Right) between ground truth and ST10 contours for all delineated substructures (n = 11). Error bars represent the standard error of the mean.

Additionally, across the 11 test cases, over half of all contours had centroid displacements < 3.0 mm, with largest shifts in the coronary arteries. The greatest centroid displacements occurred in the superoinferior direction (predominantly superior). Three out of 12 substructures (left main coronary artery, pulmonary vein, and right coronary artery) had statistically significant differences in volumes between ST10 and manually generated contours ($P < 0.05$). Figure 11 highlights agreement for all auto-segmented cardiac substructures as compared to ground truth.

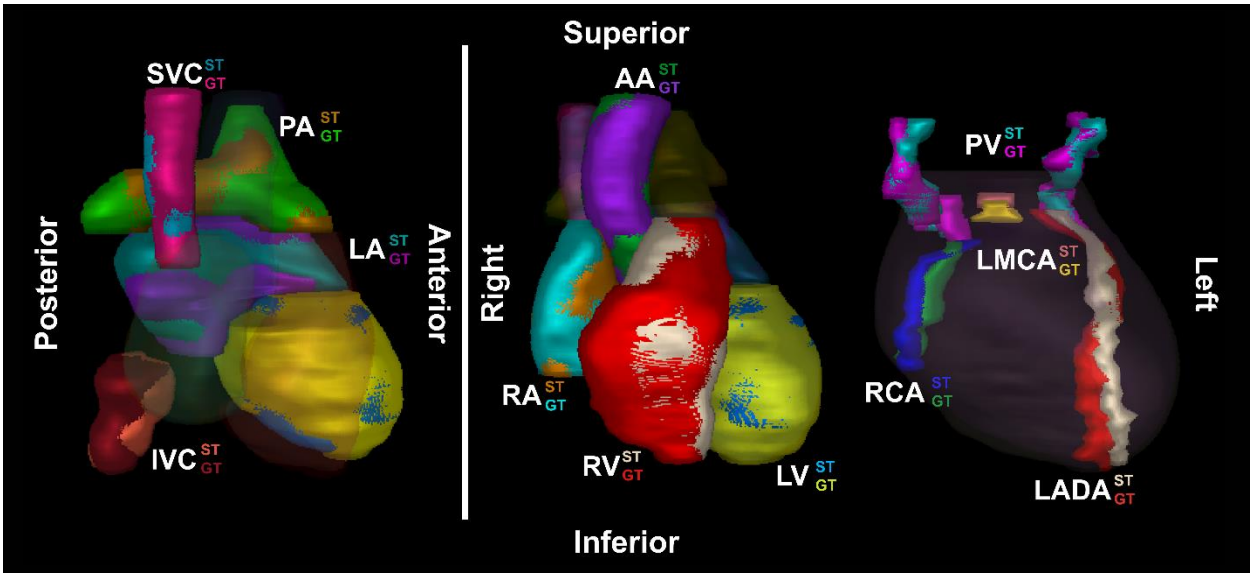


Figure 11: Three-dimensional rendering of substructures showing agreement between manually drawn ground truth (GT) contours and STAPLE 10 (ST) generated contours.

Qualitative Contour Grading using ST10

Physician consensus scores for the heart, ventricles, PA, RA, SVC, IVC, and AA were found to require only minor modifications, typically at the inferior boundary (average score: 4.2 ± 0.5). The pulmonary vein (PV) and LA scored between 3 and 4, requiring moderate modifications. Major modifications were necessary for the LADA (1.8 ± 0.8) and RCA 2.4 ± 0.9) although propagated contours were deemed useful for localization. The left main coronary artery (LMCA) yielded the lowest average score (1.4 ± 0.5), suggesting inadequate segmentation. The highest scoring segmentations occurred for the LVs and AAs, with 3 and 2 subjects requiring no modifications, respectively. Excluding the coronary arteries, average consensus scores across all validation patients were greater than 4 (Table 3), suggesting only minor modifications were necessary.

Dosimetric Assessment

Figure 12 shows a representative dose volume histogram (DVH) for a test subject including the LV, LADA, and heart, as these structures fell within the tangential fields. All other cardiac substructures received negligible radiation dose (mean dose < 1.5 Gy, results not shown). The propagated contour yielded a $LADA_{mean}$ of 23.1Gy and $LADA_{max}$ of 44.9Gy, which was within 3.4% and 0.1% of ground truth, respectively ($LADA_{mean}$ of 22.4Gy, $LADA_{max}$ 44.9 Gy).

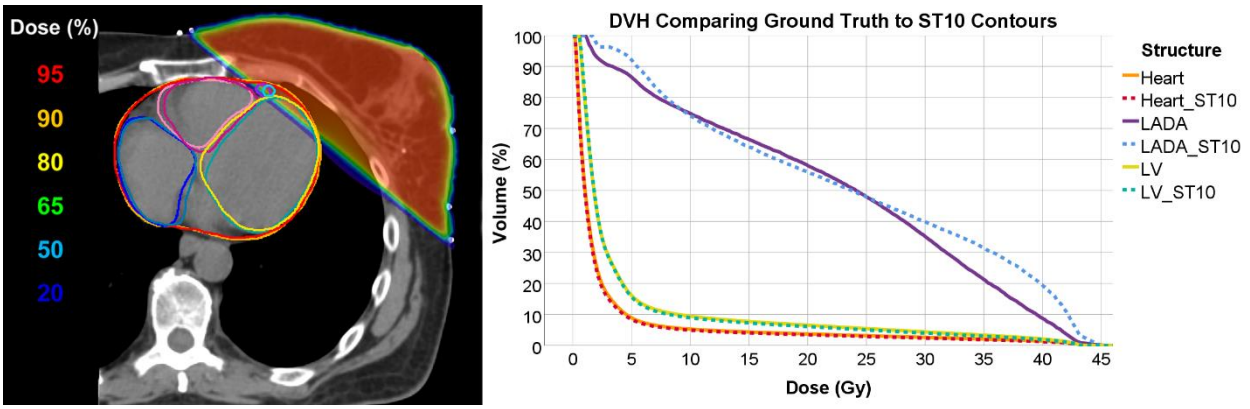


Figure 12: Left: Axial cross section of a treatment planning CT for a representative validation patient showing contours generated from STAPLE 10 (ST10) and ground truth, as well as percentage dose delivered to the left breast (substructure colors not represented in the dose volume histogram (DVH): Dark Blue-RA, Denim Blue-RA_ST10, Pink-RV, Magenta-RV_ST10). Right: Corresponding DVH for the same validation patient.

The Wilcoxon signed rank test showed no statistically significant differences between ST10 and ground truth contours for the minimum, mean, and maximum dose to the chambers and great vessels ($P > 0.05$). Additionally, there were no statistically significant differences in $LADA_{max}$, MHD, and LV-V5 ($P > 0.05$). However, there was a significant difference in dose for the $LADA_{mean}$ ($P < 0.05$). Excellent estimation of the dose to the heart and LV was observed across the 11 test subjects for the propagated contours

proving them to be robust for dosimetric endpoints. The summation of the primary and boost treatment plans for the 11 validation subjects yielded no statistically significant differences in dosimetric endpoints between ST10 and ground truth contours for the LV-V5 (14.9 ± 7.0% vs. 15.3 ± 7.3%), MHD (2.7 ± 1.0 vs. 2.8 ± 1.0 Gy), and $LADA_{max}$ (46.2 ± 6.9 vs. 43.2 ± 9.0 Gy) ($P > 0.05$). However, the difference in $LADA_{mean}$ was statistically significant (22.5 ± 11.2 vs. 18.3 ± 10.0 Gy) ($P < 0.05$).

Discussion

This work has optimized and validated a hybrid MRI/CT contouring atlas for cardiac substructure segmentation with the overarching goal of applying it to non-contrast enhanced CTs for RTP and dose assessment. After a promising segmentation approach was identified (i.e. ST10), accurate delineations were obtained for the heart, chambers, and great vessels (10 of 13 structures), although the coronary arteries were not adequately segmented (DSC < 0.3). While the current retrospective dosimetric evaluation focuses on cardiac substructures for left-sided breast cancer RT, the atlas may be applied to other disease sites, such as advanced stage lung or esophageal cancer, which can be explored in future work.

Although cardiac substructure atlases have been described in the literature, to our knowledge, none have included hybrid MR/CT information for propagation to CT. STAPLE was recently applied to delineate heart chambers on non-contrast enhanced CT images via a fused contrast-enhanced CT[132]. With the introduction of MRI into our atlas, our work outperformed that of Zhou for the heart chambers (average improvements in DSC and MDA of 0.12 ± 0.02 and 2.8 ± 0.5 mm, respectively). A multi-atlas MV method was used to automatically segment cardiac chambers on CT angiography scans from a

large multicenter/multivendor database with little improvement in segmentation accuracy with > 5 atlas matches[45]. When applying MV in our work, a slight improvement in median and interquartile range was observed when using up to 15 atlas matches (Figure 9, right) at the expense of computational time (~10 minutes/dataset). While Zhuang *et al.* did incorporate MRI into their cardiac substructure atlas, they focused mostly on the large structures such as the heart, chambers, AA, and PA[47] and also used contrast-enhanced CT. Our segmentation accuracy for ST10 was comparable for the same structures (DSC and MDA within 0.01 and 0.5 mm, respectively) after applying the MRI/CT atlas to standard RTP CTs.

One limitation is that ground truth contours were generated by a single radiation oncologist. In their atlas implementation for automatic cardiac substructure segmentation on contrast enhanced CT, Zhou *et al.* found that inter-observer variability increased for the PV and coronary arteries where DSCs were less than 0.50 and MDAs were larger than 4.0 mm across experts[132]. Even though ground truth segmentations in this work were generated by a single radiation oncologist, contour verification by a radiologist and our consensus scoring provided additional clinical interpretation by multiple observers. Another limitation of this work is that the performance of the atlas has yet to be evaluated for deep-inspiration breath hold, which has been shown to provide additional cardiac sparing for left-breast cancer cases[151]. Finally, the volumetric T2-weighted cardiac MRI scans were not optimized for RTP (slice thickness=8 mm) although in-plane resolution was 0.7 x 0.7 mm^2. Thinner slice thicknesses will improve contouring accuracy for small volumes, however at the expense of reduced signal-to-noise ratio.

The coronary arteries (i.e. LADA, RCA, and LMCA) were most challenging for our atlas, which is consistent with other studies[132,133,152]. Potential causes of this include complex and low contrast anatomy and image resolution limitations. Although we accounted for respiratory motion by utilizing a local cardiac confined registration, cardiac motion may have adversely impacted the MR/CT fusion accuracy which may introduce additional uncertainty in small structures, such as the arteries. In these cases, manual segmentation was difficult and required expertise. Additionally, significant motion from respiration and the cardiac cycle may present challenges in identifying the coronary arteries as they can often appear indistinct or noncontiguous[141]. A recently reported contouring atlas using landmarks like the atrioventricular and interventricular grooves has also shown to be useful in segmenting the coronary arteries without the use of contrast[152]. The 9 outliers (Figure 9, right) are due to PV segmentations from 2 validation patients where, in both cases, the atlas overestimated the volume and did not reach the inferior extent of the ground truth contours. Additionally, the 3 extreme outliers (DSC < 0.4) (Figure 9, right) were attributed to an inadequate IVC segmentation on a single patient, likely due to the liver appearing homogeneous on non-contrast enhanced CT. Thus, artery and vein segmentation will be further addressed in future work via the application of deep convolutional networks, which have shown promise for ventricle segmentation[153].

The retrospective dosimetric evaluation revealed that for whole breast RT, few cardiac substructures may require assessment. Nevertheless, the maximum dose to the LADA (46.2 ± 6.8 Gy) was substantial, with possible consequences of acute cardiac events[28] and ischemic heart disease[30,150]. Future work may include the evaluation of treatment planning strategies and extend this work to other disease sites that may benefit

from cardiac substructure sparing (e.g. esophagus, lung, or lymphoma). Additionally, further development of this atlas may incorporate the inclusion of cardiac valves and segments of the LV[152].

Conclusion

Overall, applications of the hybrid MRI/CT atlas offer future potential for robust cardiac substructure sparing using standard simulation CTs that are in routine use for treatment planning (i.e., non-contrast CT/4DCT) when an MRI is unavailable. As virtually all patients receiving RT have a CT-SIM as needed for accurate dose calculation, our approach offers strong potential for widespread application. Our hybrid MR/CT atlas shows promise for cardiac substructure segmentation for use in routine treatment planning and dose assessment.

CHAPTER 4 "CARDIAC SUBSTRUCTURE SEGMENTATION WITH DEEP LEARNING FOR IMPROVED CARDIAC SPARING"

Introduction

Increased risks of radiation-induced heart disease including acute (pericarditis) and late (congestive heart failure, coronary artery disease, and myocardial infarction) cardiotoxicities have been linked to dose from thoracic RT for lymphoma, lung, breast, and esophageal cancers[5-7,16]. Radiation-induced heart disease presents earlier than previously expected, beginning only a few years after RT and with elevated risk persisting for ~20 years[16]. Importantly, dose escalation evaluation for locally advanced non-small cell lung cancer in RTOG 0617[26] revealed that the volume of the heart receiving ≥ 5 and ≥ 30 Gy were independent predictors of survival[26]. Further, heart dose/volume metrics are significantly associated with a patient's quality of life[27].

At present, dosimetric evaluation is currently limited to simplified heart volume/dose relationships, such as those recommended by QUANTEC, where the heart is considered a single organ. It is currently recommended that < 10% of the heart receive > 25 Gy with the clinical endpoint of long-term cardiac mortality[18]. Despite having whole-heart dose limits, evidence suggests that dose to sensitive cardiac substructures may lead to cardiac toxicities[16,28,29] including cardiomyopathy, coronary artery disease, pericardial, and conduction system diseases[9]. Specifically, an increased rate of cardiac events and ischemic diseases have been associated with increased radiation dose to the LV[28], LA[29], and LADA[154]. Patel *et al.* found that the maximum dose > 10 Gy to the LADA was a significant threshold for increased odds of developing coronary artery calcification. When compared to MHD, maximum dose to the LADA had a stronger association with coronary artery calcification onset[32]. However, following these dosimetric thresholds is

currently limited by the poor visualization and ability to delineate these sensitive cardiac substructures on non-contrast CT-SIM scans.

Several studies have assessed atlas-based segmentation of cardiac substructures in RT[48,132,155] to avoid the time consuming (6-10 hours/patient) and tedious task of manual delineation[47]. However, most atlases fail in segmenting the coronary arteries, with DSCs between ground-truth and auto-segmentation of the LADA ranging from 0.09-0.27[48,132,155,156]. Incorporating multiple imaging modalities (i.e. contrast enhanced CT and MRI) has improved visualization and yielded successful chamber and great vessel segmentation, yet coronary artery segmentation remains an unmet need[48,132].

Recently, DNNs, such as U-Net[119], have shown great promise for generating accurate and rapid delineations for RT[120]. Here, a DNN learns a mapping function between an image and a corresponding feature map (i.e. segmented ground-truth). Payer *et al.* implemented a U-Net for substructure segmentation and obtained a DSC of 94% in the aorta as compared to ground-truth[157]. Various DNNs have been applied to medical image segmentation[120], specifically for cardiac substructure segmentation. These include deep CNNs with adaptive fusion[92] or multi-stage[157] strategies, as well as generative adversarial networks (GANs)[158]. Adaptation of these segmentation strategies have greatly improved cardiac chamber[157] and pulmonary artery[92] segmentations on contrast enhanced CTs (DSCs > 85%). Additionally, deep residual learning techniques are currently being used to generate cardiac substructure segmentation models that are robust against the presence or absence of image contrast[159]. However, most of these models have not been applied to conventional CT-SIM images and have yet to implement segmentations of the PV and coronary arteries.

The current work builds upon recent DNN results to develop an efficient and accurate cardiac substructure deep learning (DL) segmentation pipeline that can be implemented into routine practice on standard, non-contrast CT-SIMs, thus requiring no additional image acquisitions. Here, training is performed via labeled MRI/CT pairs inputted into a 3D U-Net coupled to predict cardiac substructure segmentations using a single non-contrast CT-SIM input. We further improve agreement to ground-truth delineations by introducing a 3D dense CRF as a post-processing step, which have been recently merged with DNNs for state-of-the-art results in medical image segmentation[160]. Overall, the overarching goal is to enable widespread implementation of DL to improve cardiac sparing in RTP accomplished via cardiac sparing trials and improved risk assessment evaluation.

Methods

Imaging and Ground-Truth Contour Delineation

Thirty-two left-sided whole-breast cancer patients, with 36 unique datasets, were consented to an Institutional Review Board approved study and underwent cardiac MRI scans (two-dimensional T2 single-shot turbo spin echo sequence, repetition time = 927.9 ms, echo time = 81 ms, voxel size = 0.7 x 0.7 x 8.0 mm^3) at EE on a 3T Philips Ingenia (Philips Medical Systems, Cleveland, OH). Imaging was completed in a single breath hold (acquisition time = 22.1 ± 4.4s). Non-contrast CT-SIM images were acquired on a Brilliance Big Bore CT simulator (Philips Medical Systems, Cleveland, OH) (voxel size = 1.1 x 1.1 x 3.0 mm^3-1.4 x 1.4 x 3.0mm^3, 120-140 kVp, and 275-434 mAs) with patients immobilized in the supine position on a Posiboard (Civco, The Netherlands). Twenty-four

patients were imaged under free breathing conditions, while the other eight underwent 4DCT.

To develop a mutual coordinate system between datasets, an automatic global rigid registration between the T2 MR (moving image) and CT-SIM (target image) images was performed in MIM (version 6.9.1, MIM Software Inc., Cleveland, OH). An automated local rigid registration was then applied via a manually drawn cardiac-confined bounding box. For both rigid registrations, normalized mutual information was used as the similarity metric as it has been shown to perform well with multimodality image registration tasks[140]. For patients who underwent 4DCT, the 50% phase was used as it most closely matched the EE MRI.

The evaluation and approval of the co-registration of the T2 MRI to the non-contrast CT was performed through visual verification by a radiation oncologist. To generate the contours, a consensus atlas was followed[141] as implemented in our previous work[48]. In brief, twelve cardiac substructures (left/right ventricles (LV, RV) and atria (LA, RA), superior/inferior venae cavae (SVC, IVC), pulmonary artery/veins (PA, PV), ascending aorta (AA), right coronary artery (RCA), left main coronary artery (LMCA), and LADA) were manually delineated by a radiation oncologist and verified by a radiologist with a cardiac subspecialty. Due to the enhanced soft tissue contrast that MRI provides, preference was given to anatomical information from the MRI.

Data Preparation

All work was performed using an NVIDIA Quadro M4000 graphical processing unit (NVIDIA, Santa Clara, CA). To improve generalizability, zero-mean normalization[161] (subtracting the mean intensity from the image and dividing by the standard deviation of

the image) was performed to maintain intensity consistency across MRI/CT datasets and patients. To generate ground-truth images, substructure masks were combined into a single image volume (ground truth (GT) image in Figure 13) with intensity values for the 12 substructures indexed every 20 grayscale values from 35 to 255 with no overlap among substructures. MR and CT images were all resampled to a 650 x 650 mm in-plane resolution using bilinear interpolation. Bilinear interpolation was also used to interpolate MR images in the z-direction to match the 3 mm CT slice thickness (final voxel size of 1.27 x 1.27 x 3 mm^3). Registered MR and CT image volumes were cropped to 64 slices (in-plane dimension of 128 x 128 pixels), centered on the centroid of the whole heart and padded with 32 blank slices both superiorly and inferiorly for a final size of 128 x 128 x 128 pixels.

Neural Network Architecture and Training

The proposed 3D U-Net, shown in Figure 13 was based on an existing architecture designed for brain tumor auto-segmentation[162] with several customizations as follows: (1) including deep supervision, (2) training using the entire 3D image volume simultaneously via multi-channel data inputs (i.e., MRI, CT, and cardiac substructure ground-truth masks), (3) optimizing hyperparameters of a Dice-weighted multi-class loss function[162], (4) utilizing deconvolution in the upsampling process, and (5) optimizing the number of feature maps used in the first layer.

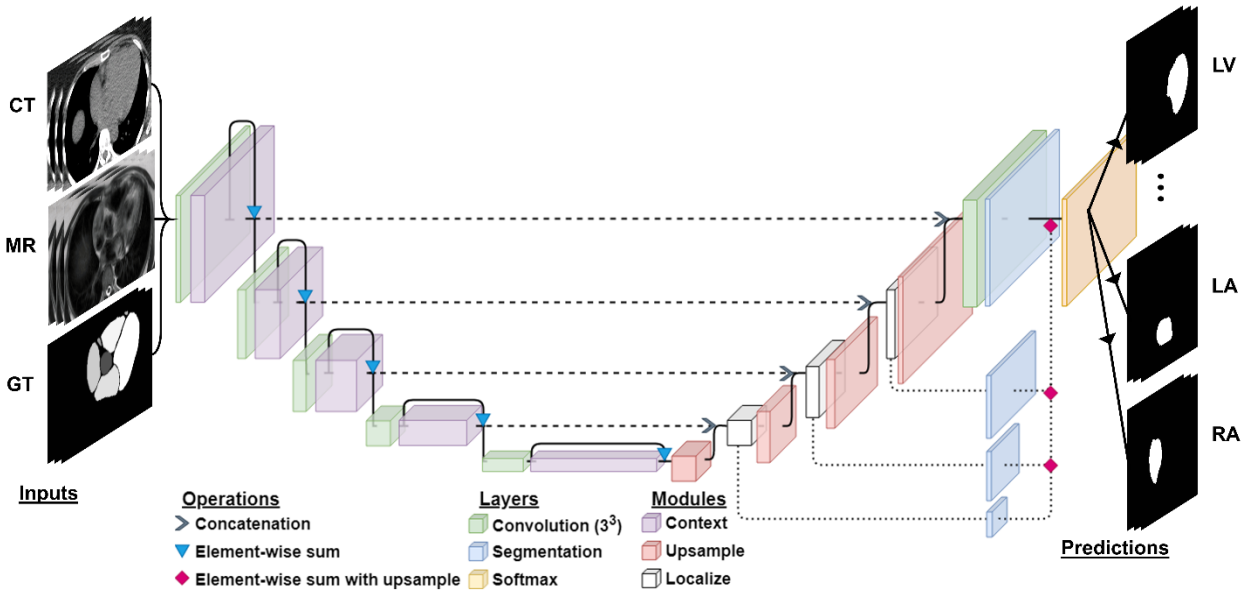


Figure 13: 3D U-Net architecture with CT and MR inputs in different image channels, along with the ground-truth (GT) labels. Prediction maps are outputted for each substructure.

This 3D U-Net is composed of a contraction pathway (Figure 13, left) to aggregate high level information using context modules and an expansion pathway (Figure 13, right) to combine feature and spatial information for localization[162]. Context modules (Figure 13, left) were composed of a dropout layer with 30% probability between two 3 x 3 x 3 convolutional layers. Deep supervision was implemented by adding segmentation layers at each step of the localization pathway (Figure 13, right). Deep supervision allows for the injection of gradient signals deep into the network[163], as it speeds up convergence and enhances training efficiency when there is a small amount of available labeled training data[162,164]. An elementwise summation with upsampling was then applied across all added segmentation layers to generate the final segmentation. As coarse segmentation results may yield unrealistic results, skip connections were applied (i.e. concatenation) by fusing earlier layers in the network where the down-sampling factor is smaller to recover

the original spatial resolution[165]. To rebuild high-resolution feature maps, deconvolution was used in the localization pathway in order to learn the upsampling[166].

To minimize model overfitting (ensuring the model remains generalizable to the hold-out dataset after being tuned to a training set), data augmentation[119] including flipping, rotating (0-30°, 1° increments), scaling (± 25%, 1% increments), and translating (10 pixels in the left-right, anterior-posterior, and superior-inferior directions) was applied. Originally proposed as a novel objective function based on DSC[167], a Dice-weighted multi-class loss function was used[162,168] to manage the different image features among substructures, as shown in Equation 9:

$$\mathcal{L}_{DSC} = -\frac{2}{|B|}\sum_{b\in B}\frac{\sum_a x_{a,b}y_{a,b}}{\sum_a x_{a,b}+\sum_a y_{a,b}} \quad (9)$$

As label encoding is not sufficient for model training, y represents the ground-truth segmentation map converted from categorical to binary variables (i.e. one-hot encoding). For training voxel a in class b, $x_{a,b}$ and $y_{a,b}$ represent the prediction and ground-truth, respectively. As a larger DSC represents better overlap between ground-truth and the prediction, the loss function is negative due to it being minimized during the training. Each value of correspondence between both the training and validation datasets to ground-truth are represented by an average across all 12 substructures.

An adaptive momentum estimation optimizer[169] was used along with randomly initialized weights. Patience (i.e., number of epochs to wait without validation loss improvement before reducing the learning) was also implemented during training. An epoch is defined as one forward pass and one backward pass (i.e., backpropagation process) of all the training samples[117]. Optimized hyperparameters included an initial

learning rate of 5 x 10^{-4}, 50% learning rate reduction, a batch size of 1, patience of 10 epochs, and 16 base filters in the first layer of the localization pathway.

Patient data was split into 25 patients for training data and 7 (11 unique datasets) patients for a hold-out data set for network testing. No hold-out datasets used for testing were implemented in the network training. Training data was split via random assignment into 80% training and 20% validation data. Paired MRI and CT data were placed into separate image channels along with indexed ground-truth labels for 25 patients to train the 3D U-Net using the entire 3D MR and CT images and all substructures simultaneously. Training was considered to be sufficiently converged when the training error between two adjacent epochs (i.e. one forward and backward pass of all the training samples) was less than 0.001[117].

All work was performed using Windows 10 operating system in Python version 3.6. The 64-bit Microsoft Windows system is equipped with a quad-core Intel® Xeon® CPU-E5-1630 v4 at 3.70GHz and 16GB of memory. The employed graphics processing unit was an NVIDIA Quadro M4000 with 8 GB of RAM and 1664 CUDA cores where Keras 2.0 was implemented with a TensorFlow backend.

Contour Post-Processing and Optimization

As coarse output maps from the DL network may containing holes and spurious predictions from neural networks are common[121], contour post-processing was performed on the 3D U-Net output using a fully connected CRF[170] that imposes regularization constraints through minimizing an energy function[113]. A 3D-CRF model was developed based on an initial two-dimensional implementation[113] and optimized to refine

segmentations by smoothing, filling holes and removing false positives, such as small remote regions.

3D-CRF was implemented on a GPU for improved computation and inference time. Inference here is with regard to the number of iterations applied to minimize the Kullback-Leibler divergence[113]. Both bilateral[171] (i.e. appearance kernel) and Gaussian[113] (i.e. smoothness kernel) pairwise energies were used to account for the grayscale intensity similarity, as well as the spatial proximity of pixels. The applied kernel involves the sum of a smoothness and appearance kernel which are shown in the following equations[113]:

$$smoothness\ kernel = \ k_s = exp\left(-\frac{|P_\alpha - P_\beta|^2}{2\theta_x^2}\right) \tag{10}$$

$$appearance\ kernel = \ k_a = \ exp\left(-\frac{|P_\alpha - P_\beta|^2}{2\theta_x^2} - \frac{|Q_\alpha - Q_\beta|^2}{2\theta_y^2}\right) \tag{11}$$

$$applied\ kernal = w_1 * k_s + \ w_2 * k_a \tag{12}$$

where P and Q represent Intensity and position vectors at pixel (α, β). The smoothness kernel works to remove small remote regions[172] and is controlled by a scaling factor θ_x. Parameter θ_y is an additional scaling factor in the appearance kernel, which controls the degree of similarity in predicted pixels. The appearance and smoothness kernel are equally weighted with weights $w_1 = w_2 = 1$.

CRF hyperparameters θ_x and θ_y were optimized automatically[173] for each substructure by stepping through different parameter values and then comparing prediction results back to ground-truth through DSC. Based on the range of utilized values found in the current literature[113,174,175], full integer values from 1-80 were stepped through for θ_x and in steps of 0.05 from 0-1 for θ_y, yielding a total of 100 individual tests. As it was shown in[113] that convergence may be reached in less than 10 iterations, 10 inference

steps per substructure prediction were used[173]. Finally, predictions both with and without CRF post-processing will be assessed by comparing agreement with ground-truth.

Network testing on the hold-out dataset was conducted using the remaining 11 test patient CTs containing the heart and thorax from seven unique patients. Binary mask segmentations were converted to contours in Digital Imaging and Communications in Medicine (DICOM) format and imported into MIM.

Evaluations and Statistical Assessment

Quantitative evaluations between DL and ground-truth segmentations were performed via DSC[176], MDA (average of the shortest distance between all voxels of the predicted and ground-truth segmentations[148]), Hausdorff distance (HD, maximum nearest neighbor Euclidean distance[177]), and centroid displacement in three cardinal axes. DL segmentations were also compared to our previously published multi-atlas (MA) results, which implemented STAPLE with 10 atlas matches[48] using a shared cohort of 11 test subjects. Lastly, qualitative consensus scoring of DL segmentations was conducted to evaluate clinical utility. Before the qualitative grading was performed, three physicians (two radiation oncologists and a radiologist with a cardiac subspecialty) reviewed DL segmentations from a patient who was excluded from the grading. The physicians were instructed on the image grading system and a grading consensus scale was established for each substructure and then applied for five unique patients. Qualitative consensus scoring was completed on five of the test subjects as evaluated in our previously published atlas study[48]. Scoring was completed using a 5-point scale[48,149] as follows: (1) not clinically acceptable, (2) clinically acceptable with major changes, (3) clinically acceptable with moderate changes, (4) clinically acceptable with minor changes, (5)

clinically acceptable. Contours were converted to a 0.25 mm high resolution display for final evaluation in MIM.

For volume size similarity assessment, 2-tailed Wilcoxon signed-ranks tests were performed between ground-truth and auto-segmented DL segmentations, with $P < 0.05$ considered significantly different. Statistical assessments using 2-tailed Wilcoxon signed-ranks tests were also used to compare DL segmentations to our previous MA method via DSC, MDA, and qualitative consensus scores.

Results

Segmentation and Post-Processing Time

The initial manual ground-truth delineations of the 12 cardiac substructures required ~ 3 hours per patient. The DL network stabilized in ~ 19.4 hours after training the network for 200 epochs, including ~ 2 hours after implementing data augmentation. Figure 14 shows the results for the training and validation datasets over the 200 epochs. The final training and validation DSC values were 83.1% and 81.5%, respectively (difference < 2%), which represents an average over all 12 substructures.

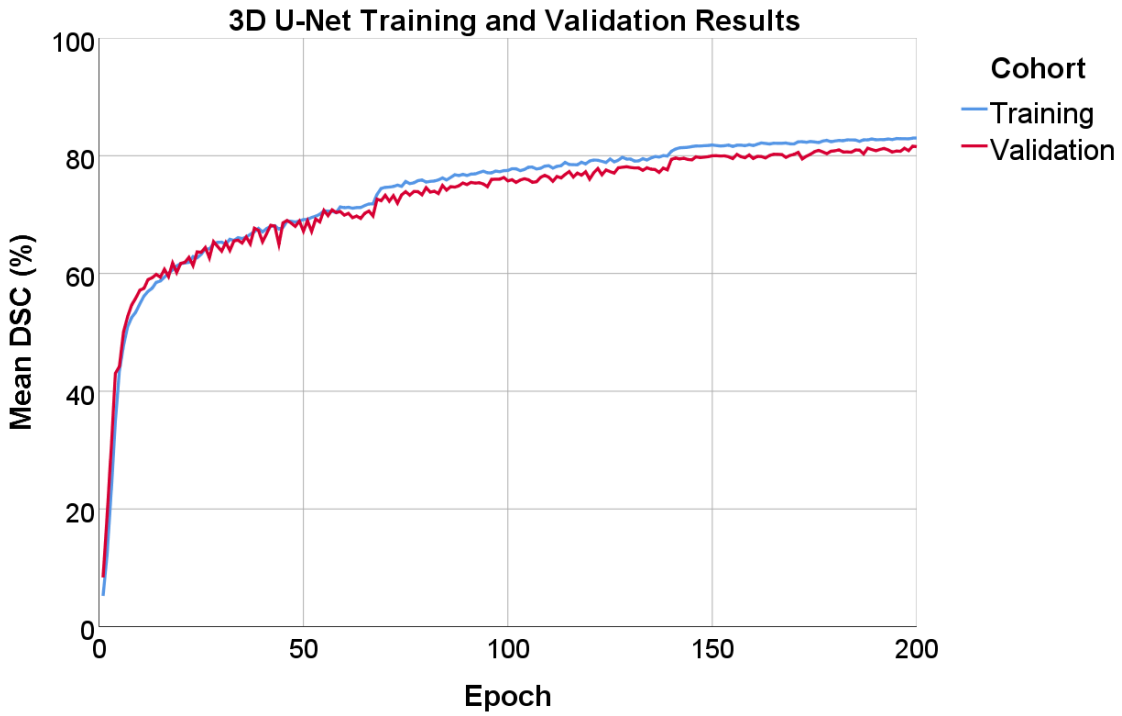


Figure 14: 3D U-Net training and validation results over 200 epochs. Values for mean Dice similarity coefficient (DSC) represent an average over all 12 substructures.

Augmentation led to an overall DSC increase of 5.0 ± 7.9% across all substructures, with greatest improvements in the coronary arteries (LMCA = 18.6 ± 15.5%, RCA = 8.7 ± 9.1%, LADA = 7.8 ± 7.1%). Substructure contour generation (12 substructures) for a new patient using a single non-contrast CT-SIM dataset input took 5.0 ± 0.6 seconds. CRF post-processing time from a single test patient using 10 inference steps for 12 substructures was 9.3 ± 0.3 seconds, for a total DL generation time of 14.3 seconds (range: 13.5-15.6 seconds).

CRF Post-Processing

CRF hyperparameter optimization revealed that differing values of (θ_x,θ_y) provided maximal DSC when three different sets of optimized parameters were employed for (1) coronary arteries and PV (2.0, 0.40), (2) superior/inferior venae cavae (2.0, 0.50), and (3)

chambers and great vessels (8.0, 0.55). CRF applications lead to an average improvement in DSC, MDA, and HD over all substructures of 1.2 ± 2.5%, 0.11 ± 0.31 mm, and 5.58 ± 14.25 mm, respectively. The LMCA had the greatest improvement in DSC (6.2 ± 6.6%, range: 1-22%) after CRF application, whereas the RV and RA saw the least improvement (0.3 ± 0.2%, range: 0.0-1.0%). The LV showed the greatest improvement in MDA (0.3 ± 0.5 mm, range: 0.0-1.2 mm) and HD (34.4 ± 23.0 mm, range: 0.1-64.7 mm) after 3D-CRF application. Lastly, after applying CRF, the mean improvement in MDA ranged from 0.04 to 0.21 mm over the 12 substructures.

Geometric Performance of Segmentation

DL segmentation results are presented in Table 4.

	Segmentation Improvement			
Substructures	**Augmentation DSC**	**CRF HD (mm)**	**Final DSC**	**Final MDA (mm)**
Chambers (LA, LV, RA, RV)	0.03 ± 0.03	10.23 ± 19.34	0.88 ± 0.03	1.53 ± 0.26
Great Vessels (SVC, PA, AA)	0.03 ± 0.05	3.59 ± 10.79	0.85 ± 0.03	1.24 ± 0.31
Inferior Vena Cava	0.00 ± 0.05	6.61 ± 15.23	0.78 ± 0.04	1.45 ± 0.45
Pulmonary Veins	0.05 ± 0.04	2.82 ± 4.41	0.77 ± 0.04	1.04 ± 0.21
Coronary Arteries				
Left Anterior Descending Artery	0.08 ± 0.07	0.16 ± 0.27	0.53 ± 0.08	1.90 ± 0.90
Right Coronary Artery	0.09 ± 0.09	5.01 ± 15.86	0.50 ± 0.09	1.97 ± 0.46
Left Main Coronary Artery	0.19 ± 0.16	0.65 ± 0.96	0.50 ± 0.18	1.27 ± 0.68

Table 4: Improvement in automatic segmentation performance in Dice similarity coefficient (DSC) after augmentation and in Hausdorff distance (HD) after implementation of conditional random fields (CRF) post-processing. The table also shows the final agreement to ground-truth via DSC and mean distance to agreement (MDA). Additional abbreviations defined in the text.

Figure 15 presents comparisons between ground-truth and DL segmentations across substructures (LMCA not shown). The best-case patient (Figure 15, right) had chamber DSCs greater than 0.90 and MDAs less than 2 mm for all substructures with favorable

results for the RCA (DSC = 0.72, MDA = 1.67 mm). MDA (Figure 17, left) across all 12 substructures was less than 2.0 mm (MDA = 1.46 ± 0.50 mm).

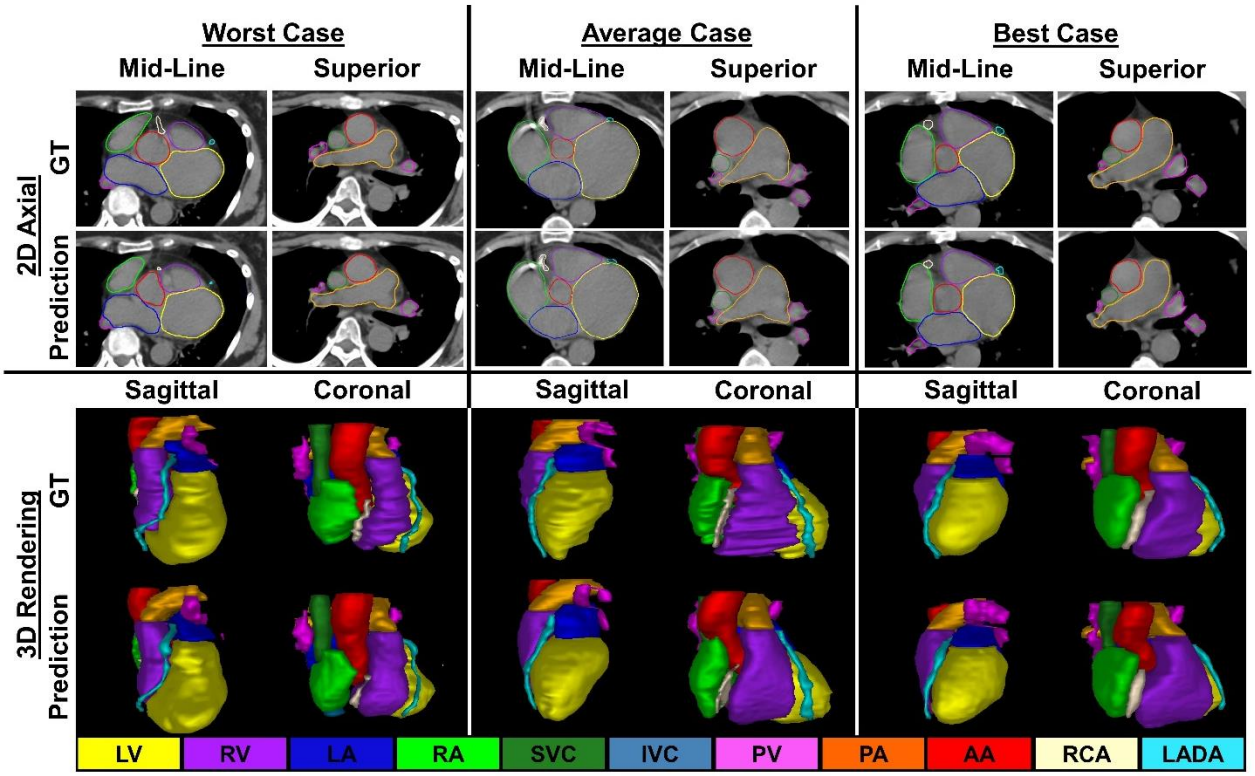


Figure 15: Comparisons between contours generated via deep learning prediction and ground-truth (GT) in both two-dimensional axial slices (top) and 3D renderings (bottom) for the worst (left), average (center), and best (right) cases.

Wilcoxon signed-ranks tests revealed no significant differences in cardiac substructure volumes between DL and ground-truth ($P > 0.05$). Figure 16 summarizes the centroid shifts in all cardinal axes. On average, the smallest displacements (< 2 mm) occurred in the anterior-posterior direction for 11 substructures. The largest displacements occurred in the superior-inferior direction.

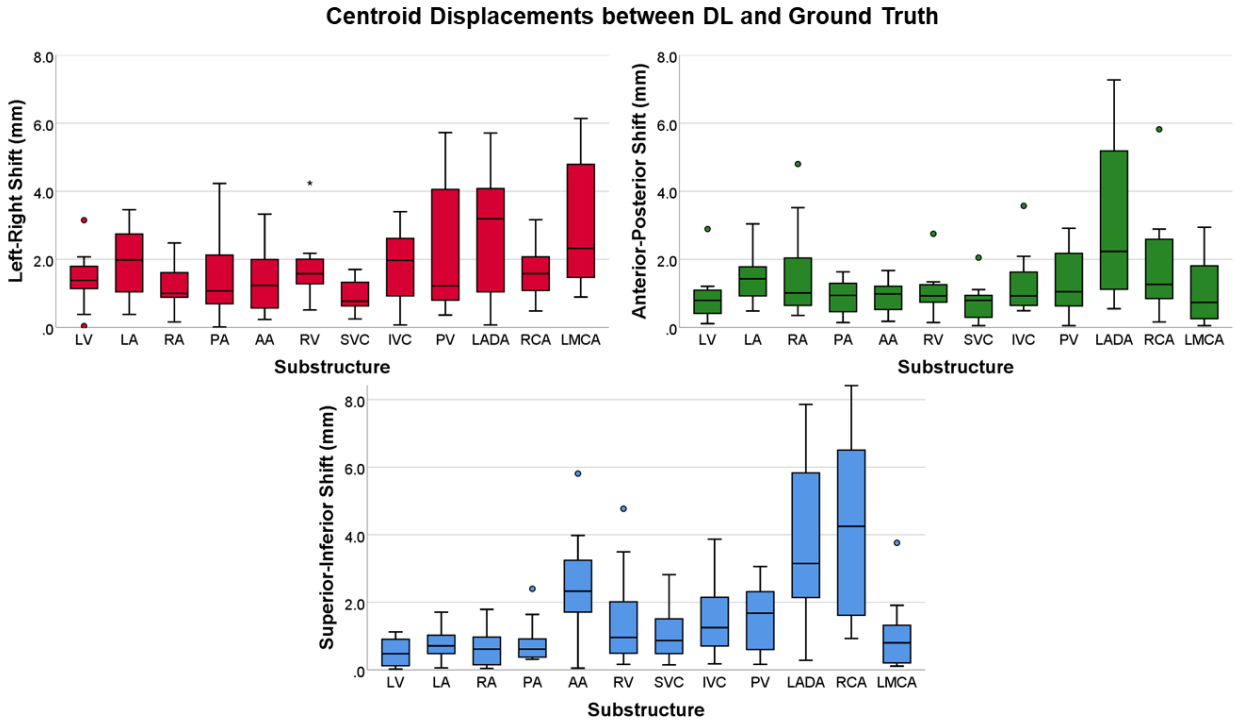


Figure 16: Substructure centroid displacements in the left-right (left), anterior-posterior (right), and superior-inferior (bottom) directions. Legend: interquartile range = box, median = line, minimum and maximum = whiskers, circles and stars = 1.5 and 3 times the interquartile range, respectively.

DL vs. MA Segmentation

Figure 17 summarizes MDA and DSC results over 11 test cases and compares DL with our previously developed MA method for the same cohort[48]. MDA and DSC for all cardiac substructures improved with DL. Specifically, DSC agreement to ground-truth increased 3-7% for chambers, 9-11% for the superior/inferior venae cavae and PV and reached 23-35% for the coronary arteries. On average, MDA improved by ~1.4mm with DL, with greatest agreement in the SVC (MDA = 0.99 ± 0.15 mm) and worst agreement in the RCA (MDA = 1.97 ± 0.46 mm). For four test CTs, our DL method yielded LMCA contours, whereas our previous atlas-based model failed to produce any segmentation.

Overall, DL provided a significant improvement ($P < 0.05$) over the previous MA method for every substructure in terms of MDA and DSC.

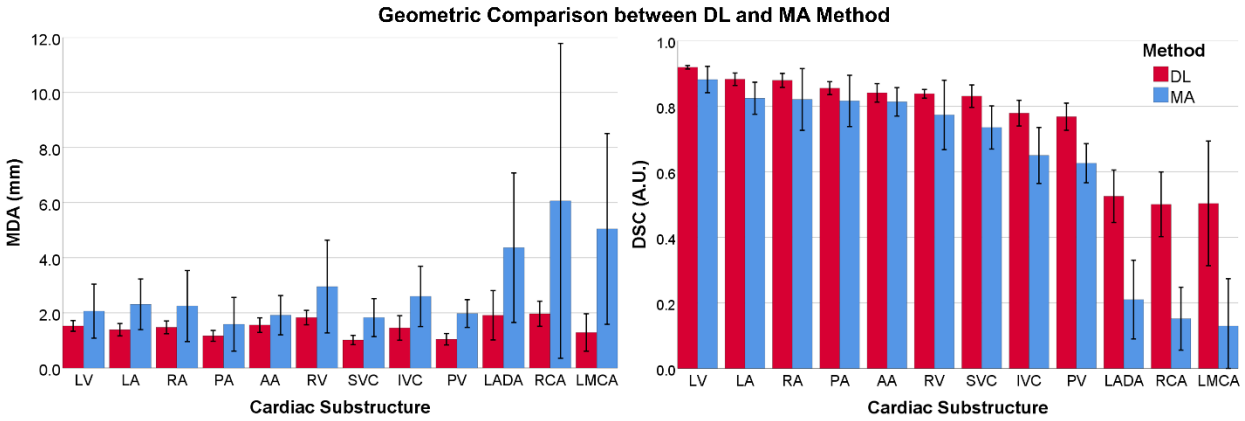


Figure 17: Agreement between manually drawn ground-truth and auto-segmentation methods (Blue: Previous multi-atlas method (MA), Red: Novel DL method) over 11 test cases. Left: Mean MDA, Right: Mean DSC.

Qualitative Analysis

Physician consensus scores are summarized in Figure 18. All patients had clinically acceptable contours (score of 5) for the LV, RA, and RV (results not shown), while the SVC, PA, and PV had clinically acceptable contours for 4/5 patients with DL. For the cardiac chambers, 6/20 comparisons between DL and MA methods were equivalent, while all others improved by at least one grade with DL. The LMCA and RCA had the lowest average scores of 3.0 ± 1.0 and 3.8 ± 0.4, respectively, with all other substructures scoring an average of ≥ 4.4. DL provided significant improvements ($P < 0.05$) over the MA method for the LADA, RCA, PV, PA, SVC, LA, RA, and RV. Improvements in 44/60 (5 patients, 12 substructures) qualitative scores were observed with DL. For only one instance, DL scored worse than MA (AA: grade 4 to 3). For two

LADA segmentations, MA yielded a grade of 1 (clinically unusable) and improved to a 5 (clinically acceptable) with DL (Figure 18, right).

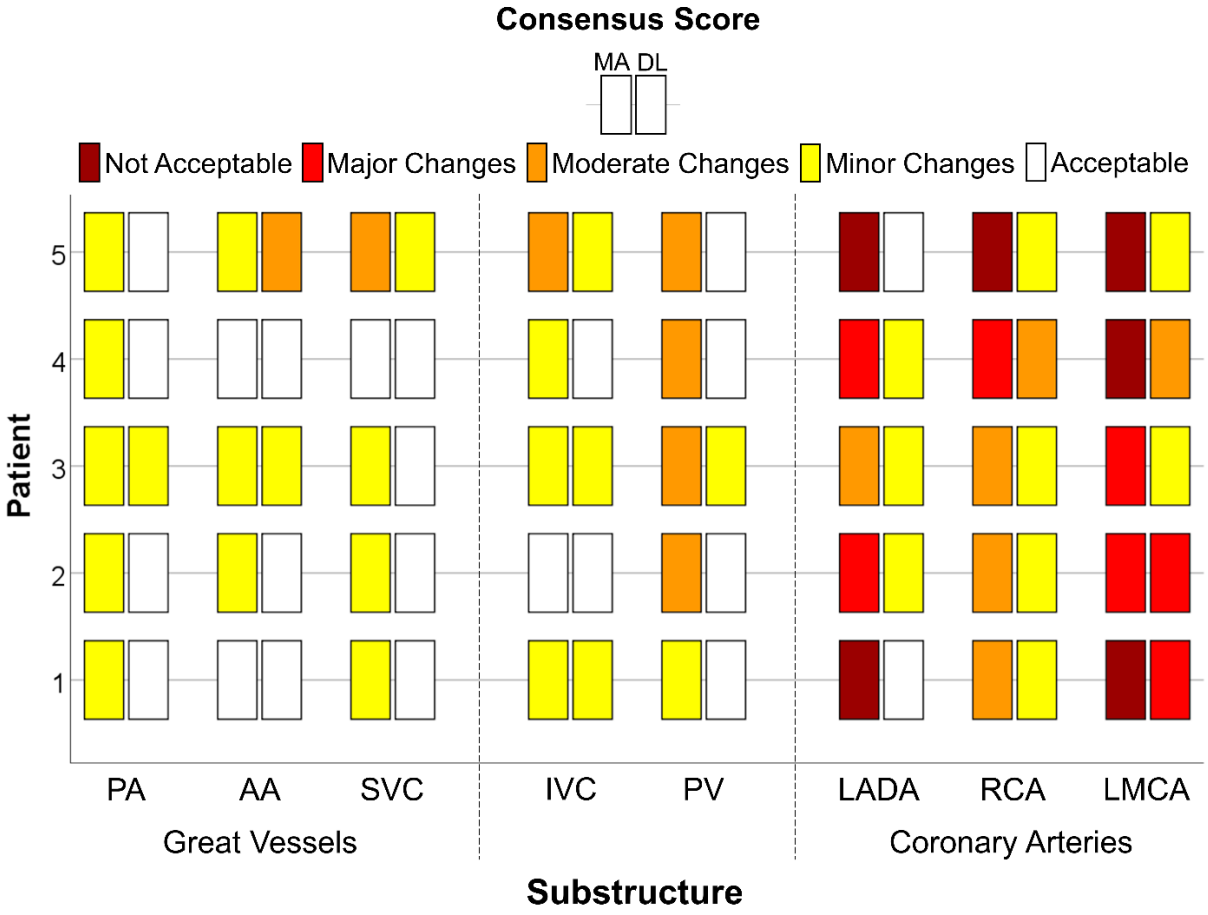


Figure 18: Qualitative consensus scoring (not clinically acceptable, clinically acceptable with major changes, clinically acceptable with moderate changes, clinically acceptable with minor changes, clinically acceptable) of five patients for the multi-atlas (MA) and deep-learning (DL) based image auto-segmentations (chambers not shown). For each substructure column, the MA and DL methods are shown on the left and right, respectively.

Discussion

This work presented a novel DL pipeline to segment sensitive cardiac substructures using a 3D U-Net with the principal goal of applying to non-contrast CT-SIM for RT planning. Data augmentation and CRF post-processing improved DL contour

agreement with ground-truth. Overall, our method provided accurate segmentations of the chambers, great vessels and PVs, and led to promising results in coronary artery segmentation on non-contrast CT-SIM datasets[178].

While cardiac substructure segmentation has been explored previously, to our knowledge, none have included paired MR/CT multi-channel data inputs to yield robust segmentations on non-contrast CT inputs. Several atlas segmentation methods have been recently published[48,132,155] and report cardiac chamber DSCs > 0.75. However, these methods have had limited success segmenting coronary arteries as atlas methods rely on image registration quality and are unable to consider large amounts of patient data due to computational demands[179]. Our work parallels recent applications of DNNs where CTCA scans specifically optimized for cardiac imaging were utilized. Here, DSC in the RA (87.8%)[157] and PA (85.1%)[92] were within 1% of our DL method, while we were within 5% of their chamber segmentation results. Our work adds to the current literature by including additional substructures and allowing for predictions to be made on non-contrast CT-SIM scans.

Data augmentation improved DL segmentation accuracy by ~5% across all substructures. Although no comparison values exist in the literature for cardiac substructure segmentation, this value is consistent with studies performed on liver lesion segmentation[180,181]. One extreme outlier (greater than 3 times the interquartile range) observed for the RV in the left-right axis occurred for the worst-case patient (Figure 15, left), where the heart was rotated clockwise and shifted posteriorly/left. While this patient's anatomy was an anomaly, this result may be addressed in the future by further augmenting the data (i.e., rotation > 30^{o}). Furthermore, both the LADA and RCA had

larger centroid shifts in the superior-inferior plane (Figure 16, bottom). This can be further visualized in Figure 15 (left), where the inferior extents of the LADAs and RCAs were underrepresented with some narrowing of these substructures observed in the midline axial slice. To address this, recent atlas-based methods have restricted the size of the LADA to 4 mm throughout its entire length[156]. Nevertheless, our DL pipeline performed well for coronary artery contours on non-contrast CTs (DSC ~ 0.50, MDA < 2.0 mm), particularly as compared to recent atlas results where coronary artery (LADA, RCA, and LMCA) DSCs ranged from 0.09-0.27[48,132,155] and had MDAs > 4 mm[132]. Coronary artery segmentations may be improved through the use of high resolution (0.78 x 0.78 x 1.6 mm^3) CTCA[92] that use contrast and yield DSCs ~ 60%[182]. Additionally, implementing a Dice loss function weighted on the inverse of the class size may improve the results for smaller substructures such as the coronary arteries. Originally proposed by Crum *et al.*[183], the generalized Dice loss function has been shown to improve hyperparameter robustness for unbalanced tasks (i.e. when each class is not represented equally in the dataset), and improve overall segmentation accuracy for small structures[184].

While rare cases involved the removal of spurious remote predictions that resided within the ground-truth delineation, 3D-CRF led to an overall improvement in segmentation agreement. The coronary arteries experienced the greatest improvement from CRF post-processing, with the LCMA improving ~ 6% in DSC. Additionally, there were improvements in MDA up to 1.21 mm and 1.96 mm for the LV and LA, respectively. Aside from removing spurious outlying points, CRFs also improved the smoothed appearance of the segmentations as needed for clinical application[185]. CRF tuning required different parameters for cardiac substructures based on size and shape, much

like the work completed by Rajchl *et al.*[186]. The improvement in segmentation agreement observed, along with the use of a 3D-CRF to remove spurious isolated regions, parallels other emerging uses of 3D-CRF post-processing in medical imaging[187,188]. Although this study implemented CRFs as a post-processing step, some current studies have integrated CRFs into the utilized neural network and have seen improved segmentation performance[173,174,189] and can be explored in future work for possible coronary artery segmentation improvement.

The overall time to generate DL segmentations on pre-processed CT-SIM data was rapid: 14 seconds for all 12 substructures. This value can be compared to Mortazi *et al.* who segmented seven cardiac substructures in ~ 50 seconds on high resolution CTCA and 17 seconds on MRI[92]. Moreover, our previous MA method required ~ 10 minutes to generate substructure contours per patient without post-processing[48].

Although the in-plane resolution was 0.7 x 0.7 mm^2, our study may have been limited by the 8 mm slice thickness of the MRI. Despite our data augmentation techniques, increasing the training sample size may further improve segmentation results. However, similar training and testing cohort sizes with augmentation have been used previously[190]. While paired cardiac MRI/CT data are commonly limited for cancer patients, the training cohort may be expanded in the future by applying our DL model to generate additional ground-truth segmentations. Data quantity may also be increased through utilizing unlabeled images for unsupervised learning via generative models such as, a cycle-[191] or a stacked-[192] GAN, which implement multiple GANs for data synthesis. Recently, Zhang *et al.*[158] proposed a novel cardiac chamber segmentation method using a GAN integrating cycle- and shape-consistency. They obtained DSCs comparable to atlas segmentations

(DSC ~ 0.75) on CT and MRI by using ~ 14% of real data and augmenting their dataset by incorporating synthetic MRI and CT data into training. Our model may be enhanced similarly by incorporating synthetic images in the network training, while also providing additional substructures, such as pulmonary veins and coronary arteries. Nevertheless, even with the current limited training dataset, our results outperform other currently available approaches. As shown in Figure 14, training and validation results increased to a point of stability with a difference of < 2% after convergence. Moreover, to further limit potential overfitting in this more limited cohort, data augmentation (i.e., flipping, scaling, rotating, and translating) and model regularization (dropout = 0.3) were implemented.

As both the CT and MR images were acquired in breath hold conditions, respiratory motion is assumed to be negligible during this study. However, due to extended scan times and heart rate, one limitation of this study is that numerous cardiac cycles are captured during imaging. Thus, the substructures are represented by their average position over the course of the scan and cardiac motion is not taken into consideration. Currently, cardiac motion is not managed clinically due to limitations in available treatment technologies. Nevertheless, the magnitude of cardiac motion is on the order of 3-8 mm[52] suggesting internal motion may be incorporated into future margin design as has been previously proposed[33,53].

As MR-guided RT and MR-only planning become more prevalent, future work will include training an MR-only model. It has been recently recommended that the LADA be included as an avoidance structure in RTP[32], thus a natural clinical endpoint of this work includes dosimetric analysis and implementing cardiac avoidance strategies via accurate and efficient cardiac substructure segmentation made possible by DL.

Conclusion

These promising results suggest that our novel DL application offers major efficiency and accuracy gains for cardiac substructure segmentation over previously published MA results, using only non-contrast CT inputs. Future work involves further refinement of coronary artery segmentation using conditional random fields as a recurrent neural network and through expanding the patient cohort. Coupled with robust margin design, improved cardiac sparing in treatment planning can be realized.

CHAPTER 5 "QUANTIFYING INTRA-FRACTION MOTION AND INTER-FRACTION SETUP UNCERTIANTIES"

Part 1 "Characterizing Sensitive Cardiac Substructure Excursion due to Respiration"

Introduction

Radiation dose to the heart from thoracic cancer treatments can increase the possibility of the patient experiencing diseases such as ischemic heart disease, cardiomyopathy, and artery atherosclerosis[193,194]. Currently, only whole heart dose estimates are considered for RTP in clinical practice[16]. However, recent studies have shown that dose to individual substructures in the heart may be better indicators of future cardiac events than whole heart dose metrics[28]. The superior soft tissue contrast that MRI provides allows for these sensitive cardiac substructures to be visualized. However, the use of MRI for thoracic cancer treatment is not the standard of care due to technical and accessibility limitations. For example, CT is more prevalent for diagnosis because it allows for lung nodule depiction down to 1-2 mm[195], whereas MRI is sensitive to lung nodules of 5-11 mm[196]. Additionally, MRI may be limited in the diagnosis of lung cancers due to susceptibility artifacts caused by several air-tissue interfaces[197]. Through the use of MRI to inform cardiac substructure delineations on thoracic CT images, our objective is to quantify substructure excursion during respiration to identify dominant axes of displacement per substructure. Based on our segmentation work described previously, we were able to incorporate sensitive cardiac substructures into the treatment planning process, however accurate radiation dose assessment may be complicated by respiratory motion influences[198]. An example of the displacement of the heart over the respiratory cycle is shown in Figure 19.

Figure 19 displays both an end-inhalation (EI, 0% phase) and end-exhalation (EE, 50%, phase) 4DCT image for a representative patient. In both the axial and sagittal views, a delineation of the whole heart on the EI phase is shown in red to reveal the cardiac displacement in various axes. Additionally, there is a difference map on the right side of Figure 19 for all cardinal axes.

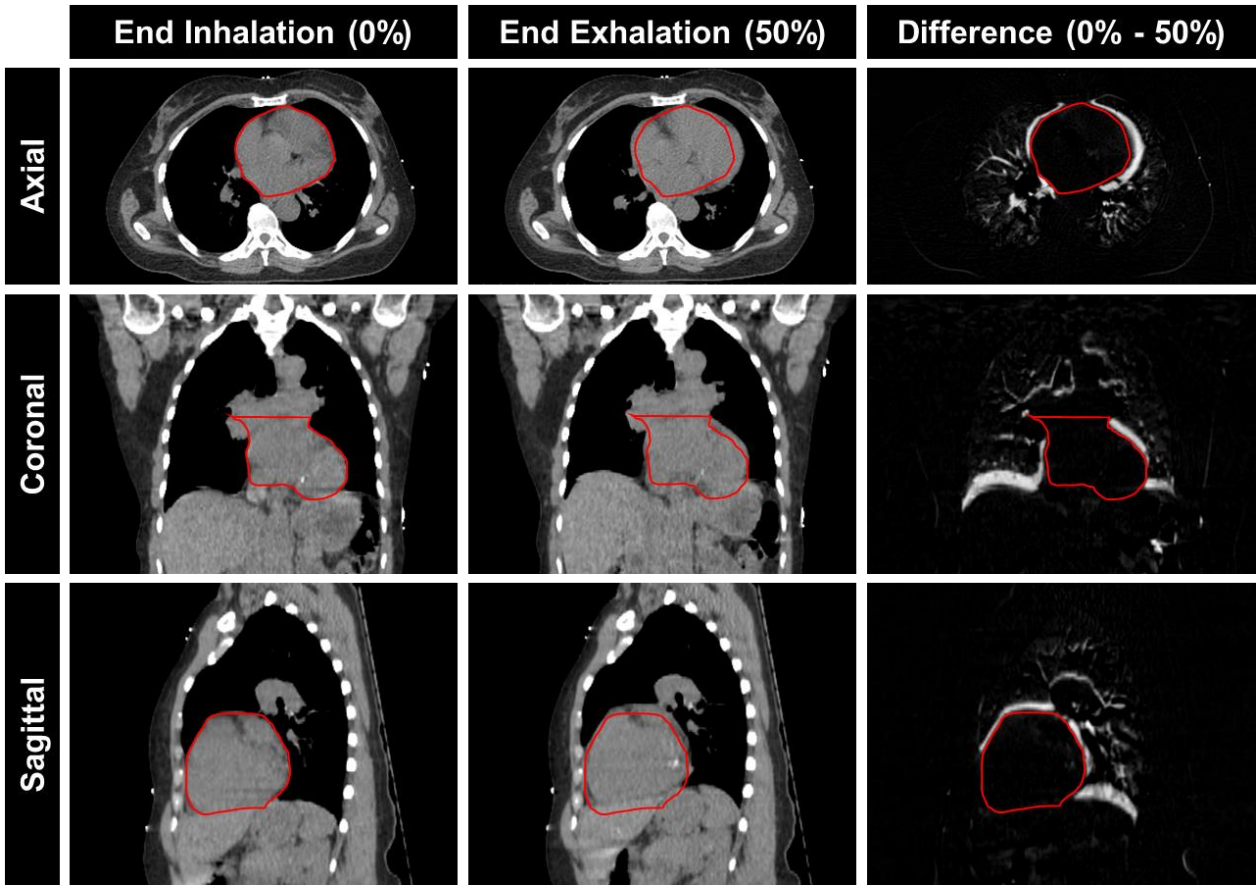


Figure 19: Example of whole heart displacement in between respiration at end inhalation (0% phase, column 1) and end exhalation (50% phase, column 2). Difference maps of the 0% minus the 50% phase are shown in column 3 with a representation in each cardinal axis. The end inhalation delineation of the heart is shown on each image in red.

Several motion management techniques, including but not limited to, gating and breath-hold, have been incorporated clinically to manage a patient's respiratory motion during radiation treatment or image acquisition[199]. DIBH techniques have been

implemented in cancer treatments to position the heart further from the irradiation field for left breast cancer treatments. However, even minor displacements can have a large dosimetric effect when using highly conformal radiation therapy techniques like intensity modulated radiation therapy (IMRT) or volumetric modulated arc therapy (VMAT)[50,200]. These dosimetric effects are heightened when small structures, like the coronary arteries, are introduced[201].

Various studies have reported on how intra-fractional motion affects dose delivery to the heart from thoracic cancer treatments. A study by George *et al.* studied the effects of intra-fractional motion for breast cancer treatments using IMRT at three different types of respiration (i.e. shallow, normal, and heavy)[202]. They found that lung and heart dose increase with respiration[202]. Yue *et al.* expanded on these findings by studying how intra-fractional motion affects changes in DVH metrics for left breast treatments using 10-phase 4DCT[203]. They found the maximum dose to the heart can vary up to 6 Gy in respiration[203]. Along with studying how intra-fractional motion in RT affects the heart in left-sided breast cancer treatments using 10-phase 4DCT, El-Sherif *et al.* included the LV, and LADA[201]. They found that even though the 95% confidence interval of the four-dimensional dose was ± 0.5 Gy for the whole heart, it varied ± 8.7 Gy for the LADA.

In a similar fashion, Guzhva *et al.* completed a study using 10-phase 4DCT on 20 patients to analyze the cardiac and respiratory combined intra-fractional motion influence on 12 cardiac substructures for patients undergoing RT for thoracic cancers[50]. Cardiac substructure segmentations were completed on the 50% phase and then deformably propagated to the remaining phases[50]. They found that the largest centroid displacements from intra-fractional motion were in the craniocaudal axis (i.e. superior-inferior), and that

the cardiac chambers experienced the smallest displacements overall (largest displacements in the coronary vessels)[50]. The work presented our study builds upon the current literature by providing population results and including additional cardiac substructures, such as the LMCA and the great vessels (i.e. SVC, AA, and PA). Additionally, the current study provides volume statistics across respiratory phases and a dosimetric analysis.

In this study, we utilized 4DCT data and DIR to quantify the excursion of 12 cardiac substructures along with the whole heart during respiration to characterize displacement and identify dominant axes of excursion. The work completed here may be used to complete a future study on generating an accurate motion model for cardiac substructures, from which a robust safety margin can be defined. These safety margins would ensure adequate cardiac sparing and potentially prevent patients from experiencing future cardiac toxicities.

Methods

Patient Cohort and Imaging

Eleven patients with cancer (8 left-sided breast cancer and 3 lung cancer patients) were retrospectively reviewed on an Institutional Review Board approved study conducted at the Henry Ford Cancer Institute. These patients either underwent 4-phase (n=8, breast cancer patients) or 10-phase (n=3, lung cancer patients) non-contrast 4DCTs. All patients were imaged with a respiratory correlated 4DCT and cardiac gated T2-weighted MR in EE. Reconstructed data was exported from the clinical scanners and de-identified for analysis. The EE phase of the 4DCT was rigidly registered with the EE

MRI and the registration was refined with an assisted alignment surrounding the heart using MIM (version 6.9.1, MIM Software Inc., Cleveland, OH).

Segmentation of Ground Truth

Thirteen cardiac substructures including the heart, left/right ventricle (LV/RV), left/right atrium (LA/RA), pulmonary vein (PV), pulmonary artery (PA), ascending aorta (AA), superior/inferior vena cava (SVC/IVC), left anterior descending artery (LADA), left main coronary artery (LMCA), and right coronary artery (RCA) were automatically segmented using multi-atlas[48] (n=8) and deep learning techniques[204] (n=3) that used hybrid MRI/CT information on the EE phase of the 4DCTs. These EE delineations were verified and if needed, corrected by a physician before propagating contours to the other phases. All physician segmentation delineations and corrections followed a recent consensus contouring guideline for cardiac substructure segmentation[141].

Physician verified contours on the EE (i.e. 50% phase) of the 4DCT were deformed to the other phases using a constrained, intensity-based, free-form DIR based on Demons which minimizes the intensity differences between two single modality datasets[205]. The Demons DIR technique, originally proposed by Thirion *et al.*[206], is widely used for its accuracy and computational efficiency[207]. This same DIR algorithm from MIM software has been implemented in several other CT to CT DIR studies and achieved high accuracy[51,208,209]. Specifically, Piper *et al.* applied a known deformation to a CT volume and found that this deformable image registration technique averaged 1.1 mm error from the gold standard[145].

The location of each substructure at end-inhalation (0%), end-exhalation (50%), and two intermediate phases (25-30%, 70-75%) were evaluated in this study. To conduct

the DIR, a box-based assisted alignment around the heart is completed. The DIR workflow creates a set of grid control points using a coarse-to-fine, multi-resolution approach and incorporates user-specified DIR locks[208]. Regularization is added to discourage folds and tears in the deformation field[208]. The multi-resolution approach prioritizes larger registration differences and then refines the registration to account for smaller local changes, which allows for anatomical alignment even when there are large differences present due to respiration[208]. With regard to expected registration uncertainty, residual registration deformation errors in the lung were 0.8 ± 0.4 mm[208]. Once DIR was used to propagate cardiac substructure segmentations to each phase of the 4DCT, final contour verification was conducted by a radiation oncologist.

Analysis for Statistical and Quantitative Comparisons

Measurements of centroid locations, volume at each respiratory phase, as well as maximum excursion between phases were exported from MIM for subsequent analysis. Maximum excursions in each direction were reported as mean ± SD. The displacements that exceeded 5 mm for each of the 3 cardinal axes were evaluated based on guidance provided by the respiratory motion management report produced by the American Association of Physicist in Medicine Task Group report number 76[210]. Paired t-tests were employed for statistical analysis of each substructure to compare volumes of the contours between all 4 phases. Any *P*-value less than 0.05 was considered as a statistically significant difference. All statistical analysis was performed using SPSS version 25.0 (SPSS, Chicago, IL, USA).

Dosimetric Analysis

Dosimetric assessment included all patients (n=11) and plans were originally completed on the 50% phase. Assessment consisted of tabulating the mean (D_{mean}) and maximum dose (D_{max}) and locating the largest dosimetric change over the respiratory phases. As different treatment sites (i.e. breast lung) are represented in this patient cohort, percent variation in dose was also used to assess population values.

Results

Cardiac Substructure Centroid Displacement Summary

Centroid displacements for the left-right (L-R), anterior-posterior (A-P), superior-inferior (S-I), and vector shifts for the patient population are shown in Figure 20. Maximal vector displacements ranged from 5-10 mm across substructures. Vector displacements were largest for the IVC and the RCA, with displacements up to 17.9 mm. Of the three cardinal axes, intra-fraction centroid displacements were observed to be the largest in the S-I axis. Maximum displacements of greater than 5 mm were found for 24.8%, 8.5%, and 64.5% of the cases in the L-R, A-P, and S-I axes, respectively. As shown by the green boxplots in Figure 20, 10 of the 13 structures considered in this study had median intra-fraction centroid displacements that were equal to or greater than 5 mm of displacement in the S-I axis, as shown in Figure 20. Further, only the S-I axis had a 95th data percentile that extended past 15 mm (IVC). For 10/13 studied cardiac structures, data for the first three quartiles was less than 5 mm in the L-R axis. With regard to outliers, seven out of eight of the substructure outliers in the L-R axis can be attributed to a single patient, as discussed further in Figure 22. Over all substructures, the A-P was the axis with the least excursion. As shown in Figure 20, median excursions for 11/13 structures were smallest

in the A-P axis. Table 5 also summarizes the L-R, A-P, and S-I maximal displacements for each cardiac substructure over the patient population.

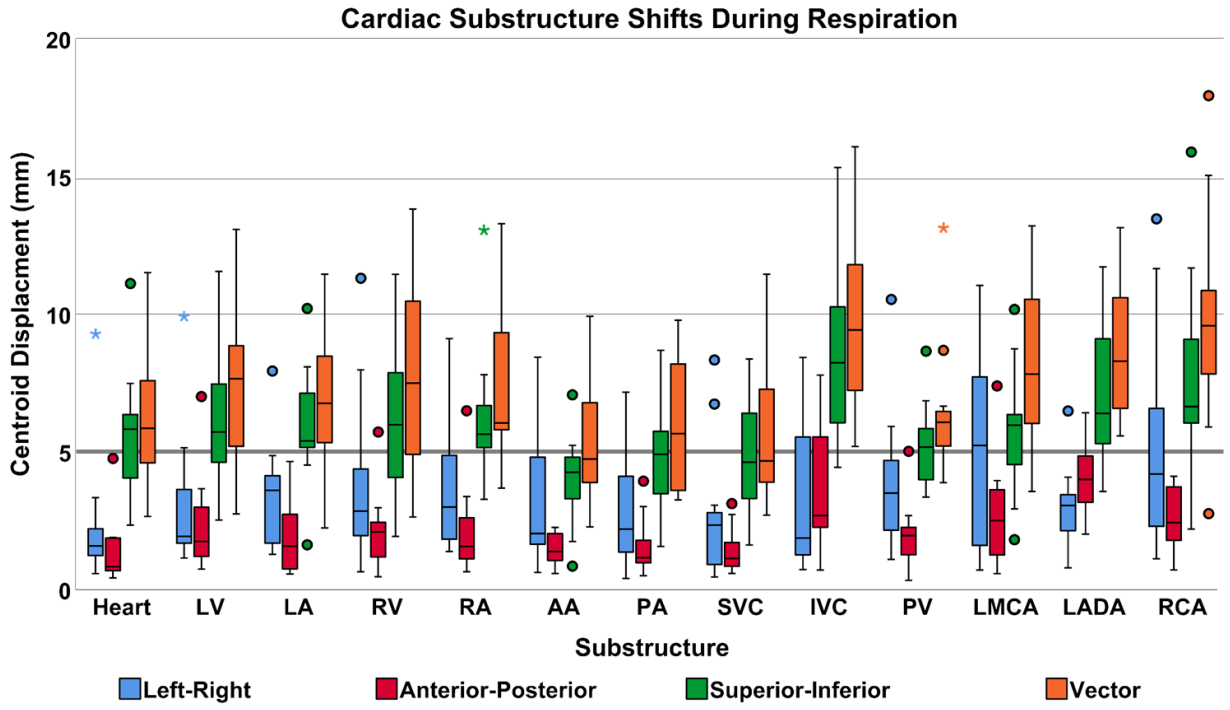


Figure 20: Intra-fraction centroid displacement comparison between all 13 structures for each direction: left-right, anterior-posterior, superior-inferior, and vector. Boxplots, thick line, and whiskers represent the interquartile range (IQR), median, and 5th and 95th percentiles, respectively. Data points displayed as a small circle represent a value greater than 1.5 times the IQR and the star represents a value greater than 3 times the IQR.

With regard to regional displacement, Table 5 reveals that the great vessels (i.e. the AA, SVC, and PA) had the least amount of excursion along each axis, also confirmed by Figure 20. Both the IVC and the RCA, cardiac substructures at the inferior aspect of the heart, had maximal centroid displacements that were greater than 15 mm. Figure 20 and Table 5 also reveal that the IVC was the substructure with the largest displacements in the S-I axis. For the RCA, nine out of 11 patients had centroid displacements that exceed 5 mm in the S-I axis while the IVC had eight patients exceed this threshold. The

RCA and IVC each had the largest maximum vector displacements of greater than 9 mm, as shown by Table 5.

Substructure	L-R (mm)	A-P (mm)	S-I (mm)	Vector (mm)
Heart (n=9)	2.5 ± 2.7	1.5 ± 1.4	5.9 ± 2.5	6.6 ± 3.1
LV	3.1 ± 2.6	2.4 ± 1.8	6.3 ± 2.8	7.6 ± 3.4
LA	3.4 ± 2.0	1.9 ± 1.3	6.0 ± 2.2	7.0 ± 2.7
RV	3.9 ± 3.1	2.1 ± 1.4	6.1 ± 3.0	7.7 ± 3.6
RA	3.7 ± 2.5	2.1 ± 1.7	6.2 ± 2.6	7.5 ± 3.1
PA	2.9 ± 2.2	1.6 ± 1.1	4.9 ± 2.0	6.0 ± 2.5
AA	3.1 ± 2.5	1.4 ± 0.6	4.0 ± 1.7	5.4 ± 2.4
SVC	2.8 ± 2.6	1.4 ± 0.9	5.0 ± 2.3	5.7 ± 2.8
PV	3.9 ± 2.6	2.0 ± 1.2	5.2 ± 1.6	6.5 ± 2.5
IVC	3.2 ± 2.7	3.8 ± 2.5	8.5 ± 3.3	9.8 ± 3.4
RCA	5.3 ± 4.1	2.6 ± 1.2	7.8 ± 3.7	9.8 ± 4.1
LADA	3.0 ± 1.5	4.1 ± 1.3	7.1 ± 2.6	8.7 ± 2.5
LMCA	5.0 ± 3.5	2.7 ± 2.0	5.8 ± 2.4	8.2 ± 3.1

Table 5: Maximum displacement of individual cardiac substructures over 11 patients throughout the respiratory cycle in each cardinal axis (left-right (L-R), anterior-posterior (A-P), and superior-inferior (S-I)) and vector displacements. Substructure abbreviations defined in the text.

Volume and Statistical Analysis

The paired t-tests revealed that out of 39 volume comparisons per patient (13 structures, 4 phases), there were 4 instances in total where $P < 0.05$ for the volume comparisons, showing reasonable maintenance of geometric and anatomical properties. The average volume of the whole heart across all patients at the 50% phase was 742.9 cc. On average, the percent difference in volume for the whole heart between the 0% and 50% phases was 1.2 ± 0.5 %. The cardiac substructures with the largest variabilities in

volume between EE and EI had average volumes greater than 70 cc (PA and cardiac chambers) and are shown in Figure 21. The RA was the cardiac substructure with the largest volume differences between EE and EI of 7.8 ± 6.5 % (range: 0.9 to 20.2 %).

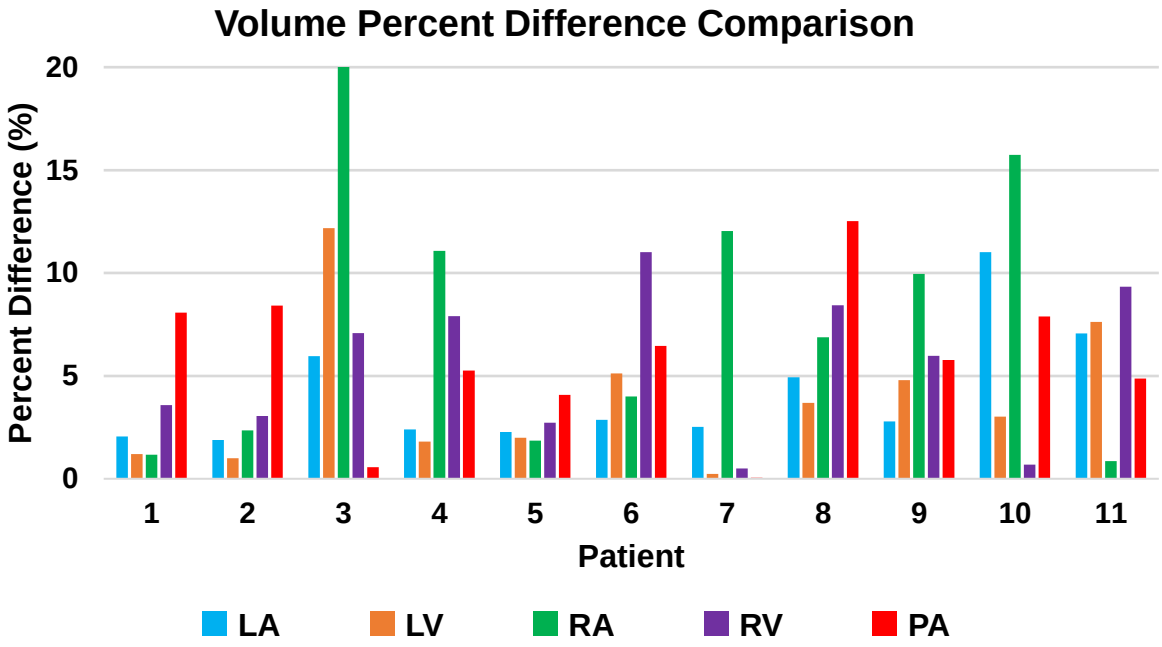


Figure 21: Volume percent difference between end-inhalation and end-exhalation for select substructures over all studied patients

Excursion analysis of the 4-phase 4DCTs revealed that 52.5% and 31.9% of maximum excursions occurred between the 0 and 50%, and the 0 and 70% (or 75%) trajectories, respectively. Maximum excursions only occurred between the 0 and 30% (or 25%) phases 6.4% (9/141 instances) of the time.

Individual Patient Cases

Figure 22 shows substructure excursion between the 0% phase (bottom row) and the 50% phase (top row) images for two representative patients. Results for these patients are also shown in both the axial and sagittal axes. Patient 1 (Figure 22, left) was selected as it experienced minimal centroid displacement for cardiac substructures over

respiration. Patient 1 had smaller than average vector displacements for 11/13 cardiac structures (LMCA had the largest vector excursion for this patient of 11.8 mm).

Even though the largest displacements were observed in the S-I axis across the population (as stated in the section above), Figure 22 also shows Patient 9 (Figure 22, right) as they experienced the largest L-R displacement across patients. This patient exhibited the largest substructure displacements in the L-R axis (contrary to the S-I tendency), with L-R intra-fraction displacements for 12 out of 13 cardiac structures exceeding 5 mm and displacements up to 13.5 mm for the LADA. In reference to Figure 20, Patient 9 (Figure 22, right) accounts for eight out of nine of the substructure outliers in the L-R axis (blue dots and stars). In the S-I axis, Patient 9 had centroid displacements for all substructures greater than 5.0 mm and up to 8.7 mm.

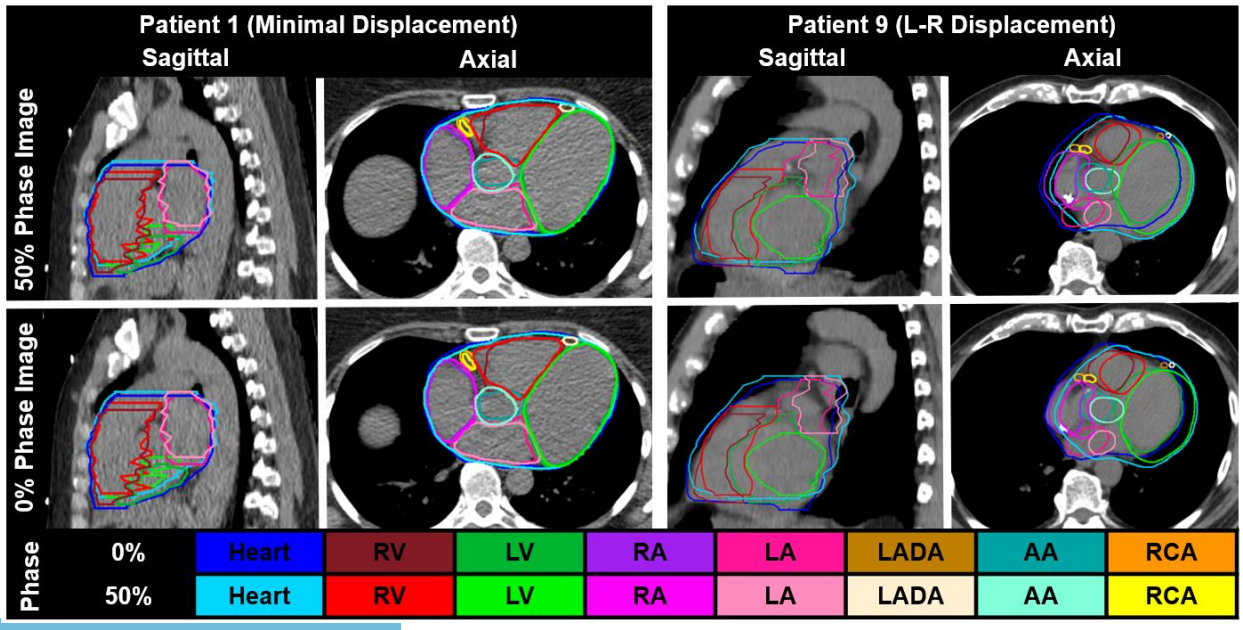


Figure 22: Two representative patients showing substructure excursion between 0% (bottom row) and 50% phase (top row) images with the contours from each phase shown on both image sets for the axial and the sagittal axes. Left: Patient 1 selected for minimal displacement over respiration. Right: Patient 9 chosen for largest left-right (L-R) displacement across patients. Cardiac substructure abbreviations are defined in the text.

Figure 23 displays results for Patient 3, where the average displacements across substructures for the L-R, A-P, and S-I directions were 5.1 ± 3.1 mm, 2.2 ± 1.6 mm, and 10.0 ± 3.8 mm, respectively. As can be seen, there were intra-fraction centroid displacements for the RA and LV in the L-R axis (axial view of the 0% phase image), while the LA and RV had negligible displacements (less than 1 mm). The sagittal view, on the left of Figure 23, emphasizes that the maximum displacements occur in the S-I axis for the LA, AA, and heart. The large excursions found for this case may be attributed to the patient's abnormal anatomy of preceding scoliosis. Additionally, the heart is rotated into the left lung as can be seen in Figure 23.

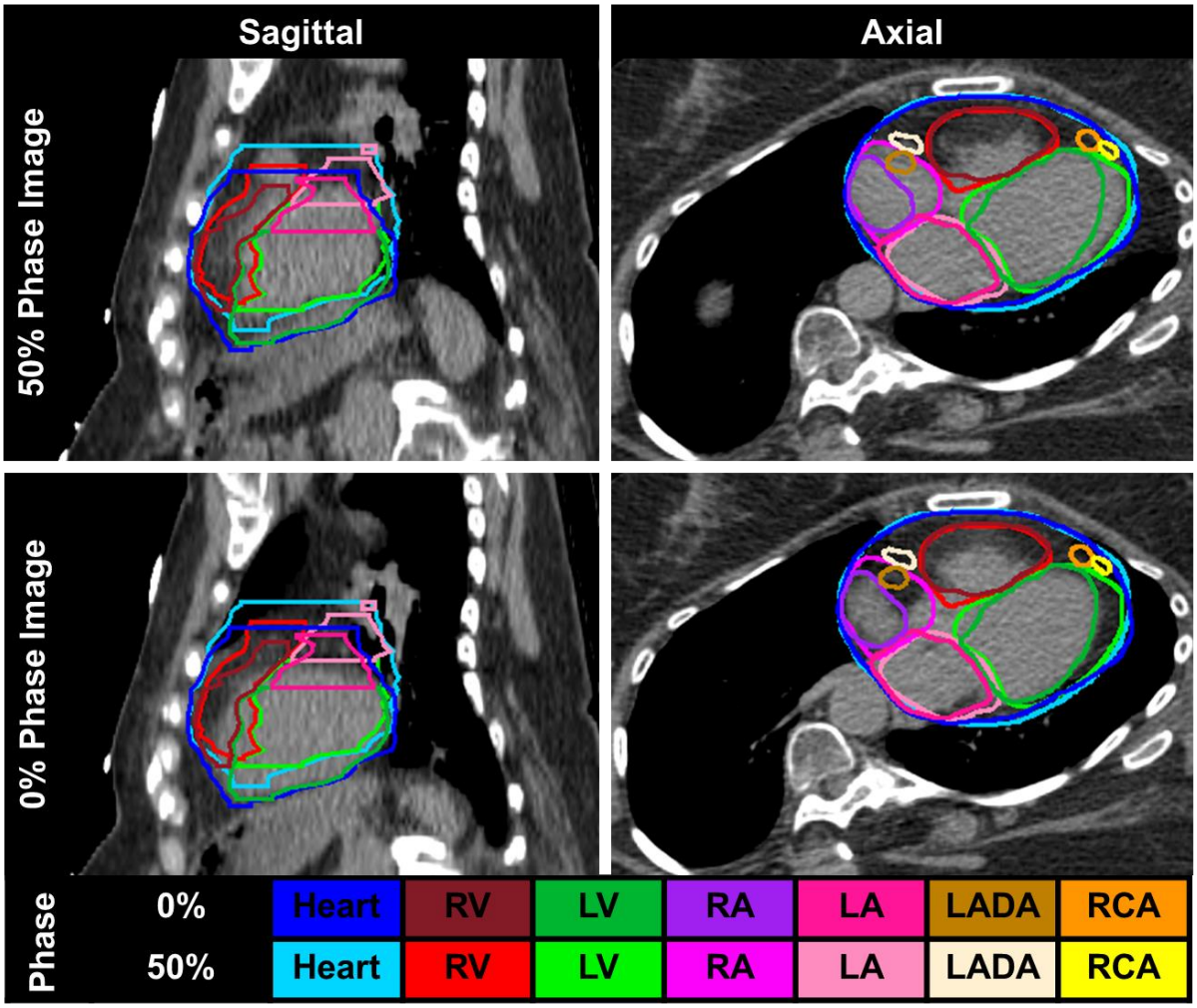


Figure 23: Representative patient (Patient 3) showing substructure excursion between 0% (bottom row) and 50% phase (top row) images with the contours from each phase shown on both image sets for the axial (right) and the sagittal (left) planes.

Dosimetric Analysis

Over the eight studied breast cancer patients, the LADA D_{mean} varied 3.03±1.75 Gy (range: 0.53 to 5.18) throughout respiration. Whereas, whole heart D_{mean} changed 0.18±0.09 Gy (range: 0.06 to 0.37). Figure 24 shows two representative breast cancer patients with the heart, ventricles, and LADA displayed over respiratory phases. Both Patient 2 and Patient 6 were treated to 42.72 Gy in 16 fractions for a stage 1A malignant

neoplasm of the left breast. For these patients the LADA (Figure 6, green) had a D_{mean} that varied >3.5 Gy in respiration.

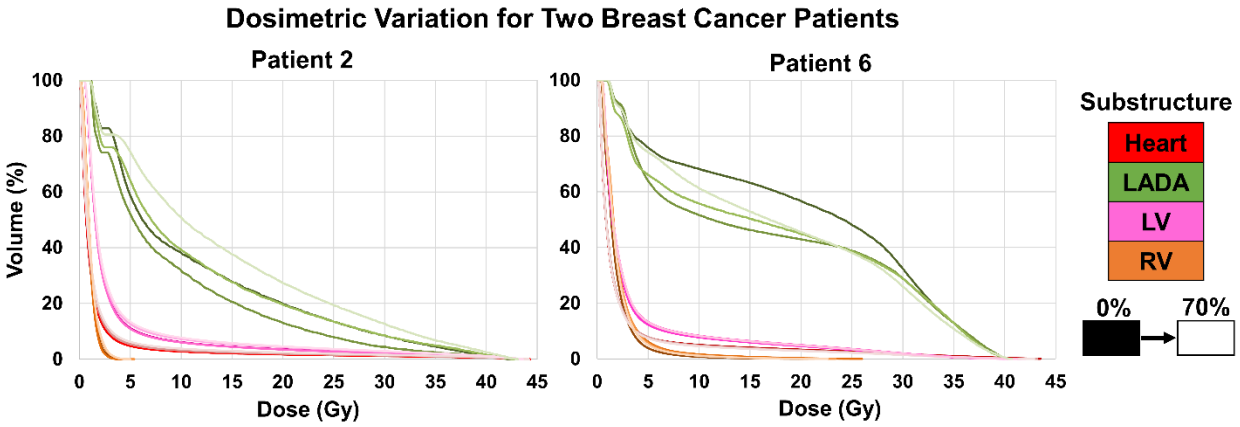


Figure 24: Dose volume histograms for Patients 2 (Left) and 6 (Right) showing the dosimetric variation across respiratory phases for the heart and select cardiac substructures. Substructure color gradient transitions from dark to light as the respiratory phases pass from the 0% to the 70% phase, respectively. Cardiac substructure abbreviations defined in the text.

Patient 9 (shown in Figure 22, right) received RT for a malignant neoplasm of the upper right lung bronchus (non-small cell lung cancer) of 60 Gy in 20 fractions. Of the three lung cancer patients studied, Patient 9 experienced the largest changes in dose across respiratory phases, which is shown in Figure 25. Patient 9 experienced an average change in D_{max} of 3.2±2.9 Gy (range: 0.46 (PA) to 9.05 Gy (RA)) across cardiac substructures. With regard to D_{mean}, the average change across substructures was 2.2±1.8 Gy. For the other two lung cancer patients, no cardiac substructure had D_{mean} differences >1.4 Gy (D_{max} differences up to 5.6 Gy).

Figure 25 shows DVHs for Patient 9 displaying the dosimetric variation across respiratory phases for the great vessels (top), the cardiac chambers (middle), and the coronary arteries (bottom). Each DVH also shows the dosimetric variation in respiration across the whole heart, which is shown in red. The great vessels and chambers had more

variation in dose throughout respiration as compared to the whole heart (Figure 25). Specifically, although changes in D_{mean} for the heart were <0.5 Gy, the SVC (shown in blue in Figure 25, top) had a D_{mean} difference up to 5.4 Gy. The bottom of Figure 25 also shows a DVH displaying coronary artery dosimetric variation in respiration. The LMCA, shown in burgundy, experienced a change in D_{mean} up to 4.6 Gy.

Patient 3, represented in Figure 23, received RT for a malignant neoplasm of the left breast (state IIIA, T2, N2) of 45 Gy in 25 fractions. Between the 25% and 75% phase, the RV experienced an 8.7 Gy change in D_{max}. Regarding D_{mean}, the LADA experienced a 4.2 Gy change between the 0% and 25% phases. For both Patient 3 and Patient 9 however, changes in D_{max} and D_{mean} for the heart were <0.5 Gy. The DVHs outlined highlight an increased dosimetric sensitivity through local dose changes that is not captured by the whole heart.

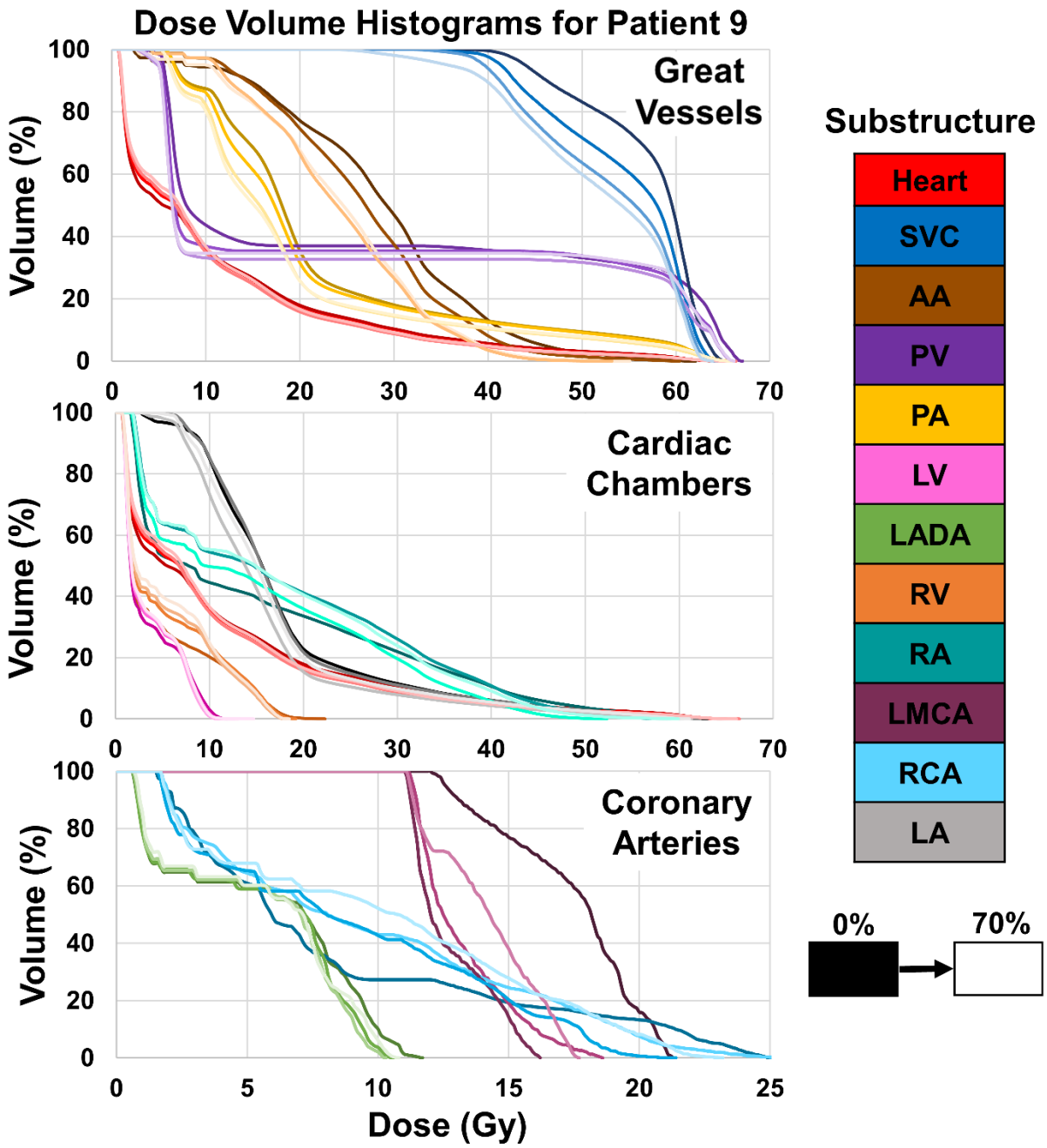


Figure 25: Dose volume histograms for Patient 9 showing the dosimetric variation across respiratory phases for the great vessels (top), the cardiac chambers (middle), and the coronary arteries (bottom). Substructure color gradient transitions from dark to light as the respiratory phases pass from the 0% to the 70% phase, respectively. Cardiac substructure abbreviations are defined in the text.

Discussion

By leveraging multi-phase 4DCT data, this work sought to quantify the intra-fractional displacement of sensitive cardiac substructures throughout the respiratory cycle. Automatic segmentations were generated on the 50% phase of the 4DCT using previously validated multi-atlas and deep learning methods[48,204]. DIR was used to propagate physician delineated ground truth segmentations of 12 cardiac substructures and the whole heart to all other 4DCT phases. Intra-fraction centroid displacements over the 4DCT data were analyzed to determine the dominant axes of excursion.

While other studies have evaluated intra-fraction heart displacement, our work helps to further quantify the primary axis of the displacement by quantitatively defining the excursion of the substructures contained within the heart. The work completed in this study is similar to that of a study by Guzhva *et al.* where they utilized 4DCT data with contour propagation in order to show the displacement experienced by the heart and its substructures due to respiration[50]. Guzhva *et al.* also propagated from the 50% phase to the other phases and then manually revised segmentations for final analysis[50]. They found that vector intra-fractional displacement of the cardiac substructures ranged from 7 to 15 mm and was dominantly in the S-I axis[50]. Similarly, our work agrees with those findings in that substructure excursion from respiration was predominantly in the S-I axis, and maximal vector displacements ranged from 5 to 10 mm. Our study improved on the work conducted by Guzhva *et al.* by considering the LMCA and the LADA as separate cardiac substructures and through the consideration of the great vessels (i.e. SVC, PA, and AA)[50]. The current study also considered radiation dose and cardiac substructure

volume at different phases in respiration. However, both studies were limited as neither accounted for inter-observer contouring variability.

With regard to the volume comparison between substructure volumes at EE and EI, it was found that the whole heart volume had small insignificant changes (~1%), which parallels results by Yue *et al.*[203]. Moreover, the dosimetric analysis in this work revealed that, although changes in D_{mean} and D_{max} for the whole heart were less sensitive to respiration (< 0.5 Gy), large dose differences for individual substructures were experienced. This study also found that cardiac substructures towards the superior extent of the heart, the great vessels (i.e. AA, SVC, and PA), had the smallest displacements in each axis, where substructures at the inferior extent of the heart, the RCA and the IVC, had the largest displacements. Limited data are available for direct comparison, however it has been reported that tumor excursion in caudal lung lobes displace the most over the respiratory cycle[211]. Additionally, in a study by Wang *et al.*, their 4DCT data showed that lung tumors with close proximity to the diaphragm experienced the most respiratory motion[212]. In the present study, the largest substructure centroid displacements occurred for the IVC, which is located at the inferior aspect of the heart, and passes through the diaphragm at the vena caval foramen[213].

One limitation of this work is that the intra-fraction motion of cardiac substructures occurred in free breathing respiration and not under DIBH conditions. There have been abundant studies confirming that DIBH reduced cardiotoxicity risk[214]. However, it has also been shown that this reduction in risk may be accompanied by large inter-fraction setup errors[215-217]. The limitation of uncertainties associated with the DIR process (as outlined by American Association of Physicist in Medicine Task Group report number 132[218]) were

mitigated as segmentations at each phase of respiration were manually verified and corrected. Another limitation is that isolating cardiac excursion was not possible with respiratory-correlated 4DCTs, and thus, the excursions presented in this work represented a contribution from both respiratory and cardiac motion. However, it has been previously reported by Tan *et al.* that the LV and coronary arteries are the most mobile cardiac substructures through the cardiac cycle, displacing between 3-8 mm between end-diastolic and end-systolic phases in three dimensions[52]. Thus, cardiac motion may be managed through incorporation in future planning organ at risk volume design. This work may also be limited in that only four phases of the 4DCT were used instead of 10. However, in a recent 10-phase 4DCT study, there was no mention that all 10 phases were required or were at all advantageous over 4-phase 4DCT[50]. Additionally, our study presents statistics on which phases the maximum centroid displacement occurred between, which is a unique contribution.

This work included both breast and lung cancer patients, which may have contributed to differing dominant axes of motion from cardiac and respiration influences. Guzhva *et al.* did find that patients with Hodgkin's lymphoma had a tendency to have larger displacements in the S-I axis than patients with lung cancer, which could be due to the comorbidities accompanying smoking or otherwise compromised lung function (i.e. a hyperinflated lung)[50]. Therefore, inconsistencies in patient anatomy could also cause uncertainty in determining the dominant axes of excursion and may be circumvented through expanding the patient cohort, or grouping by disease site. Nevertheless, this work was done in order to validate the need for consideration of cardiac substructures through

the incorporation of a motion model. This would provide an opportunity to decrease cardiotoxicity risk during radiotherapy treatment.

Conclusion

This work characterized the independent intra-fraction displacement of the cardiac substructures through the respiratory cycle. This work has importance for possible cardiac substructure PRV generation for patients who are unable to comply with breath-hold conditions for thoracic cancer treatments. Future work to determine the dosimetric effect of sensitive cardiac substructure displacement in respiration is warranted.

Part 2 "Inter-Fraction Cardiac Substructure Displacement Assessed Via MR-Guided Radiation Therapy"

Introduction

RT doses to the heart are strongly linked to cardiac toxicities such as coronary heart disease, heart failure, and even cardiac death[15,16]. Cardiotoxicity is often reported for breast, lung, and esophageal cancers as well as Hodgkin's disease[5,6,8]. In RTOG 0617[26] evaluating dose escalation for locally advanced NSCLC, volumes of the heart receiving ≥ 5 and ≥ 30 Gy were independent predictors of a patients' quality of life[27] and overall survival[26]. Yet, the heart is complex and dose to substructures (e.g., coronary arteries, ventricles, atria, great vessels, etc.) contained within the heart have been strongly linked to radiation-induced cardiac morbidity[30] and future acute coronary events[28,219]. Thus, recent attention has been focused toward local radiation dose deposition to substructures contained within the heart. Sub-analysis of RTOG 0617 revealed that atrial, ventricular, and pericardial doses showed a stronger association with overall survival than using standard whole heart dose metrics[40,41,220]. Furthermore, an association has been found between the incidence of coronary stenosis to the radiation dose received by the coronary arteries[221], thus underscoring the need for cardiac substructure-specific dose assessment.

However, one significant challenge with assessing dose to cardiac substructures is that they are difficult to discern on non-contrast treatment planning CTs and not typically considered in the treatment planning process[18] due to their limited soft tissue contrast as shown in Figure 26 (center). MRI, on the other hand, substantially improves the visibility of cardiac substructures (Figure 26, right)[43,44]. Magnetic resonance-guided RT (MRgRT) offers significant advantages compared to x-ray based technologies for delineation,

localization, tumor tracking, and adaptive radiation therapy. Tumor and OAR visualization using 0.35 T MRgRT has been shown to be superior to cone-beam computed tomography (CBCT)[222]. MRgRT allows for simultaneous tracking to monitor intra-fraction motion during treatment delivery[223] while avoiding radiation exposure due to the continuous imaging.

In a prospective Phase 1 trial for MR-guided adaptive radiation therapy for ultra-central lung cancer, the proximity of the lesion to the heart triggered plan adaptation for multiple treatment fractions, suggesting that inter-fraction displacement of the heart may be substantial[224]. Prior studies have shown that the average inter-fraction displacements of the whole heart and the LADA are typically < 7 mm in each orthogonal direction with the S-I displacement typically the greatest due to diaphragm motion[49,225,226]. To date, limited data are available to quantify inter-fraction displacement of other substructures other than the LADA. This study sought to leverage longitudinal MRgRT data to quantify inter-fraction displacements of 12 cardiac substructures to facilitate safety margin design.

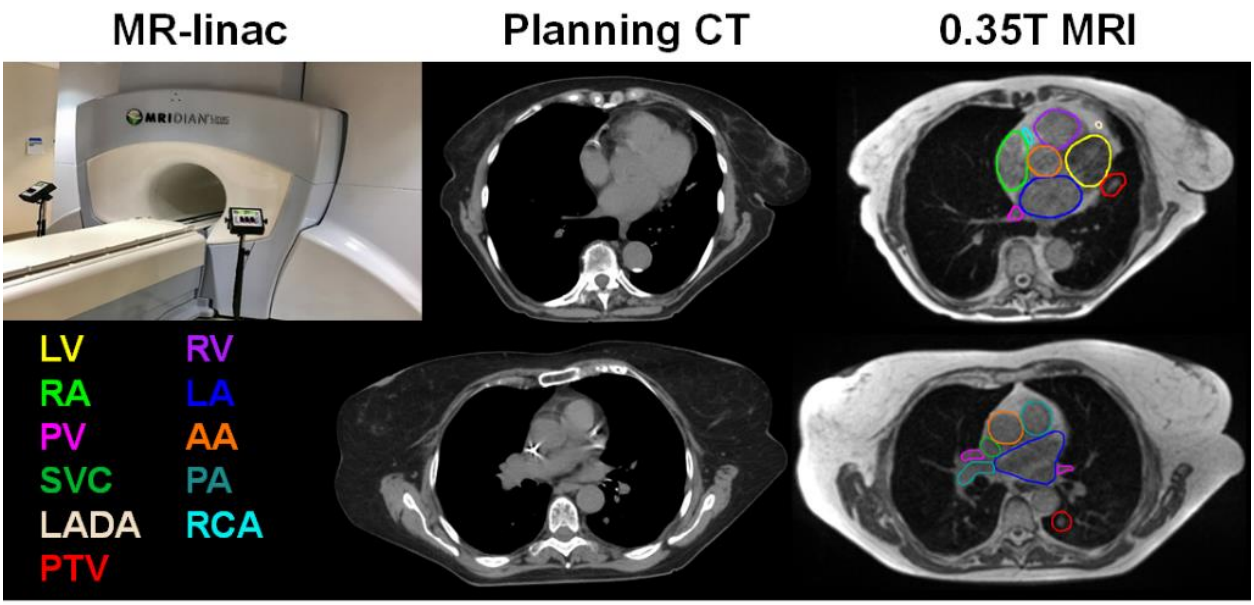


Figure 26: (Top Left) ViewRay MR-linear accelerator, and two patient examples showing: (Middle) Treatment Planning CT that is low contrast and does not show sensitive cardiac substructures, (Right) 0.35 Tesla MR dataset with cardiac substructure contours evident and delineated via deep learning-based segmentation. Abbreviations defined in the text.

Methods

Patient Methods for Low-field MRI

After obtaining Institutional Review Board approval, we retrospectively analyzed radiotherapy treatment data for 20 patients who underwent daily MRgRT using a ViewRay MRIdian 0.35T MR-linac (ViewRay, Mountain View, CA) (Figure 26, top left). Eleven patients had lung masses (64% metastatic and 36% primary bronchogenic malignancies), five had mediastinal and chest wall lesions, and the remaining four patients had liver tumors (i.e. hepatocellular carcinoma). SBRT was prescribed to 75% of the study cohort (4-5 fractions/patient) and the remaining were treated with conventionally fractionated IMRT (14-30 fractions/patient). All cases underwent 0.35 T MRI simulation (MR-SIM) on the MR-linac in the same respiratory condition used for treatment (11 imaged in EE, 7 in

EI, and two in free breathing). Breath-hold scans were conducted using a 17-25 second acquisition (1.5×1.5×3.0 mm^3 resolution) while the free breathing scans utilized a 3-minute acquisition (1.5×1.5×1.5 mm^3 resolution). For each image acquisition, a b-SSFP sequence was utilized (TrueFISP, Siemens, MAGNETOM Avanto, Syngo MR B19). TrueFISP is commonly used in cardiac imaging due to its high signal-to-noise ratio and imperviousness to motion artifacts[227,228]. The first 3-4 daily MR scans for all cases were evaluated for cardiac substructure displacement (total = 79 fractions).

Cardiac Substructure Segmentation

Inter-fraction motion was assessed for 12 cardiac substructures including the left/right atria (LA, RA), ventricles (LV, RV), superior/inferior venae cavae (SVC, IVC), ascending aorta (AA), pulmonary artery/veins (PA, PV), LADA, right coronary artery (RCA), and left main coronary artery (LMCA). Of the 20 total patients studied, initial cardiac substructure segmentations were generated on MR-SIM datasets using a previously validated cardiac substructure segmentation atlas[48] for 11 patients. For the remaining 9 patients, cardiac substructures were automatically generated using a three-dimensional deep learning U-Net[204] that was developed at a later date. The deep learning U-Net was implemented as it yielded reductions in substructure generation time and improved segmentation accuracy as compared to the atlas method. After the substructures were segmented on the initial MR-SIM image, the outputted segmentations were validated by one of two radiation oncologists. Contours underwent final verification by the more experienced of the two radiation oncologists with manual modifications made as needed to ensure clinically viable segmentations were rendered regardless of the initial segmentation approach.

A commercially available intensity-based free-form DIR algorithm (MIM Software, Cleveland, OH) was used to propagate contours from the MR-SIM image to each daily MRI, yielding a total of 3-4 registrations per patient (79 unique daily MRIs across all patients). Final propagated contours were again verified by one of two radiation oncologists and corrected as needed. In order to assess the inter-fraction substructure displacement due to daily patient positioning, a final translation-only rigid registration between the MR-SIM image and each daily MRI was performed by a physicist with an emphasis on aligning the PTVs. Figure 27 displays an example of substructure variation in position at breath-hold between MR-SIM and daily fractions for a representative patient.

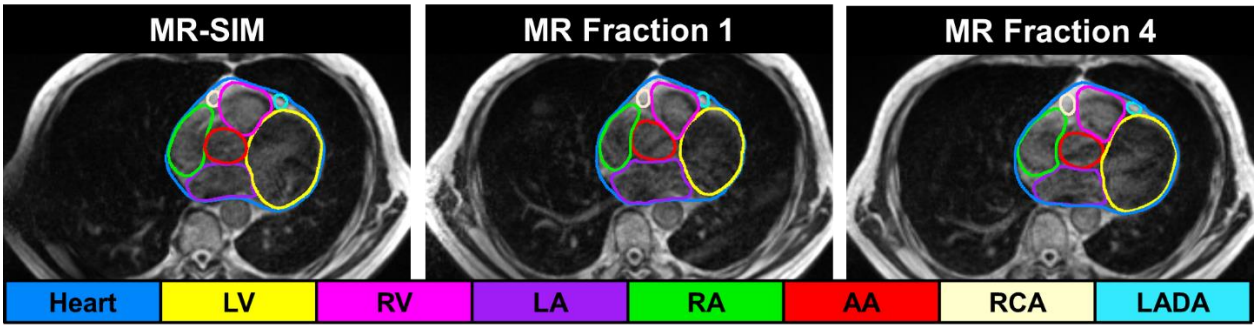


Figure 27: Example of substructure variation in position at breath-hold between MR simulation (MR-SIM, left), daily fraction 1 (center), and daily fraction 4 (right) for a representative patient. Substructure abbreviations are defined in the text.

Statistical Analysis and Data Extraction

A MIM workflow was developed to export centroid and volume information for the substructures after tumor-based rigid registrations. Inter-fraction differences for each cardiac substructure were quantified via centroid analysis in each cardinal direction, as well as in vector displacement. A two-tailed Wilcoxon signed ranks test was used to test for statistically significant differences in volume between the MR-SIM and each daily MRI,

for each substructure. A *P*-value less than 0.05 was considered statistically significant. All statistical analyses were conducted using SPSS version 26.0 (SPSS, Chicago, IL).

Planning Organ at Risk Volume Generation

ICRU 62 recommends the inclusion of a margin for OARs for the consideration of geometric uncertainties in RT[229,230]. This margin, called the planning organ at risk volume (PRV), is a measure to account for both systematic and random uncertainties in the radiation therapy process[230]. Random errors in RT can cause the dose distribution to blur by displacing high/low isodose lines closer/further from the region of interest[230]. Systematic errors displace the entire dose distribution closer or further from the region of interest[230]. Using the MR-SIM as the reference, mean centroid displacements and SD for each patient and substructure were calculated. The systematic error (Σ) was calculated by taking the SD of the mean displacement and the random error (σ) by calculating the root-mean-square of the SD[53]. Of note is that a measure of uncertainty induced by breathing is not included in this calculation as 18 out of 20 patients were treated in breath-hold conditions[231]. The PRV was calculated to accommodate daily setup variations based on a previous study by McKenzie *et al.*[230] with coefficients selected where the PRV maximum dose does not exceed the OAR maximum dose in 90% of cases[230,231].

$$PRV = 1.3 * \Sigma + 0.5 * \sigma \quad (13)$$

Results

Patient Population Results

Across the heart and substructures, inter-fraction displacements for 18.5% (L-R), 17.4% (A-P), and 23.1% (S-I) fractions were greater than 5 mm. Fewer than 3.7% of all structures displaced at least 10 mm in any direction over the studied fractions, and these

were often due to lack of compliance with breath-hold conditions. Table 6 summarizes the maximum excursions across the patient population for all cardiac substructures in the L-R, A-P, S-I, and vector displacements.

Structure	L-R	A-P	S-I	Vector
Heart (average ± SD mm)	2.5 ± 1.6	1.6 ± 1.7	3.1 ± 2.3	4.8 ± 2.4
% within 10 mm	100.0	100.0	100.0	94.3
% within 5 mm	91.4	91.4	85.7	57.1
Left/Right Ventricles	3.0 ± 2.8	2.9 ± 3.0	2.9 ± 2.5	5.9 ± 3.7
	96.2	96.8	97.5	89.2
	83.5	82.3	80.4	48.7
Left/Right Atria	3.0 ± 2.5	2.8 ± 2.8	3.2 ± 2.9	5.9 ± 3.8
	96.8	96.8	96.2	88.6
	82.3	82.9	81	46.8
Great Vessels	2.7 ± 2.4	2.3 ± 2.4	3.1 ± 2.5	5.5 ± 3.2
	97.5	98.3	98.3	92.4
	87.8	86.9	80.2	51.9
Left Anterior Descending Artery	3.6 ± 3.5	4.7 ± 4.4	3.9 ± 3.4	8.2 ± 5.1
	94.9	89.9	94.9	74.7
	75.9	65.8	65.8	30.4
Right Coronary Artery	3.7 ± 3.0	3.7 ± 3.3	4.2 ± 3.1	7.5 ± 4.3
	96.2	93.7	93.7	74.7
	69.6	73.4	65.8	36.7
Left Main Coronary Artery	4.1 ± 4.2	2.8 ± 2.8	3.4 ± 3.0	6.9 ± 4.7
	91.1	96.2	94.9	84.8
	78.5	87.3	74.7	43.0

Table 6: Average displacement for heart substructures for all studied MRI guided radiation therapy fractions with respect to the MRI simulation. Abbreviations Left-Right (L-R), Anterior-Posterior (A-P), Superior-Inferior (S-I).

For the chambers, the median absolute displacements were 2.4, 1.8, and 2.4 mm in the L-R, A-P, and S-I directions, respectively. The RCA shifted similarly in all axes (median shifts 3.3-3.9 mm) whereas the LADA had the highest A-P, S-I, and vector shifts of all substructures evaluated. The great vessels (i.e. SVC, PA, and AA) showed a tendency to have larger displacements in the S-I direction, with 44.7% of shifts being

greater than 3 mm, whereas only 35.4 and 26.6% of displacements were observed in the L-R and A-P directions, respectively. Larger S-I displacements likely reflect the larger axial MRI slice thickness (3 mm vs. 1.5 mm) for most (18/20) patients.

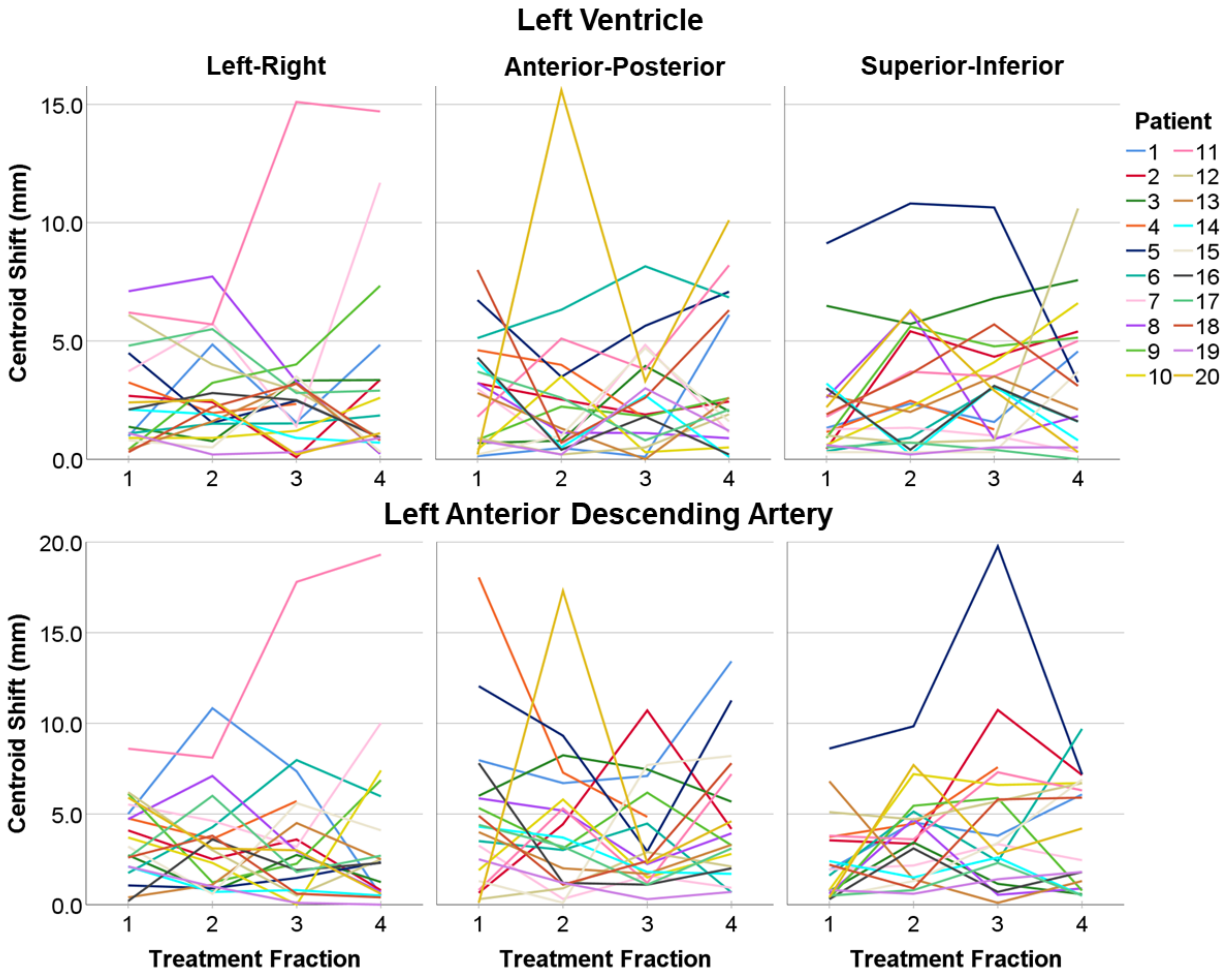


Figure 28: (Top row) Left ventricle and left anterior descending artery (bottom row) displacement across all treatment fractions with respect to positioning at MR simulation across each cardinal axis.

Figure 28 shows the absolute centroid shifts from the MR-SIM for the LV (top) and LADA (bottom) across all four treatment fractions and in the L-R (Figure 28, left), A-P (Figure 28, center), and S-I (Figure 28, right) axis. Patients 7 and 11 exhibited the largest shifts (greater than 10 mm) in the L-R axis for the LV. Patient 7, who underwent SBRT

for an enlarged mediastinal lymph node is shown in the top row of Figure 29. Figure 29 shows the MR-SIM image (left) and the fraction 4 MR image (right) localized to the centroid of the LV. Note the marked movement of LV, RV, LADA, RCA and RA in contrast to the heart that encountered minimal movement. Substructure delineations from each scan are represented on both images and reveal substructure shifts after a translation-only rigid registration to align the PTVs. Similarly, Patient 11 exhibited the largest displacement of the LV between MR-SIM and the fourth treatment fraction in the L-R axis. Patient 11 underwent RT for a malignant neoplasm of the lower left lung lobe and is represented in the bottom row of Figure 29.

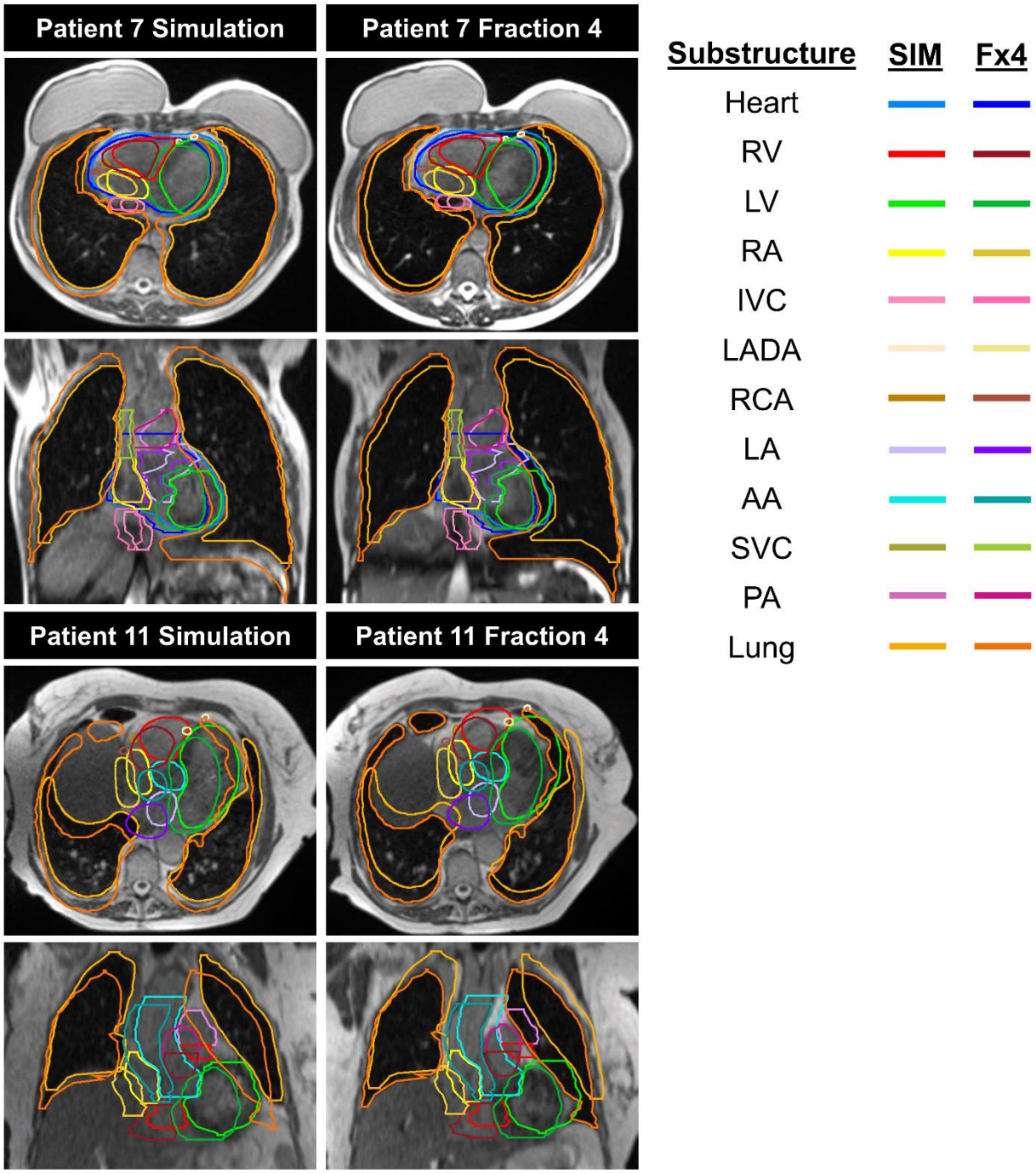


Figure 29: Patients 7 (top two rows) and 11 (bottom two rows) who experienced large left-right shifts between MR-simulation (left) and daily treatment (right). For each patient, both axial and coronal views are displayed for both the MR-simulation and the daily treatment. The lung contours highlight the lack of breath-hold compliance.

While most LADA shifts shown in Figure 28 were less than 10 mm, fraction 3 for patient 6 had a LADA centroid shift greater than 18 mm. Patient 6 received SBRT for a pulmonary nodule and is represented in Figure 30. Figure 30 shows both the axial (top row) and sagittal (bottom row) axes at the centroid of the LADA with the planned treatment dose. Note the marked movement of RV, LV, and LADA displacement, particularly in the inferior direction, moving the substructures further away from the high dose region. Substructure delineations from both the MR-SIM and fraction 3 images are represented on both images and represent substructure shifts after a translation-only rigid registration to align the PTVs.

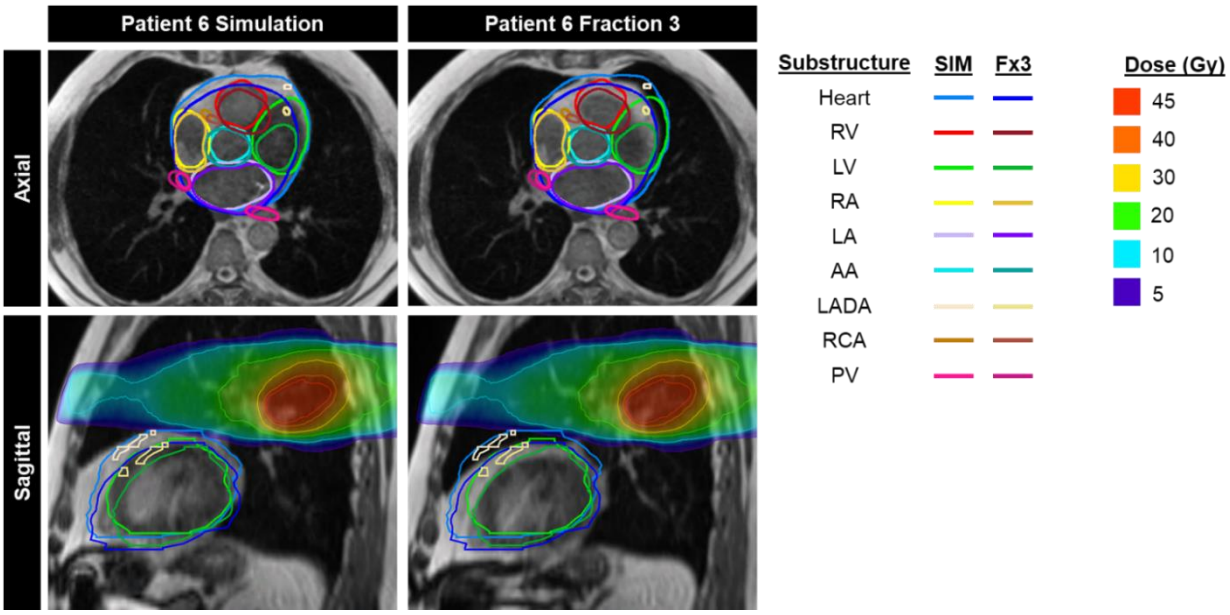


Figure 30: Displacement of cardiac substructures and planned dose between the 0.35 T MR simulation on axial (top row) and sagittal (bottom row) axes compared to fraction 3 0.35 T MRI for Patient 6 undergoing stereotactic body radiation therapy for a pulmonary nodule. Substructure abbreviations defined in the text.

Patient 20 received SBRT for an anterior liver dome hepatocellular carcinoma and is represented in Figure 31. SBRT localization at our institution includes an initial bony alignment conducted at breath-hold and then breath-hold compliance is assessed using

landmarks such as the dome of the diaphragm or liver. Once confirmed, a soft tissue match is then conducted in the tumor region. In the case of Patient 20, Fraction 2, the clinical net soft tissue correction from bony alignment for this case was 1.35 cm in the lateral direction. Figure 31 shows the displacement of select cardiac substructures between the MR-SIM (Figure 31, top left) and the second treatment fraction (Figure 31, top right) along with the planned dose. Figure 31 also highlights a DVH illustrating the expected dose to cardiac substructures at both timepoints. As can be seen in Figure 28, Patient 20 had the largest centroid displacement for the LV over any other displacement in the A-P axis. Additionally, the second fraction of this patient's treatment, shown in Figure 31, had centroid displacements for the LV greater than any other fraction for this patient by more than 5 mm. The DVH shown in Figure 31 shows that after the alignment of the treatment target shown in red, the substructures in the fraction 2 placement received additional radiation dose. More specifically, the mean dose to the IVC increased by 3.6 Gy when comparing MR-SIM planning dose to the second treatment fraction.

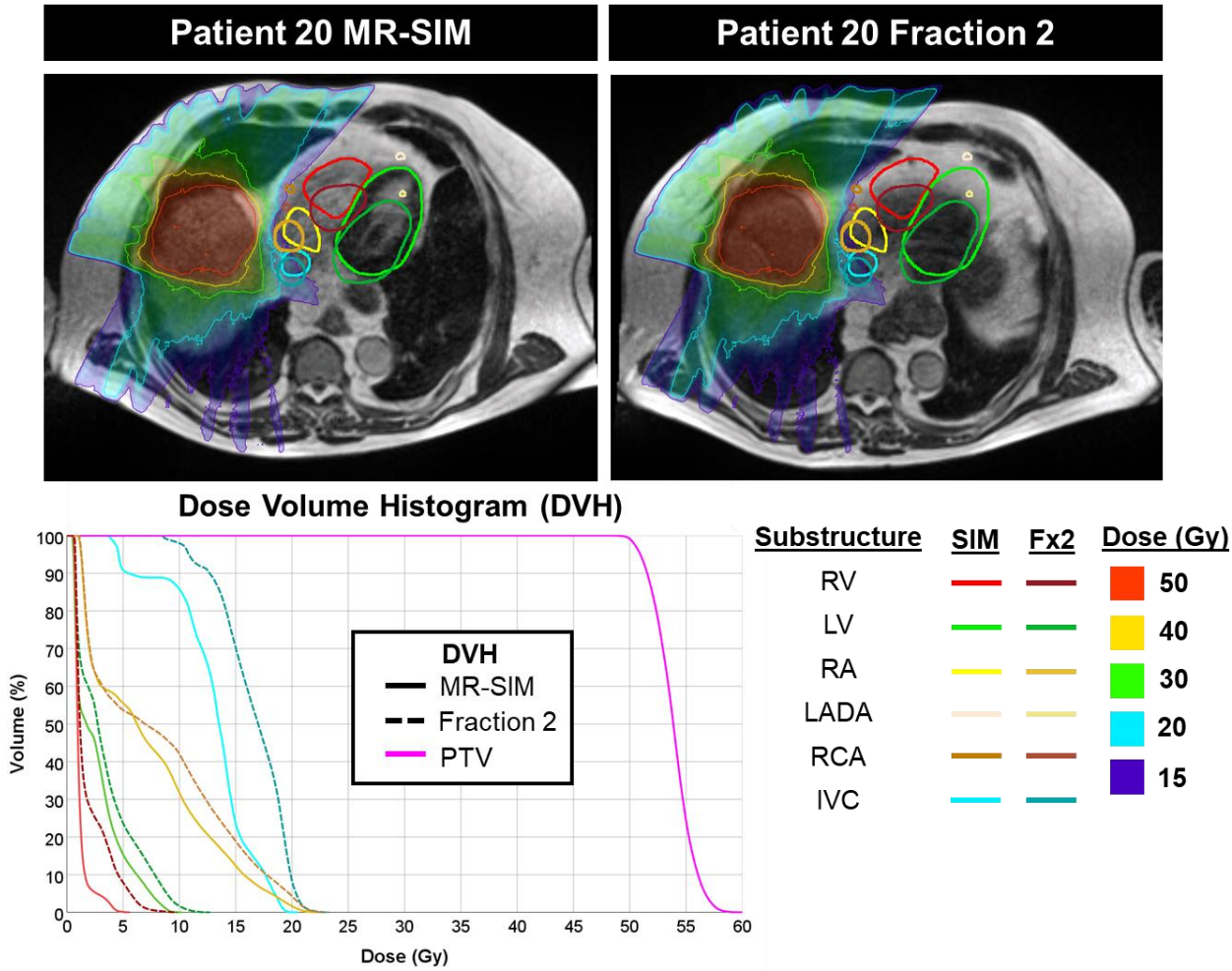


Figure 31: Top Row: Displacement of cardiac substructures and planned dose between MR-simulation (MR-SIM) in an axial view compared to fraction 2 MR for Patient 20 receiving stereotactic body radiation therapy for anterior liver dome hepatocellular carcinoma. Bottom: Dose Volume Histogram (DVH) showing planning dose to cardiac substructures at both timepoints. Substructure abbreviations defined in the text.

Planning Organ at Risk Volume Calculation

Table 7 shows the calculated systematic and random errors across cardiac substructures, as well as the PRV.

	Systematic Error (Σ)			Random Error (σ)			PRV (mm)		
Substructure	**L-R**	**A-P**	**S-I**	**L-R**	**A-P**	**S-I**	**L-R**	**A-P**	**S-I**
LV	2.19	2.25	1.97	2.16	2.30	2.07	4	4	4
LA	2.29	2.24	2.72	1.79	2.05	1.91	4	4	4
RV	2.16	1.87	1.70	1.89	1.97	1.92	4	3	3

RA	1.86	2.28	2.04	1.64	2.14	2.23	3	4	4
PV	2.17	1.67	1.66	2.33	1.92	2.13	4	3	3
PA	1.91	1.46	1.81	1.98	2.02	1.86	3	3	3
AA	2.12	1.74	1.93	1.60	1.88	1.98	4	3	3
SVC	1.76	2.19	1.61	1.62	1.72	2.29	3	4	3
IVC	1.27	1.88	2.67	2.20	2.12	2.92	3	4	5
RCA	2.37	2.32	2.33	2.17	2.81	2.43	4	4	4
LADA	2.74	2.53	2.69	2.51	3.23	2.60	5	5	5
LMCA	3.11	1.86	2.31	3.24	2.39	2.23	6	4	4

Table 7: Systematic (left) and random error (center) used to calculate the planning organ at risk volume (PRV) (right) across 12 cardiac substructures for the population (n = 20)

As shown by Table 7, isotropic margins of 4 mm were determined for the LV, LA, and RCA. The largest isotropic margin of 5 mm was calculated for the LADA, while the LMCA had a 6 mm margin in the L-R axis. The great vessels (i.e. SVC, PA, AA) experienced the smallest PRVs with the majority of axes being 3 mm.

Discussion

By leveraging MRgRT, this work sought to quantify the inter-fraction displacement of sensitive cardiac substructures over the SBRT treatment course. Recent MRgRT advances and auto-segmentation work were utilized to accurately and efficiently delineate 12 cardiac substructures. Centroid shifts over unique longitudinal MRgRT data were analyzed and allowed for safety margin design.

The trends observed in this current work are largely consistent with those from prior studies. One study by Jagsi *et al.* investigated the inter-fraction reproducibility of the LADA of 10 patients who underwent adjuvant RT for breast cancer under active breathing control[49]. The displacements of the LADA from the planning CT scan to that from 11 treatment fractions were assessed at EE and DIBH states with spine-based image

registration[49]. They reported the long-term reproducibility, defined as the SD, of the LADA position to be 4.5 mm in the L-R, 3.3 mm in the A-P, and 6.0 mm in the S-I axis at EE, and 2.7 mm in the L-R, 3.4 mm in the A-P, and 6.8 mm in the S-I axis during DIBH[49]. In the current study, the displacements of the LADA were 3.6 ± 3.5 mm (L-R), 4.7 ± 4.4 mm (A-P), and 3.9 ± 3.4 mm (S-I) which are consistent with the values reported by Jagsi *et al.*[49]. Two prior studies investigated the inter-fraction displacement of the heart. One series by Alderliesten *et al.* included breast cancer patients who received adjuvant RT at DIBH[215]. The heart position relative to the breast surface, as captured by surface imaging with AlignRT system, was measured with a planning CT and daily CBCTs[215]. Based on data from 378 fractions of 20 patients, the displacement of the heart was 2.1 ± 2.0 mm (L-R), 0.8 ± 3.3 mm (A-P), and -2.2 ± 7.8 mm (S-I)[215]. In a similar study by Comsa *et al.*, the largest average "shift of heart position" (measured as the distance between the heart and the chest on daily CBCT images) at moderate DIBH in five breast cancer patients was reported to be 6.2 mm[226]. In comparison, our study reported a heart displacement of 2.5 ± 1.6 mm (L-R), 1.6 ± 1.7 mm (A-P), 3.1 ± 2.3 mm (S-I), and 4.8 ± 2.4 mm (vector), which agree with these prior studies. However, our study is unique as it is the first study to report inter-fraction displacements of several cardiac structures other than the heart and LADA.

PRVs for 12 cardiac substructures were calculated in this study. Li *et al.* utilized 20-phase electrocardiogram gated data and a reference of the end-systolic phase to determine PRVs for coronary arteries[53]. Using the method of margin calculation mentioned here (previously proposed by McKenzie *et al.* for small and/or serial organs[230]), they calculated a range of margins between 3-8 mm[53]. Similarly, Topolnjak *et*

al. studied the geometrical uncertainty of the heart using left-sided breast cancer patients[231]. They calculated PRVs from inter-fractional motion of the whole heart, derived from CBCT images, to be 1.6, 1.4, and 2.1 mm in the L-R, A-P, and S-I, respectively[231]. The values from these two studies can be compared to the 3-6 mm range of PRVs derived from the current study. The discrepancy from Li *et al.* is likely due to their coronary artery segmentations being standardized at 2 mm diameter, whereas our coronary artery segmentations were not standardized and were larger (~4-6 mm). Whereas, Topolnjak *et al.* had smaller PRV values as they only assessed the whole heart.

One limitation of our study is that the utilized MR images were not cardiac gated and therefore did not take into consideration cardiac motion. As the image acquisition time ranged from 17 seconds to 3 minutes, numerous cardiac cycles were captured throughout the course of imaging and therefore, the cardiac substructures are represented by their average position over the course of the scan. The cardiac motion captured in these scans may have presented challenges in identifying the coronary arteries as they can become indistinct and noncontiguous[141]. This was managed by consensus ground truth segmentations being generated by two radiation oncologists. Although cardiac motion can be on the order of 3-8 mm[52], it is currently not taken into consideration clinically as cine-angiography or echocardiography are required[232]. Yet, as the cardiac substructure excursions from this work are on the order of cardiac motion, the internal motion may be merged with the inter-fraction variability in a future safety margin design. Topolnjak *et al.* incorporated a term for the PRV calculation to accommodate respiratory motion which is not addressed in this work as 18 out of 20 patients were treated in breath-hold[231].

Another limitation of this work is that only 0.35 T TrueFISP MRgRT datasets acquired at breath-hold were evaluated. To translate the work to other field strengths such as the 1.5 T offered with the Unity MR-linac (Elekta AB, Stockholm, Sweden), the deep learning segmentation pipeline would need to be retrained for the higher field MRI datasets for the MRI sequence of interest. Additionally, the increased signal-to-noise ratio at higher field strengths may provide an improvement to automatic segmentation accuracy. Nevertheless, the inter-fraction substructure displacements quantified at a low MRI field strength under breath-hold conditions may be applied regardless of field strength.

As current studies are considering cardiac substructures in RT, it is crucial that inter-fraction variation is considered. Patel *et al.* has recently recommended that the LADA is incorporated as a cardiac substructure to avoid during the RT planning process due, as radiation dose to it has been strongly tied to coronary artery calcification[32]. A study by Aldridge *et al.* found that the integration of cardiac substructures into the re-optimization of retrospective thoracic RT plans drastically reduced radiation dose to sensitive cardiac substructures[233]. This reduction was achieved with a negligible increase in plan complexity while maintaining PTV coverage and clinical endpoints for other critical OARs[233]. Examples of cardiac substructure sparing through VMAT was shown by Ferris *et al.* who achieved significant improvements in mean dose to the chambers, great vessels, and coronary arteries[234]. These studies outline the importance in considering cardiac substructures in planning which will require adequate setup and motion margin consideration.

The increased presence of RT indications for peri-cardiac tumors, especially metastatic lesions, mandate the development of realistic cardiac sparing margins. Anisotropic cardiac substructure-specific planning margins for OARs may be warranted to accommodate differential inter-fractional shifts. While these results suggest that these margins may need to be patient-specific, more precise margin definition will require further confirmation in a larger cohort stratified by respiratory status and accounting for systematic and random uncertainties.

Conclusion

This exploratory work quantified the inter-fraction displacement of critical cardiac substructures and is a first step in deriving substructure-specific safety margins to ensure highly effective cardiac sparing. Individual cardiac substructure displacement demonstrated variability in magnitude and dominant axis, suggesting that anisotropic substructure-specific PRVs may be warranted. These findings require validation in a larger cohort for applications in prospective clinical trials.

CHAPTER 6 "TREATMENT PLANNING COMPARISONS AND TRANSLATING TECHNOLOGIES TO AN MR-LINAC"

Part 1 "Incorporating Sensitive Cardiac Substructure Sparing into Radiation Therapy Planning"

Introduction

Cardiac toxicity is a major complication of cancer treatment and can occur during, shortly after, and even many years after treatment has been delivered. Long-term follow up of patients undergoing thoracic radiation, such as lymphoma, lung, breast, and esophageal cancers, has shown that in particular, RT can lead to radiation-induced cardiac toxicities such as congestive heart failure, pericardial effusion, coronary artery disease, and myocardial infarction[5-7].

Yet, when a patient's RT plan is created, only simple whole heart metrics (i.e. MHD) are routinely considered for cardiac risk assessment in the current standard of care. The QUANTEC report assesses dose to the heart as a whole and recommends less than 10% of it receives greater than 25 Gy (in 2 Gy fractions) to keep the conservatively estimated risk of long term cardiac mortality less than 1%[18]. Importantly, these whole-heart dose metrics do not provide any information on where dose is distributed.

The heart is a complex organ and dose to its substructures (e.g., coronary arteries, ventricles, atria, great vessels, etc.) have been strongly associated with radiation-induced cardiac morbidity[30] and future acute coronary events[28,219]. For example, dose to the LADA has been linked to an increased risk of myocardial infarction[31] and development of coronary artery calcifications[32]. Similarly, higher doses at the base of the heart (i.e. ascending aorta, superior vena cava, and pulmonary artery) are associated with lower rates of patient survival[35]. Importantly, recent RTOG 0617 sub-analyses suggest that

dose to the atrial and ventricular cardiac substructures are more strongly associated with survival than assessing dose/volume relationships to the entire heart volume[40,41,235]. In a recent study by van den Bogaard *et al.*[28], dose to the LV-V5 predicted major coronary events better than MHD. A study by Hoppe *et al.* highlighted the importance of quantifying substructure dose as the MHD becomes less correlated to substructure dose with increasingly conformal delivery[236]. Furthermore, a study by Jacob *et al.* outlines how the MHD does not accurately predict dose to the LV and coronary arteries[237].

To date, reducing dose to sensitive cardiac substructures has been severely limited because they are not readily visible on standard x-ray-based imaging used for both RT planning (i.e. CT-SIM) and RT delivery (i.e. CBCT). Thus, leveraging the superb soft tissue contrast of MRI may be advantageous, as MRI improves cardiac substructure visibility[43,44]. Furthermore, the recent introduction of MR-linacs (Figure 32, left) has yielded improved tumor and critical structure visualization at 0.35 T MRI as compared to CBCT[222]. MRgRT allows for continuous anatomical visualization of the patient's heart and target volume throughout treatment which may offer advantages for improved cardiac sparing. Therefore, to advance towards mitigating cardiotoxic side effects from RT, approaches for considering cardiac substructures during treatment planning are urgently needed.

This work sought to apply a multimodality workflow (treatment planning CTs coupled with low-field MR-linac MRIs) to integrate sensitive cardiac substructures into treatment planning. This multi-modality workflow allowed us to quantify potential dosimetric advantages for improved cardiac sparing through plan re-optimization and for cases that may benefit, beam angle modifications.

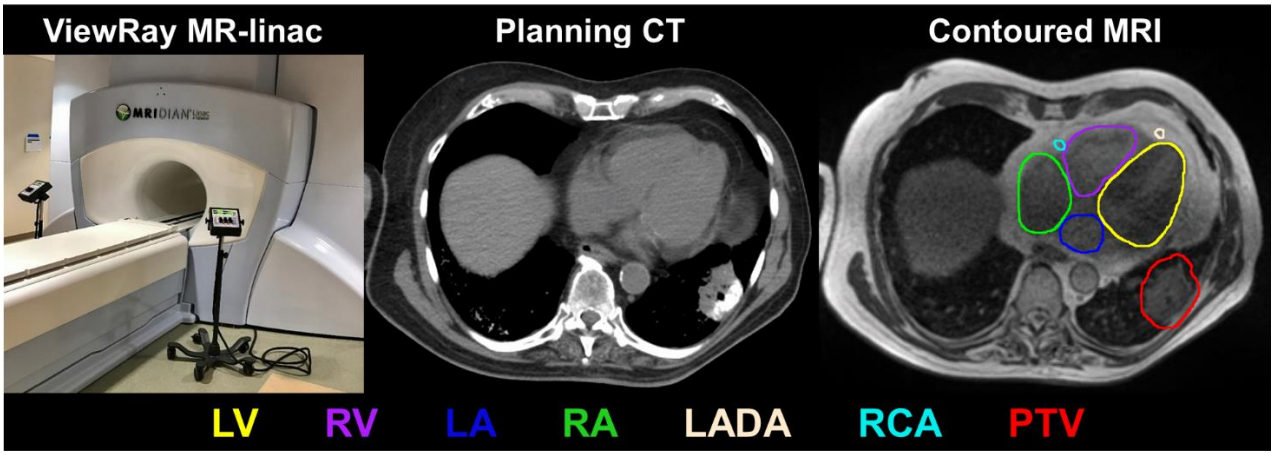


Figure 32: (Left) ViewRay 0.35T MR-linac, (Middle) treatment planning CT, (Right) 0.35T MR dataset with cardiac substructure contours evident and delineated. PTV: planning target volume (malignant neoplasm of lower left lung bronchus). Cardiac-related abbreviations are defined in the text.

Methods

Patient Cohort and Image Acquisition

Fifteen patients with 16 pericardial lesions (i.e. 16 individual plans) who underwent MRgRT for upper thoracic treatments of the lung, mediastinum, and esophagus were retrospectively reviewed on an Institutional Review Board approved study. Of these, 11 were treated with stereotactic body radiation therapy (3-5 fractions to a total dose of 30-50 Gy), 2 underwent conventional fractionation (25-35 fractions to a total dose of 50-70 Gy), and the remaining three were moderately hypo-fractionated (14-20 fractions to a total dose of 36-60 Gy). Patients were imaged in various breathing states (7 end-exhalation, 7 end-inhalation, 2 free-breathing) on a 0.35 T ViewRay MRIdian linear accelerator (ViewRay, Mountain View, CA).

All patients were imaged with a b-SSFP (i.e. TrueFISP) acquisition sequence (Siemens, MAGNETOM Avanto, Syngo MR B19) with 15/16 patients with mobile tumors undergoing daily 17-25 second MRIs ($1.5\times1.5\times3$ mm^3) under breath-hold conditions. One

patient with a left chest wall lesion could not tolerate breath-hold and thus underwent a 175 second free-breathing MRI for treatment planning. TrueFISP is commonly used in cardiac imaging due to high signal-to-noise ratio and imperviousness to motion artifacts[227,228]. All treatment planning was conducted and dose was calculated on a non-contrast CT-SIM in a manner similar to what has been reported for MRgRT of thoracic lesions[224]. All CT-SIMs were acquired on a Brilliance Big Bore CT Simulator (Philips Medical Systems, Cleveland, OH) with a 3 mm slice thickness. MR and CT-SIM sessions were conducted on the same day and patients were immobilized in the supine position using molded vacuum cushions.

Segmentation and Registration

Assessed cardiac substructures included the heart, left/right ventricles (LV, RV), atria (LA, RA), superior/inferior venae cavae (SVC, IVC), ascending aorta (AA), pulmonary artery/veins (PA, PV), left anterior descending artery (LADA), right coronary artery (RCA), and left main coronary artery (LMCA). For 11 patients, a cardiac substructure segmentation atlas[48] automatically generated the cardiac substructures on the CT-SIM dataset for treatment planning with the final contours displayed on the low-field MRI at Figure 32, right. For the remaining 5 patients evaluated at a later date, automatic cardiac substructure segmentation on the CT-SIM was performed using a three-dimensional U-Net[204], a deep learning model that improved the accuracy and substructure generation time as compared to the atlas method.

While automatic segmentation methods (i.e. multi-atlas and deep learning methods) provided initial substructure contours on the CT-SIM datasets, a radiation oncologist consulted the co-registered low field MRI to modify and confirm the final

contours used for treatment planning. As shown by the lack of contrast in the planning CT (Figure 32, center), the enhanced soft tissue contrast from the MRI assisted the generation of more reliable cardiac substructure delineations on the corresponding planning CT. Co-registration involved an automatic rigid registration based off a manually drawn, local, cardiac confined bounding box. Normalized mutual information was used as the similarity metric as it has been shown to accurately align multi-modality images[140].

Treatment Planning

For all patients, the CT-SIM was used as the primary image set for treatment planning as has been reported in the literature for MRgRT of thoracic lesions[224]. The co-registration of the low-field MR image to the CT-SIM to elucidate the cardiac substructures was a critical step in allowing the physician to verify the cardiac substructure auto-segmentations. Step-and-shoot IMRT planning was used to generate all 16 RT plans at a dose rate of 600 cGy/minute. The MR-linac utilizes a fast Monte Carlo dose calculation algorithm[238] and plans were calculated using a 1x1 mm dose grid with 1% dose uncertainty[239]. Plans were prescribed to 95% of the planning target volume with total doses for the original treatment plans varying from 30-70 Gy delivered in 4-35 fractions. The original treatment plans for all patients included clinical dose constraints for whole heart endpoints. Specifically, less than 15.0 cc of the whole heart was to receive 24-42 Gy and a maximum dose no greater than 30-40 Gy (ranges were dependent on prescription dose). All clinical treatment plans met physician objectives using standard QUANTEC[240,241] and TG-101[31] dosimetric endpoints for OARs.

Along with adding substructure segmentations retrospectively to the original clinical treatment plans for dose assessment, all plans were re-optimized to spare cardiac

substructures (SPARE plan). Strategies for substructure sparing included evaluating the original plan to identify which cardiac substructures were near the PTV and thus received the most dose. Optimization objectives were then added with increased priority on the substructures receiving higher doses. If the dose limit was unachievable, constraints were relaxed with the overall objective to minimize substructure dose. If the dose to a particular substructure was minimal in the original plan, an additional objective was added in the IMRT optimization to ensure consistency was maintained.

In addition to adding substructures to the optimization, possible further cardiac sparing improvement was also assessed through modifying the beam arrangement (New Angles plan) after the substructures had already been incorporated into the optimization. Each plan was evaluated for the presence of beams entering or exiting the heart that could potentially be removed or modified to further spare the heart and substructures. The proximity of the patient's lesion to the heart was evaluated as a potential reason for a patient benefiting from beam angle modification. IMRT techniques were used for all SPARE and New Angles plans with the substructures integrated into the optimization while maintaining tumor volume coverage and minimizing OAR dose. Table 8 outlines the dosimetric considerations during plan optimization, derived from the literature, when cardiac substructures were included. All plans were converted using the equivalent dose to 2 Gy fractions (EQD2, $\alpha/\beta = 2$) for evaluation.

Substructure	Mean Dose	Maximum Dose	Additional Endpoint
Right Ventricle	-	Minimize[42]	V45[42]
Left Ventricle	-	Minimize[42]	LV-V5[28] V45[42]
Left Atrium	8.5 Gy[35] Minimize[42,242]	Minimize[37]	V45[42]
Right Atrium	8.5 Gy[35] Minimize[42]	-	V45[42]
Superior Vena Cava	8.5 Gy[35]	-	D90[37]
PA, PV, AA	8.5 Gy[35]	-	
Left Anterior Descending Artery	Minimize[243]	< 10 Gy[32] Minimize[243]	V45[42]
RCA, LMCA	-	-	V45[42]

Table 8: Summary of cardiac substructure sparing utilized in planning optimization for the re-optimization (SPARE) plan and the New Angles plan. Abbreviations defined in the text.

Dosimetric and Statistical Assessment

Original, SPARE, and when applicable, New Angle plans were exported from the ViewRay planning system and imported into MIM (version 6.9.4, MIM Software Inc., Cleveland, OH) for automated evaluation. Dosimetric assessment included mean doses, LV-V5, and Dose to 0.03 cc ($D_{0.03cc}$, surrogate for maximum dose) for 12 cardiac substructures and the whole heart. To ensure clinical acceptable plans were still achieved, differences in PTV coverage and dose to the OARs were also assessed. Lastly, total MU and treatment time were evaluated and compared to the original clinical treatment plan as metrics of plan complexity. Dosimetric and planning data were summarized via mean ± SD. As the data was not normally distributed, dosimetric

comparisons for each metric were conducted using a 2-tailed Wilcoxon signed ranks test with $P < 0.05$ considered statistically significant. Statistical assessments were conducted in SPSS version 25.0 (SPSS, Chicago, IL).

Results

Contour Generation and Plan Complexity

The treatment time per fraction (a metric of plan complexity) across the 16 patients after plan re-optimization was 6.57 ± 3.50 minutes (range: 2.60 to 12.41) for the clinical treatment plan and was 6.93 ± 3.27 minutes (range: 2.75 to 11.99) after re-optimizing ($P > 0.05$). The percent difference between the original and re-optimized delivered MUs was 1.7 ± 11.3 % (range: -21.6 to 15.8%) which were not statistically different ($P > 0.05$).

Four patients benefited from New Angles plans where the number of original treatment beams (range: 7 to 11) shifted between -1 and +3 (range: 8 to 14). For two of the four patients, lesions were directly adjacent to the heart (i.e. a pericardial lymph node and a malignant neoplasm of the lung (Figure 37)). The other two patients presented with upper lung lobe lesions that were greater than 9 cm away from the heart. The average treatment time for these patients after beam angle modification was 6.12 ± 3.68 minutes which was not significantly different ($P > 0.05$) from the original treatment time for these 4 patients (6.54 ± 3.31 minutes). Lastly, the number of MUs for patients benefiting from beam angle modification were 9.5 ± 16.8% (range: -16.6 to 23.8%) different on average from the original plan ($P > 0.05$).

Cardiac Substructure Sparing

The radiation dose to the whole heart after plan re-optimization met all clinical objectives[240,241]. All sparing plans significantly reduced the MHD ($P < 0.05$) with an

average reduction of 0.7 ± 0.6 Gy (range: 0.1 to 2.5 Gy). Further, $D_{0.03cc}$ to the heart was reduced by 8.6 ± 12.1 Gy (range: -8.6 to 39.9 Gy) across all patients after plan re-optimization ($P < 0.05$).

Figure 33 outlines a subset of dose objectives from Table 8 representing the difference in radiation dose received by the LADA, LA, and LV between the original and clinical treatment plans across all 16 patients. Re-optimized SPARE plans reduced LADA mean and $D_{0.03cc}$ (0.0 to 63.9% and 0.0 to 17.3 Gy, respectively, Figure 33 left). For the 5 patients that had $LADA_{0.03cc}$ doses greater than 10 Gy (threshold for coronary artery calcification[32] presented in Table 1), 4 were brought below 10 Gy after re-optimization (average reduction for these patients was 13.4 ± 7.0 Gy). $D_{0.03cc}$ for the remaining patient was reduced from 29.0 to 11.2 Gy. Moreover, patient 6 (Figure 33) had a 4 Gy reduction in mean dose to the LADA after plan re-optimization. Similarly, $D_{0.03cc}$ to the LV was reduced in 14 cases (range: 0.05 to 12.85) with 10 patients having greater than 1.5 Gy reductions. A large reduction (> 7%) in LV-V5 was observed in 6 cases (Figure 33, right) and was significantly reduced over all patients ($P < 0.05$). LA mean dose (Figure 33, center) was either equivalent or reduced (average reduction 0.9 ± 1.2 Gy) for all SPARE plans. For Patient 3, the left atrial mean dose was reduced below 8.5 Gy which has been shown to be a threshold associated with decreased survival[35]. Lastly, the left atrial maximum dose that has been significantly associated with non-cancer death[37] was reduced by 2.3 ± 6.4 Gy across all 16 patients.

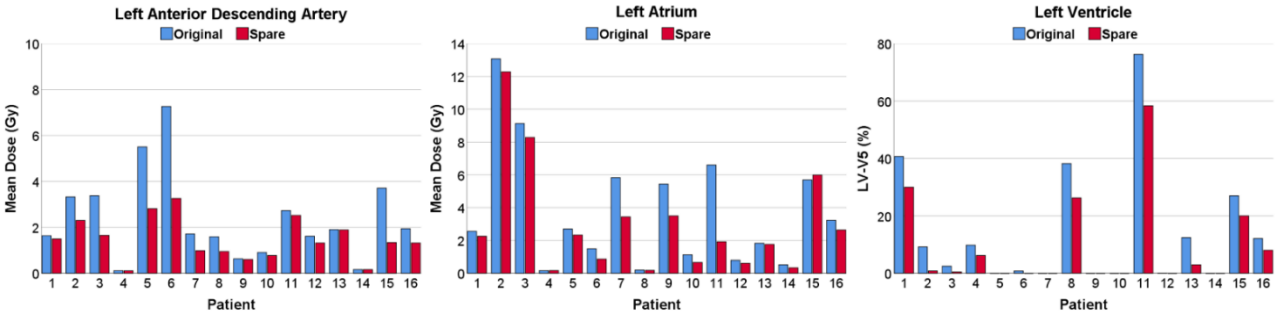


Figure 33: Dose sparing possible by incorporating cardiac substructures into IMRT optimization during MR-guided radiation therapy planning. The mean dose for all 16 patients is shown for the left anterior descending artery (left) and the left atrium (center). The left ventricular volume receiving 5 Gy (LV-V5) is shown on the right.

Table 9 summarizes the change in mean dose and $D_{0.03cc}$ to all cardiac substructures. The mean doses to all substructures and the heart were significantly reduced after re-optimization ($P < 0.05$). Moreover, $D_{0.03cc}$ was significantly reduced after plan re-optimization in 8/12 substructures, as well as for the whole heart. The LMCA mean and $D_{0.03cc}$ doses were reduced for all patients and was the substructure with the largest reduction in mean dose across all patients (average reduction in LMCA mean dose: 1.13 ± 1.15 Gy). Lastly, the volume of the heart receiving 25 Gy (V25) was significantly reduced on average (n = 12 patients who met the V25 threshold) by 1.08 ± 1.47% ($P < 0.05$).

Further cardiac substructure dose sparing beyond re-optimization was achieved for 4 patients with beam angle modification where the mean dose reduction across all substructures was 0.6 ± 0.4 Gy (highest reduction in PA of 1.5 ± 2.0 Gy). The $D_{0.03cc}$, mean dose, and V25 to the heart were further reduced by 5.4 ± 4.1 Gy, 0.5 ± 0.7 Gy, and 4.2 ± 2.9%, respectively. For the LV, after re-optimization coupled with beam angle modification, $D_{0.03cc}$ and LV-V5 were further reduced by 2.1 ± 2.9 Gy and 2.0 ± 1.9%, respectively. Lastly, the SVC D90 improved 3.3 ± 4.0% after the beam angles were modified.

Organs at Risk (OARs) and Planning Target Volume (PTV) Coverage

All re-optimized plans met the original clinical prescription dose to the PTV while doses to the OARs met all objectives based on clinically acceptable guidelines[240,241]. Table 9 outlines the average change in the mean dose and $D_{0.03cc}$ for the PTV and OARs. Across all patients, the esophagus had a negligible change in mean dose after plan re-optimization (0.25 ± 0.70 Gy, $P > 0.05$). Additionally, differences in clinical endpoints such as the volume of the lung receiving 20 Gy (V20) and volume of the esophagus receiving 35 Gy (V35) were negligible after re-optimization ($P > 0.05$). No statistically significant changes were observed in the mean dose, $D_{0.03cc}$, and other clinical endpoints for the PTV and OARs ($P > 0.05$).

		Average Change After Re-optimization	
Structure		**Mean Dose (Gy)**	**$D_{0.03cc}$ (Gy)** **Other Clinical Endpoint**
	PTV	0.37 ± 1.85	1.95 ± 3.67 **PTV95:** 0.03 ± 0.21 Gy
Organs at Risk	Spinal Cord	0.06 ± 0.25	0.38 ± 1.37
	Total Lung	-0.01 ± 0.26	1.32 ± 2.70 **V20:** 0.03 ± 0.71 %
	Esophagus	0.25 ± 0.70	0.39 ± 4.06 **V35:** 0.53 ± 2.46 % (n=5)
Heart and Substructures	Heart	-0.68 ± 0.60*	-8.57 ± 12.06* **V25:** -1.08 ± 1.47* % (n=11)
	LV	-0.53 ± 0.70*	-3.27 ± 4.08* **LV-V5:** -6.33 ± 5.57* % (n=12)
	LA	-0.85 ± 1.22*	-2.30 ± 6.42
	RV	-0.55 ± 0.74*	-4.12 ± 4.81*
	RA	-0.52 ± 0.94*	-1.38 ± 4.47
	AA	-0.83 ± 1.13*	-2.23 ± 3.42*
	PA	-0.95 ± 1.60*	-2.84 ± 8.53
	PV	-0.89 ± 1.09*	-2.71 ± 5.69*
	SVC	-0.57 ± 1.19*	-1.08 ± 3.65* **D90:** -0.10 ± 1.23 Gy*
	IVC	-0.16 ± 0.38*	-0.74 ± 1.91
	LADA	-0.91 ± 1.18*	-4.05 ± 5.32*
	LMCA	-1.13 ± 1.15*	-1.31 ± 1.55*
	RCA	-0.65 ± 1.26*	-1.64 ± 3.38*

Table 9: Change in $D_{0.03cc}$ and mean dose after plan re-optimization for the planning target volume (PTV), heart and its substructures, and other organs at risk. The asterisk indicates significant reduction in dose after re-optimization. N = 16 for all structures except for the esophagus where n = 10. For the heart V25, esophagus V35, and LV-V5, results were reported only for structures with a non-zero value for the corresponding dosimetric endpoint. There were no significant increases in dose after re-optimization. Abbreviations are defined in the text.

For the four patients that benefited from beam angle modification, negligible changes were observed for all of the PTV D95 metrics (range: 0 to -0.30 Gy) and 3 out of 4 patients' $D_{0.03cc}$ (< 0.5 Gy). However, one patient had an increase in $D_{0.03cc}$ of 3.7%, or 6.2 Gy, with beam angle modification when compared to the original clinical treatment

plan. Negligible changes (< 1%) in clinical endpoints were observed for the esophagus (V35 and V50) and lungs (mean dose and V20) as compared to the original clinical treatment plan while the spinal cord $D_{0.03cc}$ was reduced by 2.3 ± 1.9 Gy with beam angle modification as compared to re-optimization alone.

Individual Patient Results

Figure 34 shows DVHs for three patients selected to represent an example of the least effective cardiac substructure sparing (Patient 1), highly effective sparing (Patient 2), and an average case (Patient 13). Each DVH shows the PTV, involved OARs, and relevant cardiac substructures for both the original clinical treatment plan and the re-optimized plan. Patient 2 benefited from beam angle modifications, and thus, that plan is represented as well. Figure 34 highlights that for the patients shown, negligible differences (< 1 Gy) were observed for the mean lung dose and $D_{0.03cc}$ to the spinal cord indicating comparable plan quality was achieved even when cardiac substructure sparing was implemented. Radiation doses to the whole heart and total lung (results not shown for all patients) were reduced for all patients after re-optimization, with even further reductions after beam angles were modified. For patient 2, the mean esophageal dose decreased by 3.0 Gy from the original clinical plan and 4.5 Gy from the reoptimized plan after modifying the beam angles, all while reducing the mean.

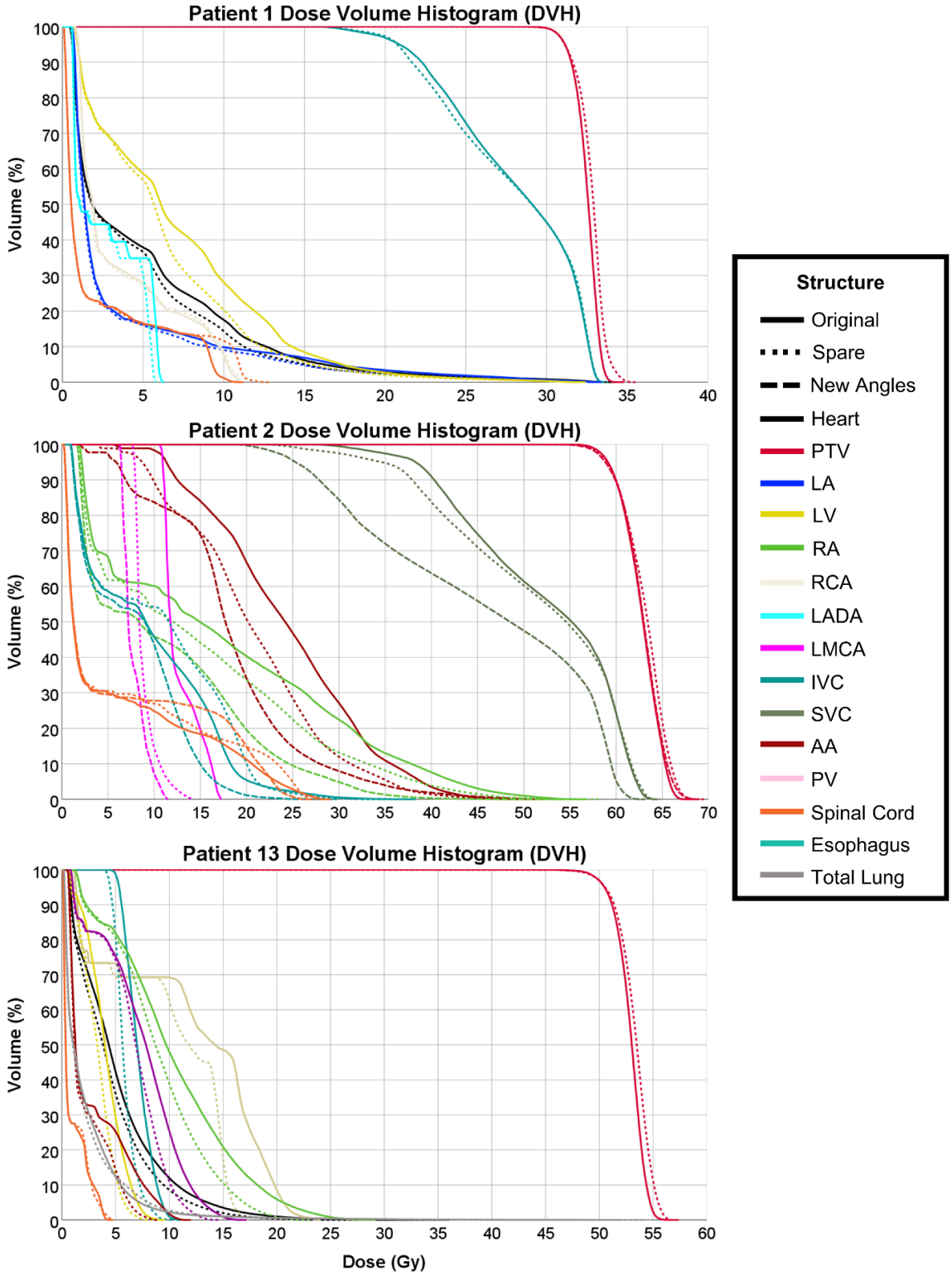


Figure 34: Dose volume histograms (DVH) for three patients of the least effective cardiac substructure sparing (Patient 1), highly effective sparing (Patient 2), and an average case (Patient 13) showing dose from the original clinical treatment plan and after re-optimization. The modified beam angle plan is also shown for Patient 2. Abbreviations defined in the text.

Figure 35 illustrates the initial clinical treatment plan of a malignant neoplasm of the lower left lung bronchus (left) treated to 48 Gy in 4 fractions and the corresponding cardiac SPARE treatment plan (right) for Patient 11. This figure highlights cardiac substructure sparing with greater than 10 Gy reductions in $D_{0.03cc}$ to the LV, LA, and PV. Specifically, note the removal of the 5 and 10 Gy isodose lines from many heart substructures (LA, AA, RA, PV, and RV) after the re-optimization.

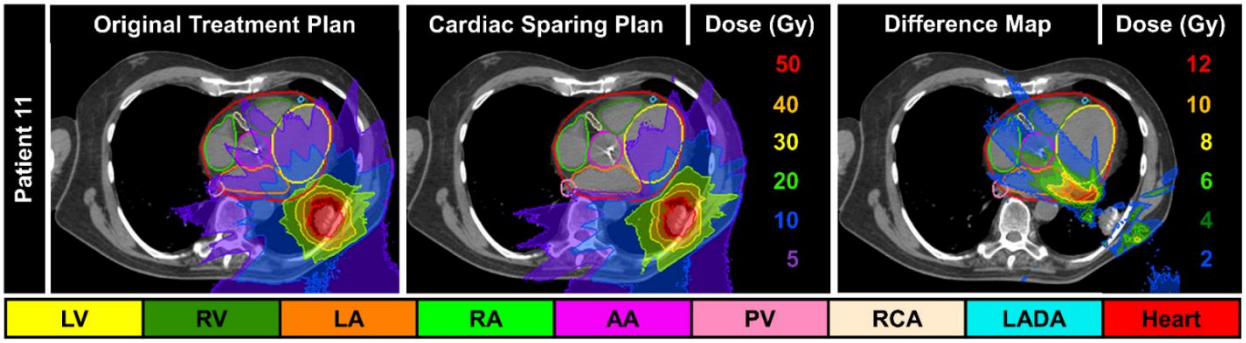
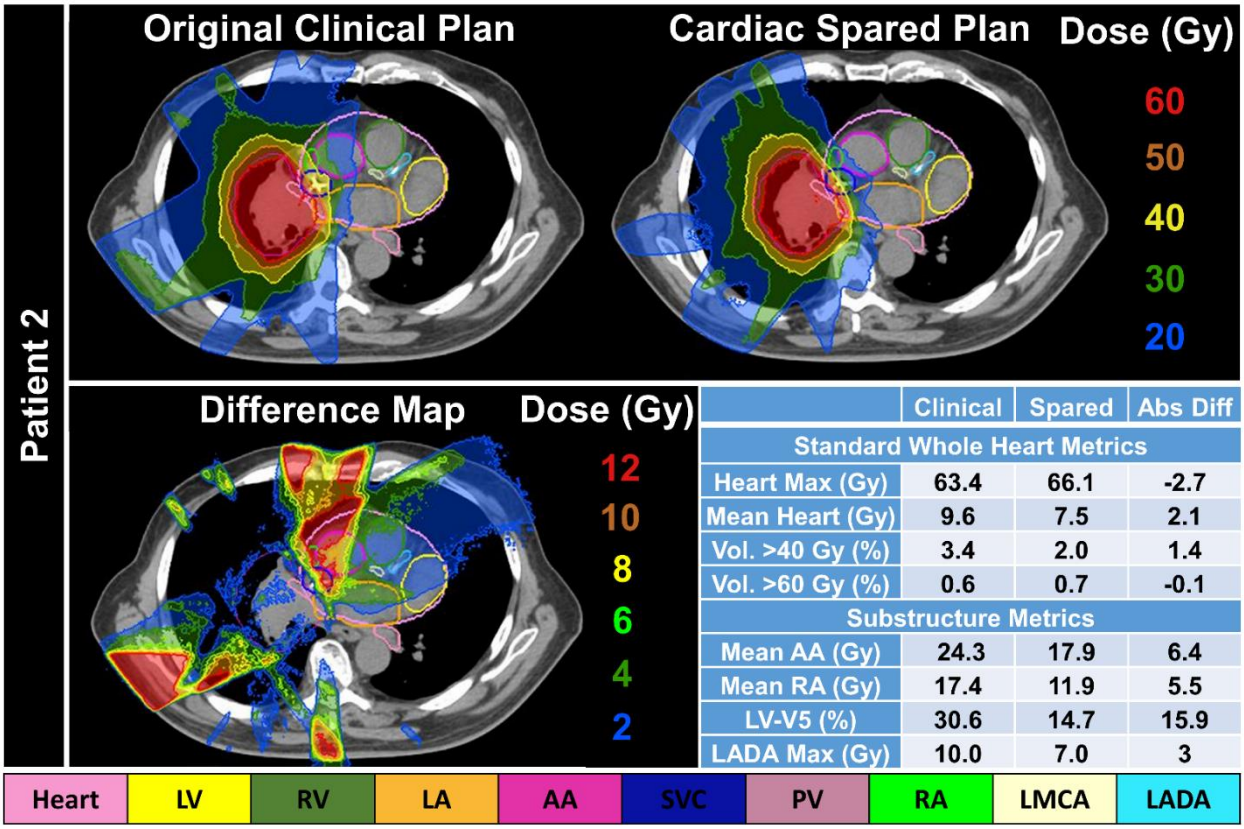


Figure 35: (Left) Initial clinical treatment plan and (Right) corresponding cardiac SPARE treatment plan. The planning target volume (PTV) is shown in red. Abbreviations defined in the text.

Figure 36 shows the clinically used radiation treatment plan for Patient 2 (DVH also shown in Figure 34) that originally met all whole-heart dose endpoints for a locally advanced lung cancer patient treated to 60 Gy in 20 fractions. Cardiac sparing after re-optimization is shown with the original clinical treatment plan shown (top left), the cardiac SPARE plan (top right), and the difference map (bottom left). The dose metric table (bottom right) highlights that standard whole heart dose metrics (< 3 Gy and < 2% absolute difference) do not reflect the local dose deposition that the substructure metrics are able to capture. For example, the LV-V5 was reduced from 30.6% to 14.7% after re-optimization. Furthermore, the mean dose to the AA was reduced by ~6 Gy and the LADA $D_{0.03cc}$ was reduced below 10 Gy (threshold for coronary artery calcification[32] presented in Table 8) with sparing.



	Clinical	Spared	Abs Diff
Standard Whole Heart Metrics			
Heart Max (Gy)	63.4	66.1	-2.7
Mean Heart (Gy)	9.6	7.5	2.1
Vol. >40 Gy (%)	3.4	2.0	1.4
Vol. >60 Gy (%)	0.6	0.7	-0.1
Substructure Metrics			
Mean AA (Gy)	24.3	17.9	6.4
Mean RA (Gy)	17.4	11.9	5.5
LV-V5 (%)	30.6	14.7	15.9
LADA Max (Gy)	10.0	7.0	3

Figure 36: Top row: (Left) Clinically treated plan for an advanced stage lung cancer patient. (Right) Cardiac substructure spared plan. Bottom row: (Left) Dose difference map (clinical less cardiac spared plan) highlighting major dose reductions to cardiac substructures. (Right) Dose metric table showing select standard whole heart dose metrics and substructure metrics. Maximum dose defined as dose to 0.03 cc volume. Abbreviations defined in the text. DVH shown in Figure 34.

Optimal beam arrangements led to further cardiac substructure dose reduction in 4 patients. Figure 37 shows the original clinical plan (left), re-optimized SPARE plan (center), and New Angles plan (right) for Patient 5 who had a left lung cancer treated to 48 Gy in 4 fractions. This figure shows that although there was a slight change for the cardiac substructures after plan re-optimization (mean reduction over all substructures: 0.2 ± 2.1 Gy), increased sparing after beam angle modification was possible (mean reduction over all substructures: 1.0 ± 1.4 Gy). For example, the mean dose to the pulmonary vein was only reduced by 0.2 Gy after re-optimization but was further reduced

by another 1.1 Gy after beam angle modification. Moreover, beam angle modification allowed for further sparing of the LADA and LA with mean dose reductions of 0.9 and 0.8 Gy, respectively, as compared to the SPARE plan.

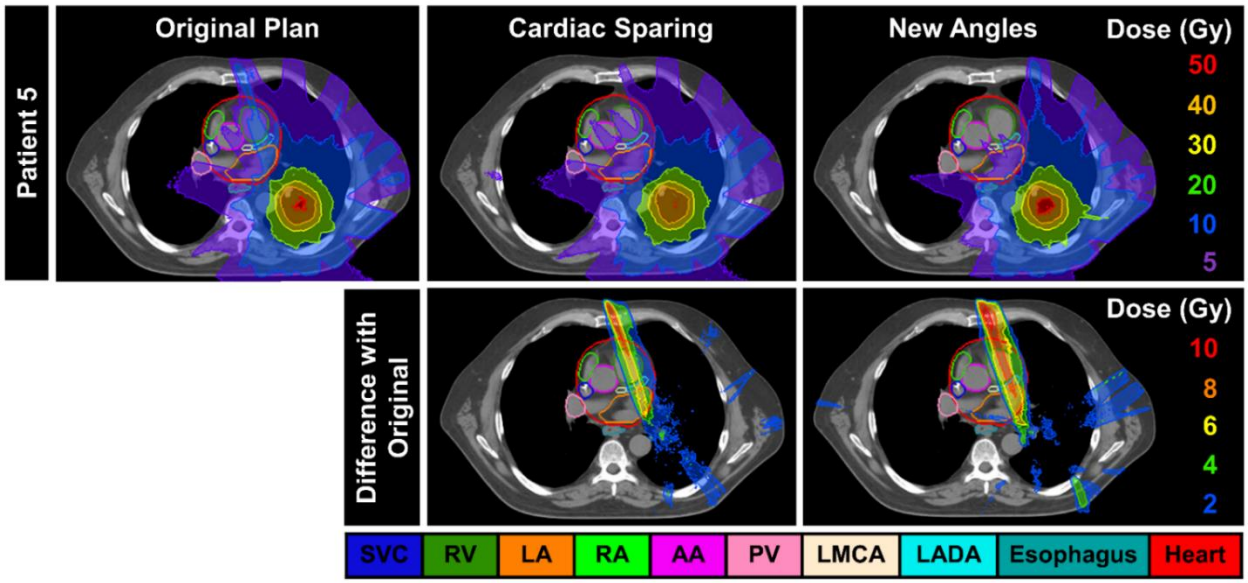


Figure 37: Top row: Original clinical plan (left), re-optimized SPARE plan (center), and New Angles plan (right) for a patient with a left lung tumor. Bottom row: Difference maps comparing the re-optimized SPARE plan and the New Angles plan to the original clinical plan. Difference maps are the original plan less the new plan. Abbreviations are defined in the text.

Discussion

This work introduced cardiac substructures into CT-based treatment planning incorporating a co-registered low-field MRI to quantify potential dosimetric advantages for improved cardiac sparing. This was completed through the retrospective re-optimization of treatment plans, as well as modifying the original beam angle arrangement to minimize cardiac substructure radiation dose, all while attempting to maintain PTV coverage and continuing to meet clinical endpoints for other critical OARs.

Even though current cooperative trials use volumetric measures based only on MHD endpoints[19,20], introducing cardiac substructure segmentation into radiation treatment planning may help better study and define radiation-induced cardiac injury. Some studies have aimed to investigate the dosimetric impact of different types of therapy on cardiac substructure sparing. A study by Ferris *et al.*[234] evaluated cardiac substructure sparing for optimized VMAT and intensity modulated proton therapy (IMPT) and found that cardiac-optimized plans led to statistically significant improvements in mean dose to the chambers, great vessels, and coronary arteries. Our findings agree with Ferris *et al.* with respect to significant reductions in the maximum dose to the LADA and RV (greater than 4 Gy on average) while maintaining or improving clinical OAR (e.g., lung, esophagus, and spinal cord) constraints and PTV coverage. Likewise, Lester *et al.*[244] created re-optimized VMAT plans to reduce radiation dose to the coronary arteries and cardiac valves. At present, few studies have integrated cardiac substructures into treatment planning optimization. Ferris *et al.* evaluated cardiac spared plans using VMAT and IMPT with CT for locally advanced non-small lung cancer patients with a conventional fractionation to 60 Gy under free-breathing conditions[234]. Lester *et al.* focused on cardiac spared planning for mediastinal lymphomas by incorporating ECG-gated CT and coronary angiography acquired at deep inspiration breath hold[244]. These patient populations were different from the present study of 11 out of 16 lung cancer stereotactic body radiation therapy cases (3-4 fractions) with 7 end-exhalation, 7 end-inhalation, 2 free-breathing to test different conditions. In addition, the present study incorporated a low-field MRI as an adjunct to treatment planning CTs whereas the Lester *et al.* study used CT angiography[244]. While MRgRT was employed in this work, the dosimetry strategies of re-

optimization using cardiac substructures and beam angle arrangement modification are applicable to other x-ray-based treatment planning modalities as our atlas and deep learning substructure segmentations work on CT-SIM image inputs.

Negligible increases in treatment time per fraction and MUs delivered after plan re-optimization were observed, suggesting similar complexity of the radiation treatment plan. Moreover, even though the modified beam angles plans involved either adding or removing beams in the revised treatment plan, the differences in treatment time per fraction and MUs delivered were negligible ($P < 0.05$).

Modifying the beam angle and number of beams used to consider cardiac substructures after the plan has been re-optimized also has the potential to increase cardiac substructure radiation sparing. However, much like the findings by Lester *et al.*[244], the results were patient specific as lesion location and proximity to the heart and its substructures played a role in if the patient would benefit from plan re-optimization and beam modification. Patients that benefited from beam angle modification varied in both the number of beams added or removed and in the proximity of the lesion to the heart (i.e. directly adjacent). So, although beam angle modification was shown to provide improvements over solely re-optimizing the plan for select cases (4/16 cases), re-optimization alone provided the majority of cardiac substructure sparing. Tumor location also plays a role in the extent a substructure is able to be spared. For example, the LA for Patient 2 was directly adjacent to the tumor volume yet the mean dose difference after re-optimization of the LA as shown in Figure 33 revealed only minor improvement (< 1 Gy) was possible. Thus, this suggests that sparing substructures closer to the tumor

volume may be difficult although accurately quantifying the dose to substructures offers value for clinical risk assessment.

There was a statistically significant sparing of mean dose to all studied cardiac substructures achieved after plan re-optimization. This may be due to the added weight in the optimizer for when the substructures are included. However, Figure 36 highlights that standard whole heart dose metrics had decreased sensitivity to a cardiac sparing treatment planning approach, whereas individual substructure endpoints clearly identified dosimetric, and clinically meaningful gains (i.e., associated with clinical outcomes). The insufficiency of quantifying the MHD alone has been recently affirmed by studies recommending the inclusion of cardiac substructures as RT treatments become more conformal (i.e. intensity modulated RT)[245,246]. For example, the LV-V5, which has been shown to be more predictive of acute cardiac events than mean heart dose[28], was reduced ~15% and the mean dose to the AA was reduced by ~6 Gy, suggesting that with confirmation in a larger cohort, further sparing may offer potential for improved survival[35]. This underscores the importance of using more sensitive metrics for dose evaluation and not simple whole-heart evaluations that are currently being implemented.

Respiratory motion was managed via breath-holding for the majority of the patients. At breath-hold, there is still the potential impact of cardiac motion which was not accounted for in this study due to not having cardiac gated 0.35 T MR-linac images. It has been shown that even under breath-hold conditions, cardiac substructures may displace ~5-7 mm throughout the cardiac cycle [247,248]. Thus, incorporating a PRV representing the variability of the cardiac substructures over a patient's imaging and treatment course will be the next step of this work. However, as substructure PRV

recommendations do not currently exist for each substructure and this study was unable to account for cardiac motion, they were beyond the scope of the current work. This may be possible through the use of van Herk's formalism[249], which was used by Levis *et al.*[248] to estimate PRVs for the coronary arteries. Additionally, while this work was based off of CT-based treatment planning that enables more widespread applicability to x-ray-based approaches, MR-only treatment planning is gaining popularity, and a future direction includes translating the work to MR-only plans. Furthermore, increasing the size of the patient cohort with varied target locations will help identify the patient geometries that will benefit most from cardiac substructure sparing. However, the size of the patient cohort in the current study is consistent with the previously mentioned studies where 7-8 patients were used[244,248]. This may be completed through applying this work to a prospective clinical trial, like that of Jacob *et al.*[129], or be applied to multi-institutional studies, such as the study recently completed by Dess *et al.*[250], to determine if cardiac substructure dosimetric sparing has an effect on clinical outcomes.

Conclusion

This work applied a multimodality workflow to enable robust dose sparing of cardiac substructures, as well as maintain PTV and OAR doses. New treatment plans did not substantially increase delivery time or required monitor units, suggesting a negligible increase in plan complexity when cardiac substructure sparing was introduced. Validation in a larger cohort with appropriate margins will offer the potential to reduce radiation-related cardiac toxicities and the dose assessment of currently overlooked radiosensitive substructures.

Part 2 "A Deep Learning Cardiac Substructure Pipeline for MR-Guided Cardiac Applications"

Introduction

Radiation dose from thoracic RT treatments is strongly associated with radiation-induced heart disease, which includes both acute and late cardiotoxicities[5,6,16]. In RT, the current standard of care is to delineate and consider the entire heart as a single organ and use simple metrics like MHD and dose/volume relationships to evaluate cardiac risks[18]. Importantly, these whole-heart dose metrics do not provide any information about where the dose is being distributed. Yet, radiation dose to individual cardiac substructures, like the LV, LA, and LADA, have been more strongly linked to late cardiac morbidities than the commonly used whole-heart dose metrics[28,41,42]. However, cardiac substructures are not currently considered in treatment planning as they are poorly visualized on standard CT simulation datasets[18]. With the improved soft tissue contrast provided by MRgRT, sensitive cardiac substructures can now be visualized in MR-guided treatment planning (Figure 38) and considered for dose evaluation. Recently, MRI has been coupled with a linear accelerator offering this increased visibility of tumors and OARs in real time[251]. Still, even with proper visualization, online delineation of cardiac substructures is impractical as it may take over four hours per patient[45,46].

Several recent atlas-based techniques have been developed for automatic cardiac substructure segmentation[48,132,155]. Yet, atlases require long processing times[179] and are heavily reliant on the quality of image registration between new patients and those existing in the atlas. DNNs including the U-Net[204], GANs[158], and CNNs[92] have also been applied for automated cardiac substructure segmentation and have shown great promise for rapidly generating contours.

Cardiac substructure segmentation for RT applications has been primarily focused on CT datasets as this is the imaging modality that is the standard of care for generating a patient's radiation treatment plan[132,155]. On CT images, studies by Zhou *et al.* and Kaderka *et al.* were able to achieve good agreement in the cardiac chambers with DSCs greater than 0.7[132,155]. The U-Net, originally designed by Ronneberger *et al.*[119] for medical image segmentation, has been applied to both MR and CT volumes for state-of-the-art segmentation results[157,204]. Through incorporating the U-Net, Payer *et al.* was able to achieve excellent agreement between aorta volumes when comparing generated segmentations to physician delineated ground truth[157]. Similarly, our previous work yielded DSCs ≥ 0.84 in the cardiac chambers by training a multi-channel (3.0 T MRI and CT) 3D U-Net for predictions on non-contrast CT[204].

Regarding MR-only results, Mortazi *et al.* conducted automated segmentation of seven substructures using a 3D b-SSFP sequence acquired in free breathing at 1.5 T in 17 seconds using a multi-planar deep CNN with adaptive fusion[92]. Similarly, Avendi *et al.* utilized a deep CNN and stacked auto-encoders combined with a deformable model-based approach to automatically segment the left ventricle[252]. This was completed using a b-SSFP sequence acquired with a dedicated eight-element cardiac coil for cardiac gating and during end-expiration breath hold[92,252,253].

There are numerous strategies to improve the variability and diversity of available data, without the need to collect new unique samples, which is referred to as data augmentation. Data augmentation has been shown in numerous studies to improve auto-segmentation accuracy, as well as minimize model overfitting[92,204,252]. The use of inter-fractional data for model augmentation has been incorporated previously, such as in the

study by Brion *et al.*, where the utilization of a patient's daily cone-beam CTs had, for the purpose of augmenting the deep learning model, been shown to significantly increase segmentation accuracy in RT[254]. However, this method of deep learning model augmentation has not been implemented for applications in MRgRT.

To address the unmet need of cardiac segmentation for MRgRT, we have implemented a novel custom DL framework to efficiently generate segmentations on low-field (i.e. 0.35 Tesla) MRI with the overarching goal of integrating cardiac substructures into daily dose assessment and to facilitate enhanced cardiac sparing. This work developed and validated an efficient and accurate deep learning pipeline for automated segmentation of 12 cardiac substructures on MR-linac images. The novelty of this work lies in the custom deep learning framework was developed for low-field MR-linac images where daily inter-fractional MR images were incorporated for data augmentation (Figure 38). The successful completion of this work will enable the integration cardiac substructures for dose assessment and improved cardiac sparing for future MR-linac implementation in prospective clinical trials using MRgRT.

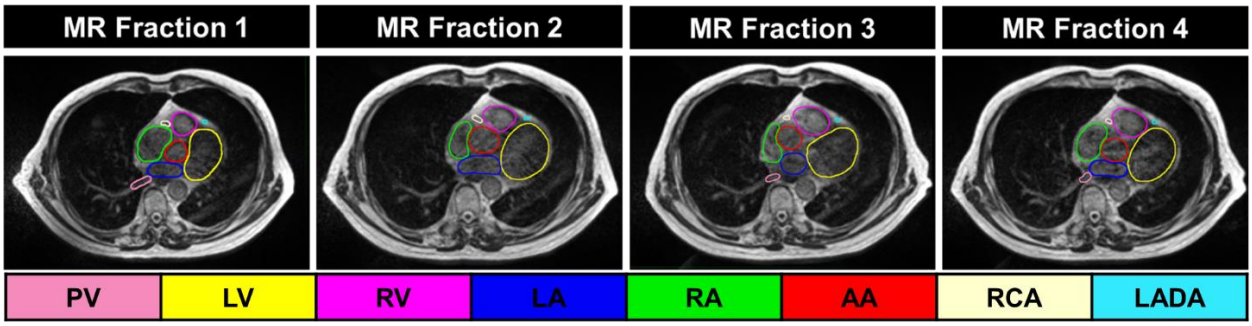


Figure 38: Axial views at the same thoracic level highlighting positional variations in cardiac substructures acquired at breath-hold between various 0.35T MR-linac fractions for a representative patient. Abbreviations: MR = magnetic resonance; PV = pulmonary vein; LV = left ventricle; RV = right ventricle; LA = left atrium; RA = right atrium; AA = ascending aorta; RCA = right coronary artery; LADA = left anterior descending artery.

Methods

Patient Cohort

Twenty-three patients who underwent RT for either thoracic or abdominal cancer were retrospectively reviewed on an Institutional Review Board approved study. Patient treatments ranged from 30-70 Gy in 4-35 treatment fractions. Treatments consisted of esophageal, right and left lung, chest wall, lymph node (pericardial), stomach (gastric lymphoma), and liver dome lesions. Of the 23 patients, 17 were SBRT and the other 6 were conventional treatments.

Imaging Methods

Patients were imaged using a balanced T1/T2 b-SSFP sequence (TrueFISP, Siemens, MAGNETOM Avanto, Syngo MR B19) on a 0.35 Tesla ViewRay MRIdian linear accelerator (ViewRay, Mountain View, CA). The TrueFISP sequence is often used in cardiac imaging due to its high signal-to-noise ratio and imperviousness to motion artifacts[227]. The four utilized scans are outlined in Table 10 where each used a flip angle of 60°, phase encoding in the anterior posterior plane, and a transversal orientation. The generalized auto-calibrating partially parallel acquisitions, or GRAPPA method, was utilized for breath-hold sequences to accelerate image acquisition[255].

Sequence	Slices	Time (s)	FOV (cm^3)	Resolution (cm^3)	TE (ms)	TR (ms)	Bandwidth (Hz/Px)	GRAPPA
1	80	17	45x45x24	0.16x0.16x0.30	1.62	3.83	385	2
2	144	25	50x45x43	0.15x0.15x0.30	1.27	3.00	599	2
3	288	172	50x45x43	0.15x0.15x0.15	1.45	3.37	535	N/A
4	240	175	54x30x36	0.15x0.15x0.15	1.45	3.38	534	N/A

Table 10: Description of utilized MR sequences and their different parameters. Abbreviations: FOV = field of view; TE = echo time; TR = repetition time; Px = pixel; GRAPPA = generalized auto-calibrating partially parallel acquisitions.

Twenty-one patients were treated under breath-hold conditions (17-25 seconds, 14 end-inhalation and 7 end-exhalation, 1.5x1.5x3 mm^3) and two under free breathing conditions (3-minute scan, 1.5 mm^3 isotropic resolution) as shown in Figure 39.

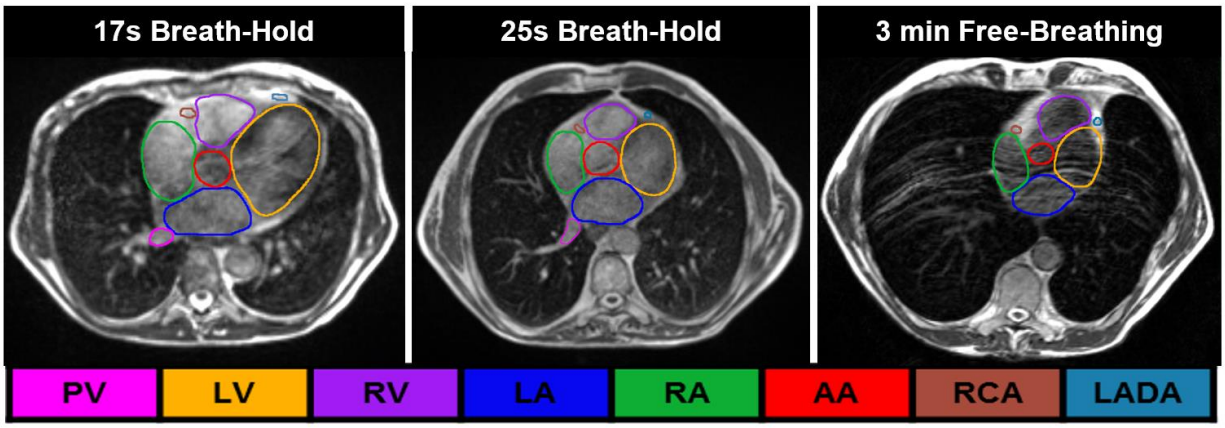


Figure 39: Axial MR slices for 3 patients illustrating end-exhalation breath-hold (Left: 17 seconds, Center: 25 seconds) and free breathing (Right: 3 minutes) conditions with physician delineations of select cardiac substructures. Abbreviations: PV = pulmonary vein; LV = left ventricle; RV = right ventricle; LA = left atrium; RA = right atrium; AA = ascending aorta; RCA = right coronary artery; LADA = left anterior descending artery.

Clinical imaging included an MR-simulation (MR-SIM) and 3-4 daily MRI series per patient for a total of n=114 unique image sets (4-5 unique daily MRIs) with an example patient's daily MRIs shown in Figure 38. To avoid preferential weighting due to the number

of fractions, if a patient underwent a treatment with more than four daily fractions, the first four were used. Ground truth segmentations of 12 cardiac substructures were delineated on each patient's MR-SIM. The manual segmentation effort of the ground truth was split between two radiation oncologists. The more experienced of the two radiation oncologists verified delineations for every patient. Studied cardiac substructures included the left/right ventricles (LV, RV) and atria (LA, RA), superior/inferior vena cavae (SVC, IVC), pulmonary artery/veins (PA, PV), ascending aorta (AA), left anterior descending artery (LADA), right coronary artery (RCA), and left main coronary artery (LMCA). Deformable image registration (DIR) was then used to transfer segmentations from the MR-SIM to each daily MRI using a commercially available free-form intensity-based registration (MIM Software, Cleveland, OH). This DIR algorithm has limitless degrees of freedom, employs regularity (i.e. penalty term weight) to ensure the deformation was smooth[145], and has provided high segmentation accuracy (i.e. lowest mean absolute difference in a comparison of eleven different deformable image registration algorithms) for previous single-modality (e.g. CT to CT or MR to MR) registrations[146]. All propagated contours were then verified and corrected as needed with a split effort between two radiation oncologists. The more experienced radiation oncologist then completed a final verification of each patient and each daily image set for continuity. MR images were resampled to 1 x 1 x 1.5 mm^3 and then cropped to a 128 x 128 in-plane resolution over 64 slices. Image cropping was competed around the heart centroid and then padded axially to 128 x 128 x 128.

Deep Learning Methods

Figure 40 shows the deep learning architecture of the modified 3-dimensional Neural Network (i.e. 3D U-Net) used in this study. This architecture is based off of a previously validated 3D U-Net[204] but has been optimized for the end point of automatic predictions on low-field MR images. In total, eighteen patients, with 90 total unique MRIs, were used to train a 3D U-Net on all 12 substructures simultaneously in 3D.

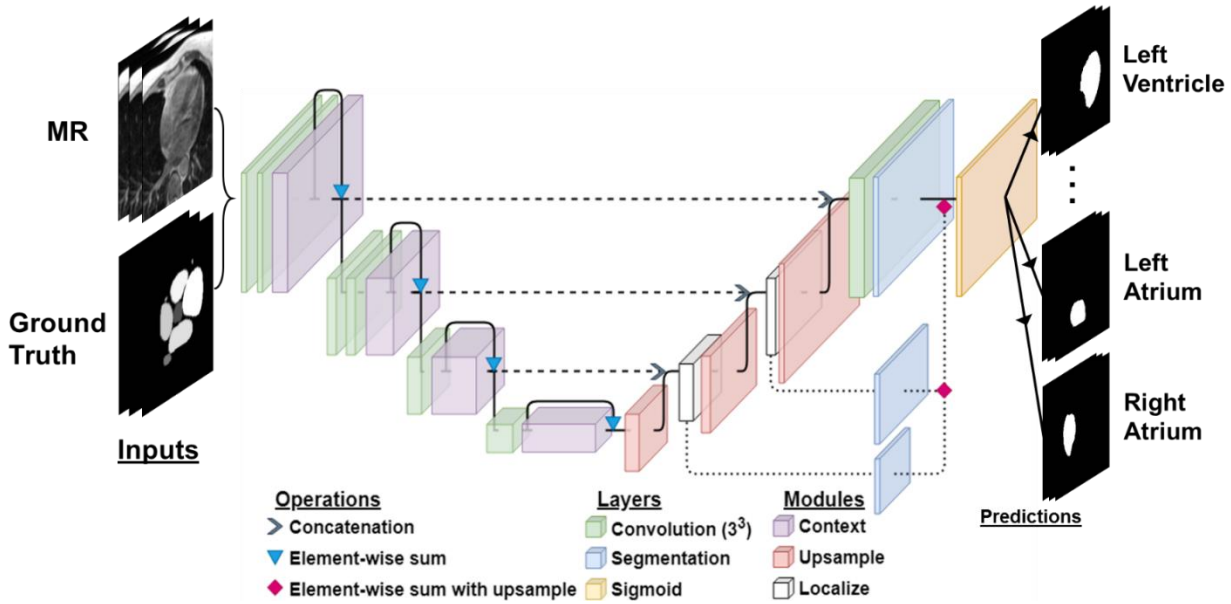


Figure 40: 3D U-Net architecture for cardiac substructure segmentation with low field MR inputs. Predictions were outputted for each substructure.

The 3D U-Net shown in Figure 40 comprises of contraction and expansion pathways. The contraction pathway utilizes context modules (previously described[204]) to aggregate high-level feature information (Figure 40, left)[162]. The expansion pathway allows for the assembly of feature information for localization (Figure 40, right)[162]. As shown by the blue boxes in Figure 40, deep supervision was also implemented and was described previously[204]. To maximize information gained from high resolution features, double convolutions were used in the first and second layers of the 3D U-Net. Strided

convolutions were used instead of maximum pooling operations in the context pathway (Figure 40, left) as they have been shown to increase model accuracy if the computational load allows[256]. An adaptive momentum estimation (Adam) optimizer[169] was used with an initial learning rate of 5e-4, a batch size of 1, and randomly initialized weights. During training, if the validation loss did not improve within 20 epochs (i.e., one forward pass and one backward pass of all the training samples), the learning rate was reduced by 50%. Training data (n = 18 unique patients) was split into 78% (n = 79 unique image sets) training and 22% validation data (n = 20 unique datasets or 4 unique patients). The 78/22 split was used to ensure only complete individual patient datasets were considered in either the training or validation datasets. Sixteen input filters were used initially, and filter number doubled along the context pathway.

A loss function based on Dice allows for a way to circumvent the imbalanced ratio of foreground to background pixels in medical images[105]. Additionally, as the training data has ground truth segmentations, optimizing DSC is what is desired throughout training. As the DSC is to be maximized to improve the segmentation result, the loss function needs to be minimized and is thus represented in Equation 14[257].

$$Unweighted\ Dice\ Loss = \sum_t 1 - DSC_t \tag{14}$$

The Dice-weighted loss function allows for multiple classes (i.e. substructures t ∈ T) to be learned during the training and is shown in Equation 15[258].

$$T_{DSC} = \frac{\sum_{t \in T} w_t * t_{DSC} + \alpha}{\sum_{t \in T} w_t + \alpha} \tag{15}$$

Where w_t is each class's predicted weight. For numerical stability, a small value α is included[258]. The specific problem of substructure segmentation involves classes that are highly imbalanced due to differences in size and shape. Furthermore, as easily

segmented substructures comprise most of the loss and dominate the gradient, we applied a variable weighting scheme using a Dice-weighted focal loss, as shown in Equation 16[257].

$$Focal\ Dice\ Loss = \sum_t w_t(1 - DSC_t^{1/\beta}) \quad (16)$$

The value β is a tunable hyperparameter and as it increases, the well segmented classes are down-weighted[259]). This focal Dice loss function emphasized the more difficult substructure segmentation tasks to manage any class imbalances as extreme class imbalances overwhelm the cross-entropy loss during training. Additionally, the focal DSC loss function helps to minimize the false known correlation between a larger volume and an improved DSC[118]. Furthermore, the focal Dice loss function has been shown to improve hyperparameter robustness for unbalanced tasks, and improve overall segmentation accuracy for small structures[260].

Materials

Experiments were performed using Python v3.6 in Windows 10 (64-bit 10 core Intel® Xeon® CPU-E5-2690 v4 at 2.60 GHz and 112 GB of random-access memory (RAM)) with an NVIDIA Tesla P100 GPU with 16 GB of RAM and 3584 CUDA cores (Keras 2.0 with TensorFlow backend). Beyond including 3-4 daily MRIs per patient, data were augmented through flipping, rotating (0-30°, 1° increments), scaling (± 25%, 1% increments), and translating (up to 10 pixels in 3 cardinal planes). These augmentation techniques, along with a 50% dropout, were implemented to minimize model overfitting.

Post-Processing, Testing, and Parameter Optimization

We further improved the agreement to ground truth delineations by introducing fully connected 3D dense CRF as a post-processing step[160,204]. The 3D CRF post-

processing has been described in detail in our previous work[204] and was used to correct for remote island segmentations and improve overall smoothness. Five unique test patients (n=5 as only MR-sim datasets were considered for testing) were held out from the training patient cohort. DSC and MDA were used to compare DL segmentations to ground truth. The learning rate schedule was tuned for optimal segmentation (i.e. initial learning rate, decay, and patience). Lastly, the β hyperparameter from the above equation was assessed at integer values from 2-5.

Statistical Assessment

Two-tailed Wilcoxon signed ranks tests were completed for statistical analysis between ground truth and auto-segmented volumes, with $P < .05$ considered statistically significant. Volumetric and geometric measures were assessed. All statistical analysis was completed in SPSS (Version 26.0, Chicago, IL).

Results

The model stabilized after training for 340 epochs (training error < 0.001) which took 32 hours to complete. Substructure contour generation for a new patient input took 20 seconds. Wilcoxon signed ranks test revealed that there 11/12 cardiac substructures did not have significant differences in volume between 3D U-Net auto-segmentations and ground truth for any cardiac substructure ($P > 0.05$). There was, however, a 0.47 cc (or 80%) average difference in LMCA volume which was statistically significant.

Shown in Table 11 are DSC results across the analyzed cardiac substructures.

Substructure	0.35T MR-linac Focal Dice Loss + CRF (n=5)	Hybrid Diagnostic 3T MR/CT-SIM + CRF[204] (n=11)	MR Literature (1.5T – 3.0T)[92,157,252]	CTCA Literature (High Resolution 0.4mm^3)[182]
Left Ventricle	0.89 ± 0.02	0.91 ± 0.01	0.88-0.92	-
Left Atrium	0.83 ± 0.02	0.87 ± 0.02	0.81-0.88	-
Right Ventricle	0.84 ± 0.02	0.84 ± 0.02	0.77-0.88	-
Right Atrium	0.84 ± 0.04	0.87 ± 0.02	0.83-0.87	-
Great Vessels	0.78 ± 0.05	0.81 ± 0.00	0.72-0.79	-
Left Main CA	0.42 ± 0.14	0.50 ± 0.18	-	0.60-0.70
Right CA	0.43 ± 0.11	0.50 ± 0.09	-	0.60-0.70
Left Anterior Descending CA	0.45 ± 0.10	0.53 ± 0.08	-	0.60-0.70

Table 11: DSC for current method using a 3D U-Net with focal Dice loss for low field MR (n = 5 test patients, column 2), previously published 3D U-Net with hybrid diagnostic 3T MR/CT information (n = 11 test patients, column 3)[204], current literature for cardiac substructure segmentation at 1.5-3.0T, and high-resolution CTCA data (voxel size 0.4 mm^3) from the literature across substructures. Abbreviations are defined in the text.

It was found that DL provided accurate segmentations for the chambers (DSC = 0.85 ± 0.01), great vessels (DSC = 0.78 ± 0.05), and pulmonary veins (DSC = 0.70 ± 0.08). DSC for the coronary arteries was 0.43 ± 0.10. The results in Table 11, column 2, reflect a β value of 4 (Equation 16), which yielded the highest DSC values across test patients. Table 11 also reveals DSC results from our previous work[204] where cardiac substructure segmentations were generated on non-contrast CTs. Lastly, Table 11 shows DSC results from the current literature for the cardiac chambers and great vessels at MR field strengths between 1.5T and 3.0 T and for coronary artery segmentations generated on CTCA.

MDA results across all cardiac substructures are shown in Figure 41.

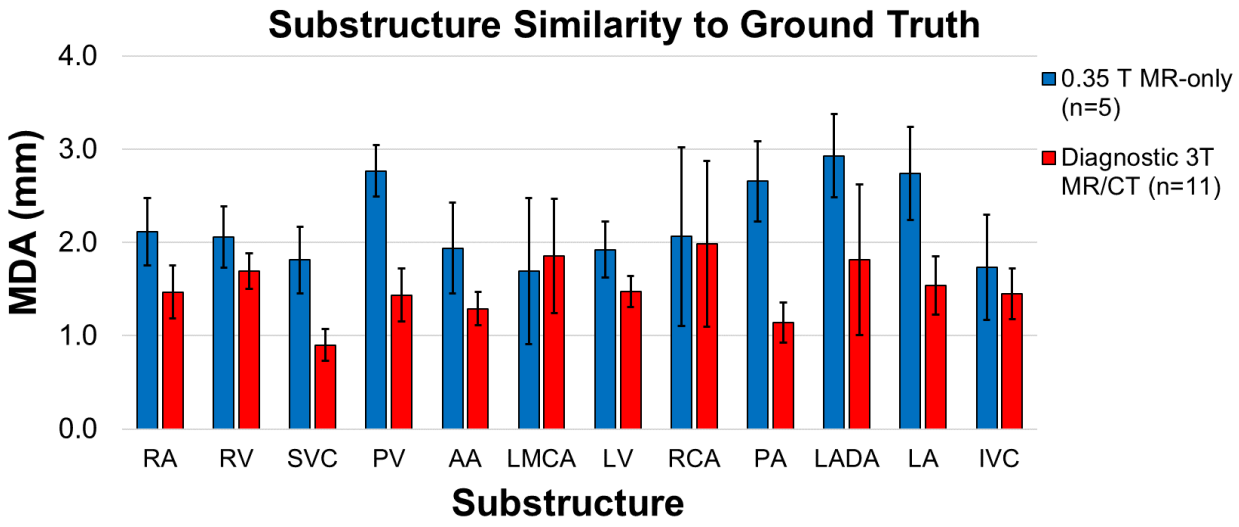


Figure 41: Mean distance to agreement (MDA) results across all cardiac substructures for the current MR-only deep learning model (blue, n = 5) and the previous hybrid MR/CT deep learning model[204] (red, n = 11).

As shown in Figure 41, MDA across all substructures was less than 3 mm. The IVC and SVC experienced the smallest MDA across test patients of 1.7 ± 1.0 mm and 1.8 ± 0.8 mm, respectively. The LADA had the largest MDA on average of 2.9 ± 1.4 mm. Centroid displacements for all cardiac substructures between the manually generated ground truth and the DL auto-segmentations in the L-R, A-P, and S-I axes are shown in Table 12. On average centroid displacements between manually drawn ground truth and DL were 2.15 ± 2.52 mm in the L-R, 2.10 ± 1.38 mm in the A-P, 3.39 ± 2.67 mm in the S-I.

Substructure	L-R (mm)	A-P (mm)	S-I (mm)
LV	0.82 ± 0.65	1.24 ± 1.20	1.04 ± 0.55
LA	1.32 ± 0.68	0.98 ± 0.57	3.26 ± 0.80
RV	1.08 ± 0.39	1.78 ± 1.68	3.40 ± 4.01
RA	0.98 ± 0.47	1.30 ± 0.76	2.70 ± 2.53
AA	0.80 ± 0.84	1.06 ± 1.52	1.62 ± 1.42
SVC	0.46 ± 0.29	0.92 ± 0.89	2.12 ± 2.54
PA	1.90 ± 1.47	5.06 ± 1.57	3.20 ± 2.29
IVC	0.70 ± 0.79	2.00 ± 2.30	2.16 ± 2.65
PV	8.98 ± 7.82	2.36 ± 3.38	3.56 ± 2.05
LADA	5.30 ± 5.74	4.72 ± 3.15	10.58 ± 5.91
RCA	2.54 ± 1.77	1.74 ± 1.23	6.18 ± 6.73
LMCA	0.96 ± 1.06	2.02 ± 1.55	0.86 ± 0.26

Table 12: Centroid displacements for all cardiac substructures between the manually generated ground truth and the DL auto-segmentations in the L-R, A-P, and S-I axes

Three-dimensional results for the best case, Patient 1, are shown in Figure 42. Patient 1 was diagnosed with esophageal cancer and was treated to 32 Gy in 13 fractions. For the cardiac chambers, great vessels, and IVC, (9/12 substructures) DSCs were greater than 0.80 for this patient. With regard to the coronary arteries (i.e. LADA, LMCA, and RCA), DSCs were greater than 0.55 and MDA was 1.60 ± 1.03 mm. When comparing the RV, AA, and LADA, for example from Figure 42, the deep learning segmentations appear more smoothed due to the application of CRF for post-processing.

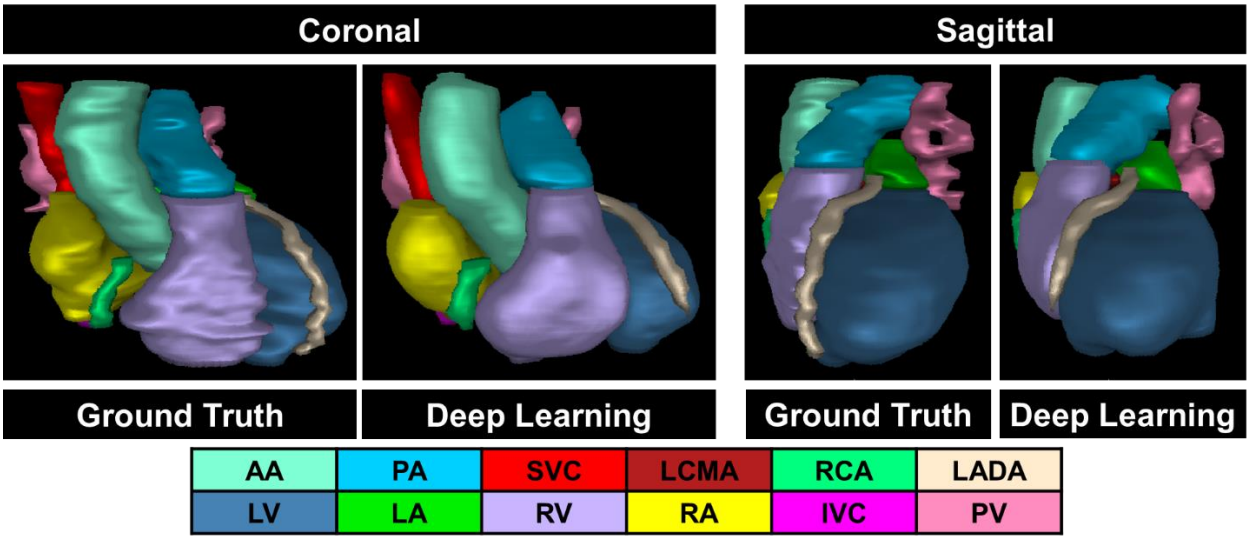


Figure 42: Three-dimensional results for the best case, Patient 1, showing both coronal and sagittal views of the manually drawn ground truth compared to the MR-only deep learning model segmentation predictions. Substructure abbreviations are defined in the text.

Shown in Figure 43 are axial images displaying results for Patient 2 and Patient 4. Each axial image shows both manually segmented ground truth segmentations, as well as DL auto-segmentations. Patient 2 was diagnosed with a malignant neoplasm of the upper right bronchus (limited stage small cell cancer) which was treated to 60 Gy in 20 fractions. The malignant neoplasm, shown in the left side of the top left image in Figure 43, caused the DL model to be unable to generate a segmentation prediction for the right PV (manually drawn ground truth shown in pink). For this reason, Patient 2 had the lowest DSC value for the PV across test patients of 0.57. Patient 4, shown on the right side of Figure 43, was diagnosed with a malignant neoplasm of the lower left lung lobe (adenocarcinoma grade 2) which was treated to 48 Gy in 4 fractions. Located in the center of both images for Patient 4 is a high intensity artifact. As nothing like this was included in the training cohort, this artifact negatively affected the PA, LA, and LMCA automatic segmentation results for Patient 4, as can be seen in Figure 43.

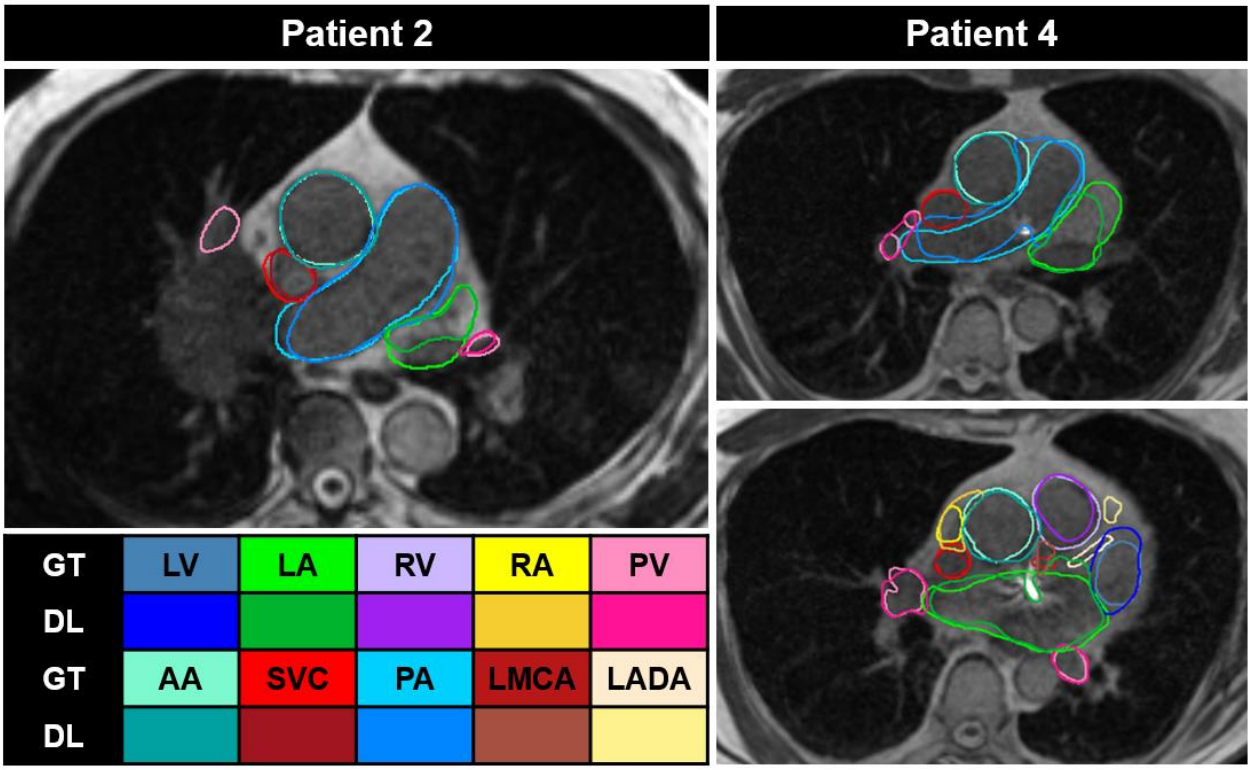


Figure 43: Axial images for two different test patients (Left: Patient 2, Right: Patient 4) revealing how abnormal patient anatomy or image artifacts affected deep learning automatic segmentation results. Each axial image shows both manually segmented ground truth segmentations, as well as DL auto-segmentations

Shown in Figure 44 is an axial MR slice comparing the manually drawn ground truth and DL segmentations for Patient 5. Patient 5 was diagnosed with liver cell carcinoma with an anterior dome lesion which was treated to 50 Gy in 5 fractions. Figure 44 also shows the delivered dose.

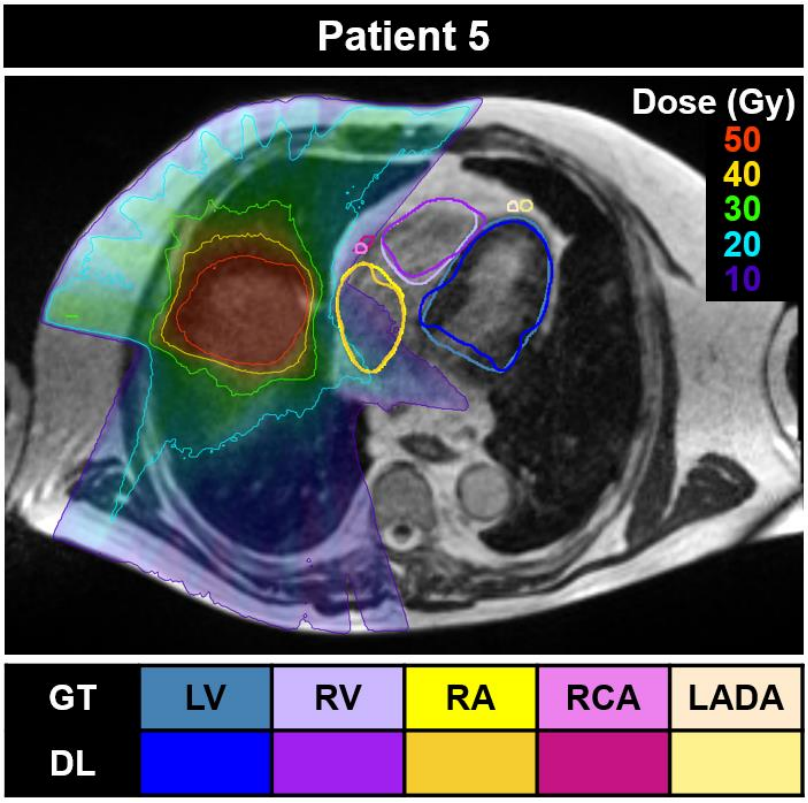


Figure 44: Axial MR slice comparing ground truth (GT) and deep learning (DL) segmentations for a patient undergoing liver treatment. Substructure abbreviations are defined in the text.

Discussion

This work developed a novel deep learning pipeline to segment cardiac substructures with the principle goal of applications in MRgRT. Several methods were used for data augmentation including the unique contribution of incorporating 3-4 daily MR images per patient to improve model accuracy and to minimize model overfitting. This pipeline lead to accurate segmentations of the chambers and great vessels and yielded promising coronary artery results on low-field MR images.

With the importance of cardiac substructures being shown in the pathogenesis of radiation-associated heart disease, automatic cardiac segmentation has been the topic of much recent exploration[132,139,204]. However, to our knowledge, it has not been applied

to images for applications in MRgRT. Additionally, for RT applications, the use of daily MR images for each patient provides a unique methodology for data augmentation. Atlas-based models have been widely explored for cardiac substructure segmentation yielding DSCs greater than 0.75 in the chambers and great vessels, but fail to segment the coronary arteries (DSCs less than 0.27)[48,132,155]. Atlas methods are limited as they rely heavily upon the quality of a registration between patients and are unable to consider large amounts of patient data[179].

This current work follows more closely current applications of deep neural networks where hybrid MR/CT information was used to train the model[204], or studies where CTCA was utilized[92,157]. When comparing to our previous study that utilized hybrid MR/CT information, shown in Table 11, cardiac chamber and great vessel results from the current study were within 0.04 DSC. Coronary artery DSC results from the present study were within 0.08 DSC, as can be seen in Table 11, which is promising as this study did not incorporate contrast-enhancement or cardiac gating. This work adds to the current available literature by allowing for substructures to be automatically generated on low-field MR images for applications in MRgRT.

To further improve automatic segmentation accuracy, fully connected 3D CRF were implemented for post-processing of cardiac substructures and was discussed in length in our previous work[204]. Several data augmentation techniques were also utilized to create variation in the model training and enhance automatic segmentation results. However, despite including 3-4 daily MRIs to augment the data for each patient, automatic segmentation agreement may improve with a larger training cohort. Nevertheless, studies with training and validation sizes similar to this one have been used previously[190,204].

Centroid displacements were largest in the S-I axis likely due to the larger slice thickness of 3 mm for each of the five test patients, as compared to the 1.5x1.5 mm in-plane resolution. As there is a known correlation between signal-to-noise ratio and MR field strength[261], similarity results using the strategy described in this work may be improved upon at higher field strengths, which should be assessed. At 1.5 Tesla, Mortazi *et al.* utilized an image resolution of $(1.6\sim2)\times(1.6\sim2)\times(2\sim3.2)$ mm[92], which is comparable to what was used in this study. When comparing DSC results for the AA, PA, and cardiac chambers, results from the current study were within 0.04 and exceeded their AA DSC by 0.07[92]. It is important to note, however, that MRIs from the current study were not ECG gated. MR-linacs do exist at field strengths of 1.5 Tesla, such as the Unity MR-linac (Elekta AB, Stockholm, Sweden). However, considering images at a different field strength would require the deep learning segmentation pipeline to be retrained as there would be variation in image intensity and relevant features.

The utilized MR images were not cardiac gated and therefore, we did not account for cardiac motion in this study. Even though cardiac displacements can be 3-8 mm[52], cine-angiography or echocardiography would be required, and they are currently not considered clinically[232]. Respiratory motion was accounted for through breath-hold techniques in 21/23 patients, but scan acquisition times ranged from 17-25 seconds and consequently captured multiple cardiac cycles. Thus, the delineated cardiac substructures used in training represent an average position over the scan course. This cardiac motion captured in the MR may introduce inter-observer variability by presenting difficulties in coronary artery segmentation as they can become indistinct and noncontiguous[141]. To combat this, cardiac substructure segmentations were verified by

two radiation oncologists. Cardiac motion may be considered in future studies through the implementation of a safety margin around the substructure which has been previously proposed[33,53].

Now that cardiac substructures can be generated efficiently on low-field MR images, dosimetric analysis and strategies for cardiac avoidance can be considered. Recently in RT planning, the LADA has been suggested as an avoidance structure[32]. Additionally, numerous recent studies have employed cardiac SBRT treatments for non-invasive radio-ablation for arrythmias[198,262], cardiac fibromas[263], and other cardiac indications[264]. Efficient and accurate automatic segmentation of cardiac substructures for MRgRT may be a useful application to cardiac SBRT studies for rapid OAR and target localization.

Conclusion

Even at the decreased in-plane resolution of the low field MR-linac, our implementation of a novel 3D U-Net has provided promising preliminary results for cardiac substructure segmentations. This significant work opens the door to conduct studies on determining cardiac substructure dosimetric limitations through a prospective clinical trial using MR-guided radiation therapy. This work also has applications in future methodologies for cardiac SBRT.

CHAPTER 7 "CONCLUSIONS AND FUTURE WORK"

Summary of Findings

In this body of work, several image processing pipelines were developed and validated to segment cardiac substructures to better quantify potential opportunities for enhanced cardiac sparing in RTP. Specifically, pipelines were generated for automatic segmentation on non-contrast CT and low-field MRI to lay the groundwork for future prospective studies on the radiation dose to cardiac substructures.

In Aim 1, the automatic segmentation of 12 cardiac substructures and the whole heart on non-contrast CT was achieved through optimization of a novel intensity-based deformable registration atlas that employed a volumetric T2 MRI rigidly registered to CT-SIM datasets. Three atlas approaches were assessed including a single-atlas method and two multi-atlas approaches (i.e. MV and STAPLE with 1-15 atlas matches (MV1-MV15 and ST1-ST15)). Segmentation performance was also assessed via consensus scoring by three physicians. When averaged over five validation patients, all three atlas approaches performed similarly with seven out of 13 cardiac structures (heart, chambers, AA, and PA) having DSC > 0.75. ST10 was deemed a promising segmentation approach as it provided the highest DSC and lowest MDA for 8/10 substructures, and half of all validation contours had centroid displacements < 3.4 mm. Dosimetric assessment of the ST10 generated segmentations revealed no statistically significant differences from ground truth for $LADA_{max}$, MHD, and LV-V5 ($P > 0.05$). Qualitative consensus grading revealed that eight substructures required minor modifications in order to become clinically acceptable. With average DSCs < 0.3, the coronary arteries were not segmented adequately but provided value for general localization. It was found that, on average,

automatic cardiac substructure segmentations for a single patient required ~10 minutes. This initial work is significant in that it offers potential to accurately localize large cardiac substructures on standard non-contrast CT for radiation sparing even when an MRI is unavailable and laid the groundwork for the next aim of the work to further improve substructure segmentation efficiency and accuracy using deep learning.

The second Aim, covered in Chapter 4, focused on further improving substructure segmentation through inputting paired MRI and CT data into separate image channels to train a 3D neural network. The network was trained in ~19 hours using the entire 3D image along with deep supervision and a Dice-weighted multiclass loss function. The hyperparameter optimization of the 3D CRF showed that maximal accuracy was achieved when cardiac substructures were grouped into three categories, which correlated to substructure size and complexity. Implementing 3D CRF as a post-processing measure improved DSC results by 1.2 ± 2.5%. Similarly, it was found that augmenting the data improved the DSC results by 5.0 ± 7.9%. For the coronary arteries, DSC improved by 26.4 ± 4.2% from the previous MA method. In total, with the DL method, the MDA across all substructures was less than 2 mm. There were four instances where the DL model was able to generate automatic segmentations for the LMCA, where the multi-atlas method failed. From the qualitative consensus grading, it was found that the DL method provided clinically acceptable segmentations in all 5 validation patients for three out of four cardiac chambers. When comparing qualitative consensus grades to the previous MA method, improvements were seen in ~75% of cases with the DL method. Overall, generation of automatic substructure segmentations took ~14 seconds per patient as compared to ~10 minutes for the multi-atlas method. This work is significant in that state-

of-the-art cardiac substructure segmentation, requiring a single, non-contrast CT input, was achieved rapidly and more accurately than other methods presently available.

The first portion of Chapter 5 was dedicated to characterizing the excursion of cardiac substructures due to respiration. Measurements of cardiac substructure centroid at 4 phases of respiration over 11 patients showed that overall shifts were largest in the S-I plane and for cardiac substructures at the inferior aspect of the heart (i.e. the IVC and RCA). The IVC and the RCA had maximum excursion displacements > 6 mm (maximum vector displacements of greater than 9 mm). Over all substructures, the A-P was the axis with the least excursion, and substructures at the base of the heart had smallest displacements (i.e. the great vessels). As shown in Figure 20, median excursions for 11/13 cardiac structures were smallest in the A-P axis (median displacements for all substructures were less than 5 mm). It was found that the whole heart volume had insignificant volume changes (~1%) between EE and EI, which parallels results by Yue *et al.*[203]. Moreover, dosimetric analysis across respiratory phases revealed that, although changes in D_{mean} and D_{max} for the whole heart were less sensitive to respiration (< 0.5 Gy), large dose differences for individual substructures were experienced. Thus, this work reveals the lack of sensitivity in only considering measurements of the whole heart. As substructure motion throughout the respiratory cycle is not considered in radiotherapy treatment planning, this work is significant in that quantifying substructure changes over the respiratory cycle allows for their potential use in improving OAR sparing during radiation treatments, especially when a patient is unable to comply with breath-hold conditions for thoracic cancer treatments.

Another challenge in the evaluation of cardiac substructure doses is the inter-fraction setup uncertainties that may occur over the treatment course and potential challenges with intra-fraction motion due to respiration as presented in Chapter 5 part 2. Because the variability of cardiac substructure position after tumor localization has not been well characterized, the work completed in part one of this Aim leveraged unique longitudinal MRgRT data to evaluate inter-fraction displacements of cardiac substructures to facilitate safety margin design. Cardiac substructures were generated utilizing the previously mentioned multi-atlas and deep learning methods for automatic segmentation. It was found that across the heart and substructures, inter-fraction displacements for 18.5%, 17.4%, and 23.1% were > 5 mm (i.e. threshold for motion management[210]) in the L-R, A-P, and S-I axis respectively. For the chambers, the median absolute displacements were less than 2.5 mm in any cardinal plane. Generally speaking, S-I displacements were larger than the other axes, largely due to the axial MRI slice thickness. Individual cardiac substructure displacement demonstrated variability in magnitude and dominant axis, where PRVs from 3-6 mm were determined as anisotropic substructure-specific margins. Thus, anisotropic cardiac substructure-specific planning margins for OARs may be warranted to accommodate differential inter-fractional shifts. This work is significant in that quantifying the inter-fraction displacement of critical cardiac substructures is a first step in deriving substructure-specific safety margins to ensure highly effective cardiac sparing. Furthermore, the cardiac substructure segmentations from this study laid the groundwork for the second portion of Aim 4 of translating automated cardiac substructure segmentation into low-field MRgRT planning.

Aim 4, the final Aim, which was described in Chapter 6, investigated translating these previously described technologies to an MR-guided environment using an MR-linac and treatment planning comparisons. The goal of the first part of this Aim was to translate automated cardiac substructure segmentation into CT-based planning for low-field MRgRT and quantify potential dosimetric advantages for improved cardiac sparing through plan re-optimization and for cases that may benefit, potential beam angle modifications. It was found that while cardiac sparing plans did reduce the MHD (0.7 ± 0.6 Gy), whole heart dose metrics do not reflect the local dose deposition that the substructure metrics were able to capture and were not as sensitive. It was also found that four out of 16 patients benefited from new beam arrangements, leading to further dose reductions. However, the results were patient specific as lesion location and proximity to the heart played a role in if the patient would benefit from plan re-optimization and beam modification. Although beam angle modification was shown to provide improvements over solely re-optimizing the plan for select cases, re-optimization alone provided the majority of cardiac substructure sparing. The new treatment plans did not substantially increase delivery time or required monitor units ($P > 0.05$), suggesting a negligible increase in plan complexity when cardiac substructure sparing was introduced. Additionally, there was no statistical difference in clinical endpoints to the PTV, lung, esophagus, or spinal cord after re-optimization. Moreover, even though the modified beam angle plans involved either adding or removing beams in the revised treatment plan, the differences in treatment time per fraction and MUs delivered were negligible. This work is significant in that it provides dose assessments of the radiosensitive cardiac

substructures that are currently overlooked in the RT clinical setting, which may help better study and define radiation-induced cardiac injury.

Lastly, part 2 of Aim 4 outlined the development and validation of a deep learning cardiac substructure pipeline for MRgRT applications. Here, cardiac substructures were automatically segmented through a novel deep learning pipeline optimized for the principle goal of applications in MRgRT on low-field MRI. The use of daily MRIs for each patient provided a unique methodology for data augmentation and led to accurate automatic segmentations of the cardiac chambers, great vessels, and IVC. Additionally, even though the utilized patient images were at a decreased resolution, not contrast-enhanced, and were not cardiac gated, promising results were also yielded for the coronary arteries. When comparing to our previous study that utilized hybrid MR/CT information, cardiac chamber and great vessel results from the current study were within 0.04 DSC and coronary artery DSC results were within 0.08 DSC. In total, it was found that substructure contour generation for a new patient input took ~20 seconds. Centroid displacements were largest in the S-I axis likely due the larger slice thickness of 3 mm for each of the five test patients, as compared to the 1.5x1.5 mm in-plane resolution. This work is significant in that it opens the door to conduct studies on determining cardiac substructure dosimetric limitations through a prospective clinical trial using MRgRT. Additionally, strategies for cardiac avoidance can be considered.

Overall, we have provided the ability to automatically and efficiently generate accurate cardiac substructure segmentations on both non-contrast CT and low-field MRI, quantify their excursion in respiration and between daily fractions, as well as through dosimetric analysis, show the utility of their inclusion in radiation therapy planning.

Limitations and Future Work

Although much has been completed to incorporate cardiac substructures into RT for improved cardiac sparing, widespread implementation is still hindered by challenges that were outside the scope of this work. Additionally, there are some improvements that may be made upon this work including the consideration of cardiac motion, coronary artery segmentation accuracy, and further refining the deep learning methodology.

To build a dynamic motion model of the heart, a 3D MRI with cine information and high temporal resolution would be required. This type of data can be acquired during a cardiac angiography scan, but it is not the standard of care in RT. Thus, the available data and sequences are not optimized for the endpoint of creating a cardiac motion model and would possibly require a prospective clinical trial which could be completed in future work. Most of the patient data from the studies covered here come from a pre-existing Breast Cancer Research Foundation Study where the utilized sequences were optimized for an endpoint unrelated to our study. The MRIs used in this body of work, both the T2-weighted single-shot turbo spin echo and the TrueFISP sequence, are taken under breath-hold conditions and are not ECG gated (i.e. they do not provide temporal data across the cardiac cycle). Therefore, respiratory motion is assumed to be negligible during the scan. However, due to extended scan times and heart rate, several cardiac cycles are captured during the scan. Thus, the heart and its substructures on the T2-weighted and TrueFISP scans are represented by their average position over the course of the scan. For the purpose of this work, the average position of the cardiac substructures, with an added PRV generated from inter-fractional motion, provides a good

surrogate for representing how the cardiac substructures might be handled clinically and for treatment planning comparisons.

Clinically, cardiac motion in RT is not managed, as dose volume parameters for the whole heart are not significantly influenced by motion from the cardiac cycle[51]. Additionally, there are uncertainties in the magnitude of the long delivery time, which can be upwards of 10 minutes, versus the frequency of the beating heart, as well as the inter-fraction setup uncertainty between treatments. Yet, there is work being done to suggest planning organ at risk volumes (3-4 mm[33,53]) as the coronary arteries and the ventricles are the most mobile regions of the heart, displacing 3-8 mm between end-diastolic and end-systolic phases[52,247,248]. Nevertheless, researchers treating atrial fibrillation and ventricular tachycardia through cardiac ablation are not considering a motion model of the heart. For example, when conducting radio-ablation for atrial fibrillation, Maguire *et al.* utilized a Synchrony system to track respiratory motion, and then added a margin to the target volume to compensate for cardiac motion[265]. Moreover, several studies state that an extra margin, on top of the one placed for machine accuracy, does not need to be generated as the cardiac motion is small and has a likely negligible impact on dosimetry[198,265-269]. Dynamic motion modeling of the heart will, however, play a significant role in treatment planning when techniques like flash RT are used at an ultra-high dose rate.

The work presented here improved upon studies like that of Luo *et al.* where coronary arteries were not considered for automatic segmentation[270]. However, there is still room for improvement in the coronary artery segmentations presented here as DSCs were less than 0.60. This is complicated as shown by the several studies that have

experienced increased inter-observer variability when segmenting the coronary arteries versus other cardiac substructures. While generating a whole heart segmentation atlas, Feng *et al.* found when segmenting the heart and left ventricle, there were contour overlaps greater that 90%[141]. However, segmentation overlaps of the RCA were only 24% on average[141]. Similarly, Zhou *et al.* had a heightened variability in coronary artery segmentations than for the chambers and whole heart[132]. They postulated that fusions of non-gated CT scans will affect smaller structures more due to imperfect image resolution[132]. Luo *et al.* also claimed that segmenting these smaller structures may also be complicated by tumor invasion and calcification[270].

Although numerous measures were taken to advance automatic cardiac substructure segmentation, including multi-channel inputs, deep supervision, 3D-CRF, etc., there may still be room for improvement. In this work, the CRF were implemented as a post-processing step. In future work, the CRF may be directly embedded into a modified 3D U-Net to create an end-to-end workflow with improved efficiency. This would require that the CRF be coded as a recurrent network, placed in the final layer[271], and trained together with the 3D U-Net. As a known drawback of skip connections is that they duplicate low resolution contents, a residual pathway may be included through the use of a modified U-Net, much like the work by Seo *et al.*[97]. For substructure accuracy improvements in small structures like the coronary arteries, as suggested by Seo *et al.*, the number of pooling operations for small structures may also be reduced as compared to those of large structures due to the resolution loss that occurs after pooling operations[97]. The selection of the convergence (i.e. loss) function to be utilized in the deep learning architecture also will have an effect on the resultant automatic

segmentation. For example, a generalized Dice loss function may be used in place of a focal Dice loss function to only consider class imbalances due to size and not segmentation difficultly[183,184]. Additionally, loss functions not based on Dice, such as mean-squared pixelwise error, have shown promise for RT applications and may be explored in future work[272]. As it is a known issue in RT that the amount of labeled and/or paired samples can be severely limited, segmentation accuracy may also be improved by leveraging the large amount of unpaired, unlabeled, multi-modality cardiac datasets available at our institution. The use of a cycle-GAN[191] architecture would allow for valuable data without manually drawn ground truth segmentations being used in the training process. As this data was collected at various sequences, parameters, etc., unsupervised domain adaptation may be used which allows for implementation on various imaging platforms[273]. Specifically, generative pixel-level domain adaptation models perform similar distribution alignment in raw pixel space in order to translate source data to the target domain[274]. Thus, a deep learning model may be generated to train two competing networks simultaneously: a generator network to synthesize data and a discriminator network to distinguish between synthesized and real data, which often outperforms atlas and convolutional models[275]. Lastly, implementing optical flow techniques to both track motion and improve segmentation accuracy may be incorporated[276]. Although these presented ideas are outside the scope of this body of work and would require curated datasets, they may be implemented in the future.

This work may also be incorporated in a prospective clinical trial (viable for studies in lung, breast, lymphoma, esophageal cancers, etc.) to evaluate the associations between cardiac substructure dose and survival, cardiac events, and quality of life

measures in a diverse patient population. These data may then further reveal that radiation doses to cardiac substructures are more precisely associated with the incidence of cardiac events and overall survival than whole heart dose metrics, which is the current standard of care. A reduction in radiation-related cardiac toxicities may be acquired using the highly effective cardiac substructure sparing techniques mentioned in this work to maximize therapeutic gain. These results may also be extended to a trial using cardiac gated RT for avoidance or targeting of atrial fibrillation[198] and ventricular tachycardia[262,277]. Lastly, future directions may include evaluating changes in radiomics as imaging biomarkers from serial heart images and integrating functional MRI to facilitate functional sparing.

REFERENCES

1. Jaffray DA, Gospodarowicz MK. Radiation therapy for cancer. *Disease Control Priorities* 2015;3:239-247.

2. American cancer society. Cancer statistics center.

3. Gunderson LL, Tepper JE. Clinical radiation oncology: Elsevier Health Sciences; 2015.

4. Altena R, Perik PJ, Van Veldhuisen DJ, et al. Cardiovascular toxicity caused by cancer treatment: Strategies for early detection. *The lancet oncology* 2009;10:391-399.

5. Ng AK. Review of the cardiac long-term effects of therapy for hodgkin lymphoma. *British journal of haematology* 2011;154:23-31.

6. Beukema JC, van Luijk P, Widder J, et al. Is cardiac toxicity a relevant issue in the radiation treatment of esophageal cancer? *Radiotherapy and Oncology* 2015;114:85-90.

7. Hardy D, Liu C-C, Cormier J, et al. Cardiac toxicity in association with chemotherapy and radiation therapy in a large cohort of older patients with non-small-cell lung cancer. *Annals of oncology* 2010;21:1825-1833.

8. Correa CR, Litt HI, Hwang W-T, et al. Coronary artery findings after left-sided compared with right-sided radiation treatment for early-stage breast cancer. *Journal of Clinical Oncology* 2007;25:3031-3037.

9. Yusuf SW, Sami S, Daher IN. Radiation-induced heart disease: A clinical update. *Cardiology research and practice* 2011;2011.

10. Bluethmann SM, Mariotto AB, Rowland JH. Anticipating the “silver tsunami”: Prevalence trajectories and comorbidity burden among older cancer survivors in the united states. In: Editor, editor^editors. Book Anticipating the “silver tsunami”: Prevalence trajectories and comorbidity burden among older cancer survivors in the united states: AACR; 2016.

11. Haasbeek CJ, Lagerwaard FJ, Slotman BJ, et al. Outcomes of stereotactic ablative radiotherapy for centrally located early-stage lung cancer. *Journal of Thoracic Oncology* 2011;6:2036-2043.

12. Ming X, Feng Y, Yang C, et al. Radiation-induced heart disease in lung cancer radiotherapy: A dosimetric update. *Medicine* 2016;95.

13. Marks LB, Yu X, Prosnitz RG, et al. The incidence and functional consequences of rt-associated cardiac perfusion defects. *International Journal of Radiation Oncology* Biology* Physics* 2005;63:214-223.

14. Gomez DR, Yusuf SW, Munsell MF, et al. Prospective exploratory analysis of cardiac biomarkers and electrocardiogram abnormalities in patients receiving thoracic radiation therapy with high-dose heart exposure. *Journal of Thoracic Oncology* 2014;9:1554-1560.

15. Walls G, Lyon A, Harbinson M, et al. Cardiotoxicity following cancer treatment. *The Ulster medical journal* 2017;86:3.

16. Darby SC, Ewertz M, McGale P, et al. Risk of ischemic heart disease in women after radiotherapy for breast cancer. *New England Journal of Medicine* 2013;368:987-998.

17. Bentzen SM, Constine LS, Deasy JO, et al. Quantitative analyses of normal tissue effects in the clinic (quantec): An introduction to the scientific issues. *International Journal of Radiation Oncology* Biology* Physics* 2010;76:S3-S9.

18. Gagliardi G, Constine LS, Moiseenko V, et al. Radiation dose–volume effects in the heart. *International Journal of Radiation Oncology* Biology* Physics* 2010;76:S77-S85.

19. Dmitriev I, Loo C, Vogel W, et al. Fully automated deformable registration of breast dce-mri and pet/ct. *Physics in medicine and biology* 2013;58:1221.

20. Kong F, Machtay M, Bradley J, et al. Rtog 1106/acrin 6697: Randomized phase ii trial of individualized adaptive radiotherapy using during treatment fdg-pet. *CT and modern technology in locally advanced non-small lung cancer (NSCLC)* 2012.

21. Nrg-gi006, phase iii randomized trial of proton beam therapy (pbt) versus intensity modulated photon radiotherapy (imrt) for the treatment of esophageal cancer. In: Editor, editor^editors. Book Nrg-gi006, phase iii randomized trial of proton beam therapy (pbt) versus intensity modulated photon radiotherapy (imrt) for the treatment of esophageal cancer; 2018.

22. Taylor CW, Wang Z, Macaulay E, et al. Exposure of the heart in breast cancer radiation therapy: A systematic review of heart doses published during 2003 to 2013. *International Journal of Radiation Oncology* Biology* Physics* 2015;93:845-853.

23. Aupérin A, Le Péchoux C, Rolland E, et al. Meta-analysis of concomitant versus sequential radiochemotherapy in locally advanced non-small-cell lung cancer

Database of abstracts of reviews of effects (dare): Quality-assessed reviews [internet]: Centre for Reviews and Dissemination (UK); 2010.

24. Ma L, Men Y, Feng L, et al. A current review of dose-escalated radiotherapy in locally advanced non-small cell lung cancer. *Radiology and oncology* 2019;53:6.
25. Machtay M, Paulus R, Moughan J, et al. Defining local-regional control and its importance in locally advanced non-small cell lung carcinoma. *Journal of Thoracic Oncology* 2012;7:716-722.
26. Bradley JD, Paulus R, Komaki R, et al. Standard-dose versus high-dose conformal radiotherapy with concurrent and consolidation carboplatin plus paclitaxel with or without cetuximab for patients with stage iiia or iiib non-small-cell lung cancer (rtog 0617): A randomised, two-by-two factorial phase 3 study. *The Lancet Oncology* 2015;16:187-199.
27. Movsas B, Hu C, Sloan J, et al. Quality of life analysis of a radiation dose–escalation study of patients with non–small-cell lung cancer: A secondary analysis of the radiation therapy oncology group 0617 randomized clinical trial. *JAMA Oncology* 2016;2:359-367.
28. van den Bogaard V, Ta B, van der Schaaf A, et al. Validation and modification of a prediction model for acute cardiac events in patients with breast cancer treated with radiotherapy based on three-dimensional dose distributions to cardiac substructures. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology* 2017;35:1171.

29. Vivekanandan S, Landau D, Counsell N, et al. The impact of cardiac radiation dosimetry on survival after radiation therapy for non-small cell lung cancer. *International Journal of Radiation Oncology* Biology* Physics* 2017;99:51-60.

30. Poitevin-Chacón A, Chávez-Nogueda J, Prudencio RR, et al. Dosimetry of the left anterior descending coronary artery in left breast cancer patients treated with postoperative external radiotherapy. *Reports of Practical Oncology & Radiotherapy* 2018;23:91-96.

31. Wang X, Pan T, Pinnix C, et al. Cardiac motion during deep-inspiration breath-hold: Implications for breast cancer radiotherapy. *International Journal of Radiation Oncology* Biology* Physics* 2012;82:708-714.

32. Patel S, Mahmood S, Nguyen T, et al. Comparing whole heart versus coronary artery dosimetry in predicting the risk of cardiac toxicity following breast radiation therapy. *International Journal of Radiation Oncology* Biology* Physics* 2018;102:S46.

33. Kataria T, Bisht SS, Gupta D, et al. Quantification of coronary artery motion and internal risk volume from ecg gated radiotherapy planning scans. *Radiotherapy and Oncology* 2016;121:59-63.

34. Hahn E, Jiang H, Ng A, et al. Late cardiac toxicity after mediastinal radiation therapy for hodgkin lymphoma: Contributions of coronary artery and whole heart dose-volume variables to risk prediction 2017;98:1116-1123.

35. McWilliam A, Kennedy J, Hodgson C, et al. Radiation dose to heart base linked with poorer survival in lung cancer patients. *European Journal of Cancer* 2017;85:106-113.

36. Wang K, Pearlstein KA, Patchett ND, et al. Heart dosimetric analysis of three types of cardiac toxicity in patients treated on dose-escalation trials for stage iii non-small-cell lung cancer 2017;125:293-300.

37. Stam B, Peulen H, Guckenberger M, et al. Dose to heart substructures is associated with non-cancer death after sbrt in stage i–ii nsclc patients. *Radiotherapy Oncology* 2017;123:370-375.

38. Hancock SL, Tucker MA, Hoppe RT. Factors affecting late mortality from heart disease after treatment of hodgkin's disease. *Jama* 1993;270:1949-1955.

39. Carmel RJ, Kaplan HS. Mantle irradiation in hodgkin's disease. An analysis of technique, tumor eradication, and complications. *Cancer* 1976;37:2813-2825.

40. Thor M, Deasy J, Hu C, et al. The role of heart-related dose-volume metrics on overall survival in the rtog 0617 clinical trial. *International Journal of Radiation Oncology• Biology• Physics* 2018;102:S96.

41. Thor M, Hotca A, Jackson A, et al. Sp-0112: Dose to cardiac substructures predicts survival in non-small cell lung cancer chemo-radiotherapy. *Radiotherapy and Oncology* 2018;127:S58-S59.

42. Gore E, Hu C, Ad VB, et al. Impact of incidental cardiac radiation on cardiopulmonary toxicity and survival for locally advanced non-small cell lung cancer: Reanalysis of nrg oncology/rtog 0617 with centrally contoured cardiac structures. *International Journal of Radiation Oncology* Biology* Physics* 2016;96:S129-30.

43. Attili AK, Schuster A, Nagel E, et al. Quantification in cardiac mri: Advances in image acquisition and processing. *The international journal of cardiovascular imaging* 2010;26:27-40.

44. Dweck MR, Puntmann VO, Vesey AT, et al. Mr imaging of coronary arteries and plaques. *JACC: Cardiovascular Imaging* 2016;9:306-316.

45. Kirişli H, Schaap M, Klein S, et al. Evaluation of a multi-atlas based method for segmentation of cardiac cta data: A large-scale, multicenter, and multivendor study. *Medical physics* 2010;37:6279-6291.

46. Zhuang X, Bai W, Song J, et al. Multiatlas whole heart segmentation of ct data using conditional entropy for atlas ranking and selection. *Medical physics* 2015;42:3822-3833.

47. Zhuang X, Shen J. Multi-scale patch and multi-modality atlases for whole heart segmentation of mri. *Medical Image Analysis* 2016;31:77-87.

48. Morris ED, Ghanem AI, Pantelic MV, et al. Cardiac substructure segmentation and dosimetry using a novel hybrid magnetic resonance and computed tomography cardiac atlas. *International Journal of Radiation Oncology* Biology* Physics* 2019;103:985-993.

49. Jagsi R, Moran JM, Kessler ML, et al. Respiratory motion of the heart and positional reproducibility under active breathing control. *International Journal of Radiation Oncology* Biology* Physics* 2007;68:253-258.

50. Guzhva L, Flampouri S, Mendenhall NP, et al. Intrafractional displacement of cardiac substructures among patients with mediastinal lymphoma or lung cancer. *Advances in radiation oncology* 2019;4:500-506.

51. Tong Y, Yin Y, Cheng P, et al. Quantification of variation in dose–volume parameters for the heart, pericardium and left ventricular myocardium during thoracic tumor radiotherapy 2018;59:462-468.

52. Tan W, Xu L, Wang X, et al. Estimation of the displacement of cardiac substructures and the motion of the coronary arteries using electrocardiographic gating. *Onco Targets Ther* 2013;6:1325-1332.

53. Li Q, Tong Y, Yin Y, et al. Definition of the margin of major coronary artery bifurcations during radiotherapy with electrocardiograph-gated 4d-ct. *Physica Medica* 2018;49:90-94.

54. Bissonnette J-P, Moseley D, White E, et al. Quality assurance for the geometric accuracy of cone-beam ct guidance in radiation therapy. *International Journal of Radiation Oncology* Biology* Physics* 2008;71:S57-S61.

55. Schneider U, Pedroni E, Lomax A. The calibration of ct hounsfield units for radiotherapy treatment planning. *Physics in medicine and biology* 1996;41:111-24.

56. Herrlin K, Ling LB, Pettersson H, et al. Gadolinium-dtpa enhancement of soft tissue tumors in magnetic resonance imaging 1990;31:233-236.

57. Brown RW, Cheng Y-CN, Haacke EM, et al. Magnetic resonance imaging: Physical principles and sequence design: John Wiley & Sons; 2014.

58. Khoo VS, Dearnaley DP, Finnigan DJ, et al. Magnetic resonance imaging (mri): Considerations and applications in radiotherapy treatment planning 1997;42:1-15.

59. Krempien RC, Schubert K, Zierhut D, et al. Open low-field magnetic resonance imaging in radiation therapy treatment planning. *International Journal of Radiation Oncology* Biology* Physics* 2002;53:1350-1360.

60. Debois M, Oyen R, Maes F, et al. The contribution of magnetic resonance imaging to the three-dimensional treatment planning of localized prostate cancer. *International Journal of Radiation Oncology* Biology* Physics* 1999;45:857-865.

61. Li BN, Chui CK, Chang S, et al. A new unified level set method for semi-automatic liver tumor segmentation on contrast-enhanced ct images. *Expert Systems with Applications* 2012;39:9661-9668.

62. Njeh C. Tumor delineation: The weakest link in the search for accuracy in radiotherapy. *Journal of medical physics/Association of Medical Physicists of India* 2008;33:136.

63. Weiss E, Hess CF. The impact of gross tumor volume (gtv) and clinical target volume (ctv) definition on the total accuracy in radiotherapy theoretical aspects and practical experiences. *Strahlentherapie und Onkologie : Organ der Deutschen Rontgengesellschaft [et al]* 2003;179:21-30.

64. Cazzaniga LF, Marinoni MA, Bossi A, et al. Interphysician variability in defining the planning target volume in the irradiation of prostate and seminal vesicles. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology* 1998;47:293-6.

65. Weltens C, Menten J, Feron M, et al. Interobserver variations in gross tumor volume delineation of brain tumors on computed tomography and impact of magnetic

resonance imaging. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology* 2001;60:49-59.

66. Valley JF, Bernier J, Tercier PA, et al. Quality assurance of the eortc radiotherapy trial 22931 for head and neck carcinomas: The dummy run. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology* 1998;47:37-44.
67. Emami B, Sethi A, Petruzzelli GJ. Influence of mri on target volume delineation and imrt planning in nasopharyngeal carcinoma. *International journal of radiation oncology, biology, physics* 2003;57:481-8.
68. Chung NN, Ting LL, Hsu WC, et al. Impact of magnetic resonance imaging versus ct on nasopharyngeal carcinoma: Primary tumor target delineation for radiotherapy. *Head & neck* 2004;26:241-6.
69. Gurney-Champion OJ, Versteijne E, van der Horst A, et al. Addition of mri for ct-based pancreatic tumor delineation: A feasibility study. *Acta Oncologica* 2017;56:923-930.
70. den Hartogh MD, Philippens ME, van Dam IE, et al. Mri and ct imaging for preoperative target volume delineation in breast-conserving therapy. *Radiation Oncology* 2014;9:63.
71. Segedin B, Petric P. Uncertainties in target volume delineation in radiotherapy–are they relevant and what can we do about them? *Radiology and oncology* 2016;50:254-262.
72. Ritter T, Quint DJ, Senan S, et al. Consideration of dose limits for organs at risk of thoracic radiotherapy: Atlas for lung, proximal bronchial tree, esophagus, spinal

cord, ribs, and brachial plexus. *International Journal of Radiation Oncology* Biology* Physics* 2011;81:1442-1457.

73. Bainbridge H, Salem A, Tijssen RH, et al. Magnetic resonance imaging in precision radiation therapy for lung cancer. *Translational lung cancer research* 2017;6:689.
74. Wachter S, Wachter-Gerstner N, Bock T, et al. Interobserver comparison of ct and mri-based prostate apex definition clinical relevance for conformal radiotherapy treatment planning. *Strahlentherapie und Onkologie* 2002;178:263-268.
75. Khoo VS, Padhani AR, Tanner SF, et al. Comparison of mri with ct for the radiotherapy planning of prostate cancer: A feasibility study. *The British journal of radiology* 1999;72:590-7.
76. Lauzon ML, Mahallati H, Frayne R. Time-efficient breath-hold abdominal mri at 3.0 t. *American Journal of Roentgenology* 2006;187:649-657.
77. Jung G, Krahe T, Kugel H, et al. Comparison of fast turbo-spin-echo and gradient- and spin-echo sequences as well as echo planar imaging with conventional spin-echo sequences in mri of focal liver lesions at 1.0 tesla. *RoFo: Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin* 1996;165:137-143.
78. Rietzel E, Pan T, Chen GT. Four-dimensional computed tomography: Image formation and clinical protocol. *Medical physics* 2005;32:874-889.
79. Lu W, Parikh PJ, Hubenschmidt JP, et al. A comparison between amplitude sorting and phase-angle sorting using external respiratory measurement for 4d ct. *Medical physics* 2006;33:2964-2974.
80. Mieres J, Makaryus AN, Redberg R, et al. Noninvasive cardiac imaging. *American family physician* 2007;75:1219-1228.

81. Edelman RR, Botelho M, Pursnani A, et al. Improved dark blood imaging of the heart using radial balanced steady-state free precession. *Journal of Cardiovascular Magnetic Resonance* 2017;18:69.

82. Abdel-Aty H, Zagrosek A, Schulz-Menger J, et al. Delayed enhancement and t2-weighted cardiovascular magnetic resonance imaging differentiate acute from chronic myocardial infarction. *Circulation* 2004;109:2411-2416.

83. Kwong RY, Schussheim AE, Rekhraj S, et al. Detecting acute coronary syndrome in the emergency department with cardiac magnetic resonance imaging. *Circulation* 2003;107:531-537.

84. Raman SV, Simonetti OP, Winner MW, et al. Cardiac magnetic resonance with edema imaging identifies myocardium at risk and predicts worse outcome in patients with non–st-segment elevation acute coronary syndrome. *Journal of the American College of Cardiology* 2010;55:2480-2488.

85. Eitel I, Friedrich MG. T2-weighted cardiovascular magnetic resonance in acute cardiac disease. *Journal of Cardiovascular Magnetic Resonance* 2011;13:13.

86. Aletras AH, Tilak GS, Natanzon A, et al. Clinical perspective. *Circulation* 2006;113:1865-1870.

87. Tilak GS, Hsu L-Y, Hoyt Jr RF, et al. In vivo t2-weighted magnetic resonance imaging can accurately determine the ischemic area at risk for 2-day-old nonreperfused myocardial infarction. *Investigative radiology* 2008;43:7-15.

88. Holland AE, Goldfarb JW, Edelman RR. Diaphragmatic and cardiac motion during suspended breathing: Preliminary experience and implications for breath-hold mr imaging. *Radiology* 1998;209:483-489.

89. Schernthaner RE, Stadler A, Beitzke D, et al. Dose modulated retrospective ecg-gated versus non-gated 64-row ct angiography of the aorta at the same radiation dose: Comparison of motion artifacts, diagnostic confidence and signal-to-noise-ratios. *European journal of radiology* 2012;81:e585-e590.

90. Michael MA, El Masry H, Khan BR, et al. Electrocardiographic signs of remote myocardial infarction. *Progress in cardiovascular diseases* 2007;50:198-208.

91. Gweon HM, Kim SJ, Kim TH, et al. Evaluation of reperfused myocardial infarction by low-dose multidetector computed tomography using prospective electrocardiography (ecg)-triggering: Comparison with magnetic resonance imaging. *Yonsei medical journal* 2010;51:683-691.

92. Mortazi A, Burt J, Bagci U. Multi-planar deep segmentation networks for cardiac substructures from mri and ct. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young A, Bernard O, editors. Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges. Cham. Springer International Publishing. 2018. pp. 199-206.

93. Wong PM, Padley SP. Cardiac computed tomography. *Medicine* 2018;46:474-479.

94. Wu W, Budovec J, Foley WD. Prospective and retrospective ecg gating for thoracic ct angiography: A comparative study. *American Journal of Roentgenology* 2009;193:955-963.

95. Ghosh S, Das N, Das I, et al. Understanding deep learning techniques for image segmentation. *ACM Computing Surveys (CSUR)* 2019;52:1-35.

96. Zaidi H, El Naqa I. Pet-guided delineation of radiation therapy treatment volumes: A survey of image segmentation techniques. *European journal of nuclear medicine and molecular imaging* 2010;37:2165-2187.

97. Seo H, Khuzani MB, Vasudevan V, et al. Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications. *arXiv preprint arXiv:191102521* 2019.

98. Mao KZ, Zhao P, Tan P-H. Supervised learning-based cell image segmentation for p53 immunohistochemistry. *IEEE Transactions on Biomedical Engineering* 2006;53:1153-1163.

99. Feng-Ping A, Zhi-Wen L. Medical image segmentation algorithm based on feedback mechanism convolutional neural network. *Biomedical Signal Processing and Control* 2019;53:101589.

100. Pal NR, Pal SK. A review on image segmentation techniques. *Pattern recognition* 1993;26:1277-1294.

101. Bhabatosh C. Digital image processing and analysis: PHI Learning Pvt. Ltd.; 2011.

102. Rohlfing T, Brandt R, Menzel R, et al. Quo vadis, atlas-based segmentation? Handbook of biomedical image analysis: Springer; 2005. pp. 435-486.

103. Cabezas M, Oliver A, Lladó X, et al. A review of atlas-based segmentation for magnetic resonance brain images. *Computer methods and programs in biomedicine* 2011;104:e158-e177.

104. Iglesias JE, Sabuncu MR. Multi-atlas segmentation of biomedical images: A survey. *Medical image analysis* 2015;24:205-219.

105. Hesamian MH, Jia W, He X, et al. Deep learning techniques for medical image segmentation: Achievements and challenges. *Journal of digital imaging* 2019;32:582-596.

106. Baillard C, Hellier P, Barillot C. Segmentation of brain 3d mr images using level sets and dense registration. *Medical image analysis* 2001;5:185-194.

107. Collins DL, Evans AC. Animal: Validation and applications of nonlinear registration-based segmentation. *International journal of pattern recognition and artificial intelligence* 1997;11:1271-1294.

108. Crum WR, Scahill RI, Fox NC. Automated hippocampal segmentation by regional fluid registration of serial mri: Validation and application in alzheimer's disease. *Neuroimage* 2001;13:847-855.

109. Dawant BM, Hartmann SL, Thirion J-P, et al. Automatic 3-d segmentation of internal structures of the head in mr images using a combination of similarity and free-form transformations. I. Methodology and validation on normal subjects. *IEEE transactions on medical imaging* 1999;18:909-916.

110. Musse O, Heitz F, Armspach J-P. Fast deformable matching of 3d images over multiscale nested subspaces. Application to atlas-based mri segmentation. *Pattern Recognition* 2003;36:1881-1899.

111. Gee JC, Reivich M, Bajcsy R. Elastically deforming a three-dimensional atlas to match anatomical brain images 1993.

112. Nguyen G, Dlugolinsky S, Bobák M, et al. Machine learning and deep learning frameworks and libraries for large-scale data mining: A survey. *Artificial Intelligence Review* 2019;52:77-124.

113. Krähenbühl P, Koltun V. Efficient inference in fully connected crfs with gaussian edge potentials. Advances in neural information processing systems. 2011. pp. 109-117.

114. Thimm G, Fiesler E. High-order and multilayer perceptron initialization. *IEEE Transactions on Neural Networks* 1997;8:349-359.

115. Jaeger H. Artificial neural networks. *Talk* 2008;17:281-289.

116. Amit G, Hadad O, Alpert S, et al. Hybrid mass detection in breast mri combining unsupervised saliency analysis and deep learning. International conference on medical image computing and computer-assisted intervention. Springer. 2017. pp. 594-602.

117. Emami H, Dong M, Nejad-Davarani SP, et al. Generating synthetic cts from magnetic resonance images using generative adversarial networks. *Medical Physics* 2018;45:3627-3636.

118. Chen C, Qin C, Qiu H, et al. Deep learning for cardiac image segmentation: A review. *Frontiers in Cardiovascular Medicine* 2020;7:25.

119. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. Springer. 2015. pp. 234-241.

120. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Medical Image Analysis* 2017;42:60-88.

121. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition. 2015. pp. 3431-3440.

122. Ibtehaz N, Rahman MS. Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation. *Neural Networks* 2020;121:74-87.

123. Feng M, Valdes G, Dixit N, et al. Machine learning in radiation oncology: Opportunities, requirements, and needs. *Frontiers in oncology* 2018;8:110.

124. Delpon G, Escande A, Ruef T, et al. Comparison of automated atlas-based segmentation software for postoperative prostate cancer radiotherapy. *Frontiers in oncology* 2016;6:178.

125. LeCun Y, Bengio Y, Hinton G. Deep learning. *nature* 2015;521:436-444.

126. Men K, Dai J, Li Y. Automatic segmentation of the clinical target volume and organs at risk in the planning ct for rectal cancer using deep dilated convolutional neural networks. *Medical physics* 2017;44:6377-6389.

127. Yang X, Wu N, Cheng G, et al. Automated segmentation of the parotid gland based on atlas registration and machine learning: A longitudinal mri study in head-and-neck radiation therapy. *International Journal of Radiation Oncology* Biology* Physics* 2014;90:1225-1233.

128. Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck ct images using convolutional neural networks. *Medical physics* 2017;44:547-557.

129. Jacob S, Pathak A, Franck D, et al. Early detection and prediction of cardiotoxicity after radiation therapy for breast cancer: The baccarat prospective cohort study. *Radiation Oncology* 2016;11:54.

130. Wright J, Blitzblau R. Cardiac morbidity and radiation therapy for breast cancer. *International Journal of Radiation Oncology• Biology• Physics* 2018;100:283-286.

131. Urbanic JJ, McGarry RC, Daly ME, et al. Locally advanced lung cancer: Is it time to take cardiac protection seriously in radiation planning? *International Journal of Radiation Oncology• Biology• Physics* 2018;100:1091-1094.

132. Zhou R, Liao Z, Pan T, et al. Cardiac atlas development and validation for automatic segmentation of cardiac substructures. *Radiotherapy and Oncology* 2017;122:66-71.

133. Thor M, Yorke E, Jackson A, et al. Succesful implementation of atlas segmentation for cardiac substructures. *International Journal of Radiation Oncology* Biology* Physics* 2017;99:E728.

134. Vennarini S, Fournier-Bidoz N, Aristei C, et al. Visualisation of the left anterior descending coronary artery on ct images used for breast radiotherapy planning. *The British journal of radiology* 2013;86:20120643.

135. Gradishar WJ, Anderson BO, Balassanian R, et al. Nccn guidelines insights: Breast cancer, version 1.2017. *Journal of the National Comprehensive Cancer Network* 2017;15:433-451.

136. Freedman GM, Lin L. Cardiac-sparing radiation therapy for breast cancer.

137. Ipsen S, Blanck O, Oborn B, et al. Radiotherapy beyond cancer: Target localization in real-time mri and treatment planning for cardiac radiosurgery. *Medical physics* 2014;41.

138. Julia White AT, Douglas Arthur, Thomas Buchholz, Shannon MacDonald, Lawrence Marks, Lori Pierce, Abraham Recht, Rachel Rabinovitch, Alphonse Taghian, Frank Vicini, Wendy Woodward, X. Allen Li. Radiation therapy oncology group breast cancer atlas for radiation therapy planning: Consensus definitions.

Available at https://www.Rtog.Org/linkclick.Aspx?Fileticket=vzjfhpabipe%3d&tabid=236.

139. Peng P, Lekadir K, Gooya A, et al. A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging. *Magnetic Resonance Materials in Physics, Biology and Medicine* 2016;29:155-195.

140. Pluim JP, Maintz JA, Viergever MA. Mutual-information-based registration of medical images: A survey. *IEEE transactions on medical imaging* 2003;22:986-1004.

141. Feng M, Moran JM, Koelling T, et al. Development and validation of a heart atlas to study cardiac exposure to radiation following treatment for breast cancer. *International Journal of Radiation Oncology* Biology* Physics* 2011;79:10-18.

142. Salkind NJ. Encyclopedia of measurement and statistics: Sage; 2007.

143. Horvat M, Nelson A, Pirozzi S. Time savings of a multi-atlas approach for liver segmentation. *Journal of Nuclear Medicine* 2014;55:1523-1523.

144. Young AV, Wortham A, Wernick I, et al. Atlas-based segmentation improves consistency and decreases time required for contouring postoperative endometrial cancer nodal volumes. *International Journal of Radiation Oncology* Biology* Physics* 2011;79:943-947.

145. Piper J. Su-ff-i-68: Evaluation of an intensity-based free-form deformable registration algorithm. *Medical Physics* 2007;34:2353-2354.

146. Kirby N, Chuang C, Ueda U, et al. The need for application-based adaptation of deformable image registration. *Medical physics* 2013;40.

147. Warfield SK, Zou KH, Wells WM. Simultaneous truth and performance level estimation (staple): An algorithm for the validation of image segmentation. *IEEE transactions on medical imaging* 2004;23:903-921.

148. Thomson D, Boylan C, Liptrot T, et al. Evaluation of an automatic segmentation algorithm for definition of head and neck organs at risk. *Radiation Oncology* 2014;9:173.

149. Kumarasiri A, Siddiqui F, Liu C, et al. Deformable image registration based automatic ct-to-ct contour propagation for head and neck adaptive radiotherapy in the routine clinical setting. *Medical physics* 2014;41.

150. Sato K, Mizuno Y, Fuchikami H, et al. Comparison of radiation dose to the left anterior descending artery by whole and partial breast irradiation in breast cancer patients. *Journal of contemporary brachytherapy* 2015;7:23.

151. Remouchamps VM, Letts N, Vicini FA, et al. Initial clinical experience with moderate deep-inspiration breath hold using an active breathing control device in the treatment of patients with left-sided breast cancer using external beam radiation therapy. *International Journal of Radiation Oncology* Biology* Physics* 2003;56:704-715.

152. Duane F, Aznar MC, Bartlett F, et al. A cardiac contouring atlas for radiotherapy. *Radiotherapy and Oncology* 2017;122:416-422.

153. Tran PV. A fully convolutional neural network for cardiac segmentation in short-axis mri. *arXiv preprint arXiv:160400494* 2016.

154. Nieder C, Schill S, Kneschaurek P, et al. Influence of different treatment techniques on radiation dose to the lad coronary artery. *Radiation Oncology* 2007;2:20.

155. Kaderka R, Gillespie EF, Mundt RC, et al. Geometric and dosimetric evaluation of atlas based auto-segmentation of cardiac structures in breast cancer patients. *Radiotherapy and Oncology* 2018.

156. van den Bogaard VA, van Dijk LV, Vliegenthart R, et al. Development and evaluation of an auto-segmentation tool for the left anterior descending coronary artery of breast cancer patients based on anatomical landmarks. *Radiotherapy and Oncology* 2019;136:15-20.

157. Payer C, Štern D, Bischof H, et al. Multi-label whole heart segmentation using cnns and anatomical label configurations. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young A, Bernard O, editors. Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges. Cham. Springer International Publishing. 2018. pp. 190-198.

158. Zhang Z, Yang L, Zheng Y. Translating and segmenting multimodal medical volumes with cycle-and shapeconsistency generative adversarial network. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018. pp. 9242-9251.

159. Haq R, Hotca A, Apte A, et al. Cardio-pulmonary substructure segmentation of ct images using convolutional neural networks. In: Nguyen D, Xing L, Jiang S, editors. Artificial Intelligence in Radiation Therapy. Cham. Springer International Publishing. 2019. pp. 162-169.

160. Kamnitsas K, Ledig C, Newcombe VF, et al. Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. *Medical Image Analysis* 2017;36:61-78.

161. Baughman DR, Liu YA. Neural networks in bioprocessing and chemical engineering: Academic press; 2014.

162. Isensee F, Kickingereder P, Wick W, et al. Brain tumor segmentation and radiomics survival prediction: Contribution to the brats 2017 challenge. In: Crimi A, Bakas S, Kuijf H, Menze B, Reyes M, editors. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Cham. Springer International Publishing. 2018. pp. 287-297.

163. Lee C-Y, Xie S, Gallagher P, et al. Deeply-supervised nets. Artificial intelligence and statistics. 2015. pp. 562-570.

164. Zeng G, Yang X, Li J, et al. 3d u-net with multi-level deep supervision: Fully automatic segmentation of proximal femur in 3d mr images. In: Wang Q, Shi Y, Suk H-I, Suzuki K, editors. Machine Learning in Medical Imaging. Cham. Springer International Publishing. 2017. pp. 274-282.

165. Mao X, Shen C, Yang Y-B. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. Advances in neural information processing systems. 2016. pp. 2802-2810.

166. Men K, Chen X, Zhang Y, et al. Deep deconvolutional neural network for target segmentation of nasopharyngeal cancer in planning computed tomography images. *Frontiers in Oncology* 2017;7:315.

167. Milletari F, Navab N, Ahmadi S-A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. 3D Vision (3DV), 2016 Fourth International Conference on. IEEE. 2016. pp. 565-571.

168. Nair AA, Tran TD, Reiter A, et al. A deep learning based alternative to beamforming ultrasound images. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE. 2018. pp. 3359-3363.

169. Kingma DP, Ba J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:14126980* 2014.

170. Larsson M, Alvén J, Kahl F. Max-margin learning of deep structured models for semantic segmentation. In: Sharma P, Bianchi FM, editors. Image Analysis. Cham. Springer International Publishing. 2017. pp. 28-40.

171. Paris S, Durand F. A fast approximation of the bilateral filter using a signal processing approach. In: Leonardis A, Bischof H, Pinz A, editors. Computer Vision – ECCV 2006. Berlin, Heidelberg. Springer Berlin Heidelberg. 2006. pp. 568-580.

172. Shotton J, Winn J, Rother C, et al. Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. *International Journal of Computer Vision* 2009;81:2-23.

173. Zheng S, Jayasumana S, Romera-Paredes B, et al. Conditional random fields as recurrent neural networks. Proceedings of the IEEE international conference on computer vision. 2015. pp. 1529-1537.

174. Zhao X, Wu Y, Song G, et al. A deep learning model integrating fcnns and crfs for brain tumor segmentation. *Medical Image Analysis* 2018;43:98-111.

175. Zhang J, Nie H. A post-processing method based on fully connected crfs for chronic wound images segmentation and identification: NUDT; 2018.

176. Dice LR. Measures of the amount of ecologic association between species. *Ecology* 1945;26:297-302.

177. Lustberg T, van Soest J, Gooding M, et al. Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer 2018;126:312-317.

178. Aleman B, van den Belt-Dusebout AW, Klokman WJ, et al. Long-term cause-specific mortality of patients treated for hodgkin's lymphoma 2003;21:3431-9.

179. Aljabar P, Gooding M. The cutting edge: Delineating contours with deep learning. *Machine Learning* 2001;2005:2013.

180. Ben-Cohen A, Klang E, Amitai MM, et al. Anatomical data augmentation for cnn based pixel-wise classification. Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on. IEEE. 2018. pp. 1096-1099.

181. Frid-Adar M, Diamant I, Klang E, et al. Gan-based synthetic medical image augmentation for increased cnn performance in liver lesion classification. *Neurocomputing* 2018;321:321-331.

182. Kjerland Ø. Segmentation of coronary arteries from ct-scans of the heart using deep learning: MS Thesis. NTNU; 2017.

183. Crum WR, Camara O, Hill DL. Generalized overlap measures for evaluation and validation in medical image analysis. *IEEE transactions on medical imaging* 2006;25:1451-1461.

184. Sudre CH, Li W, Vercauteren T, et al. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations Deep learning in medical

image analysis and multimodal learning for clinical decision support: Springer; 2017. pp. 240-248.

185. Fu Y, Mazur TR, Wu X, et al. A novel mri segmentation method using cnn-based correction network for mri-guided adaptive radiotherapy. *Medical Physics* 2018;45:5129-5137.

186. Rajchl M, Lee MCH, Oktay O, et al. Deepcut: Object segmentation from bounding box annotations using convolutional neural networks. *IEEE Transactions on Medical Imaging* 2017;36:674-683.

187. Kamnitsas K, Bai W, Ferrante E, et al. Ensembles of multiple models and architectures for robust brain tumour segmentation. In: Crimi A, Bakas S, Kuijf H, Menze B, Reyes M, editors. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Cham. Springer International Publishing. 2018. pp. 450-462.

188. Jiang H, Guo Y. Multi-class multimodal semantic segmentation with an improved 3d fully convolutional networks. *Neurocomputing* 2019.

189. Liu Z, Li X, Luo P, et al. Semantic image segmentation via deep parsing network. Proceedings of the IEEE international conference on computer vision. 2015. pp. 1377-1385.

190. Trullo R, Petitjean C, Ruan S, et al. Segmentation of organs at risk in thoracic ct images using a sharpmask architecture and conditional random fields. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). IEEE. 2017. pp. 1003-1006.

191. Zhu J-Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of the IEEE international conference on computer vision. 2017. pp. 2223-2232.

192. Huang X, Li Y, Poursaeed O, et al. Stacked generative adversarial networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017. pp. 5077-5086.

193. Hull MC, Morris CG, Pepine CJ, et al. Valvular dysfunction and carotid, subclavian, and coronary artery disease in survivors of hodgkin lymphoma treated with radiation therapy. *Jama* 2003;290:2831-2837.

194. Rutqvist LE, Lax I, Fornander T, et al. Cardiovascular mortality in a randomized trial of adjuvant radiation therapy versus surgery alone in primary breast cancer. *International Journal of Radiation Oncology• Biology• Physics* 1992;22:887-896.

195. Hochhegger B, Marchiori E, Sedlaczek O, et al. Mri in lung cancer: A pictorial essay. *The British journal of radiology* 2011;84:661-668.

196. Biederer J, Beer M, Hirsch W, et al. Mri of the lung (2/3). Why… when… how? *Insights into imaging* 2012;3:355-371.

197. Hochhegger B, Marchiori E, Irion K, et al. Magnetic resonance of the lung: A step forward in the study of lung disease. *J Bras Pneumol* 2012;38:105-15.

198. Blanck O, Ipsen S, Chan MK, et al. Treatment planning considerations for robotic guided cardiac radiosurgery for atrial fibrillation. *Cureus* 2016;8.

199. Matsuo Y, Onishi H, Nakagawa K, et al. Guidelines for respiratory motion management in radiation therapy. *Journal of radiation research* 2013;54:561-568.

200. Beck RE, Kim L, Yue NJ, et al. Treatment techniques to reduce cardiac irradiation for breast cancer patients treated with breast-conserving surgery and radiation therapy: A review. *Frontiers in oncology* 2014;4:327.

201. El-Sherif O, Yu E, Xhaferllari I, et al. Assessment of intrafraction breathing motion on left anterior descending artery dose during left-sided breast radiation therapy. *International Journal of Radiation Oncology* Biology* Physics* 2016;95:1075-1082.

202. George R, Keall P, Kini V, et al. Quantifying the effect of intrafraction motion during breast imrt planning and dose delivery. *Medical physics* 2003;30:552-562.

203. Yue NJ, Li X, Beriwal S, et al. The intrafraction motion induced dosimetric impacts in breast 3d radiation treatment: A 4dct based study. *Medical physics* 2007;34:2789-2800.

204. Morris ED, Ghanem AI, Dong M, et al. Cardiac substructure segmentation with deep learning for improved cardiac sparing. *Medical Physics* 2019.

205. Johnson PB, Padgett KR, Chen KL, et al. Evaluation of the tool “reg refine” for user-guided deformable image registration. *Journal of applied clinical medical physics* 2016;17:158-170.

206. Thirion J-P. Image matching as a diffusion process: An analogy with maxwell's demons 1998.

207. Muyan-Ozcelik P, Owens JD, Xia J, et al. Fast deformable registration on the gpu: A cuda implementation of demons. 2008 International Conference on Computational Sciences and Its Applications. IEEE. 2008. pp. 223-233.

208. Piper J, Nelson A, Harper J. Deformable image registration in mim maestro evaluation and description. *Cleveland, OH: MiM Software Inc* 2013.

209. Fukumitsu N, Nitta K, Terunuma T, et al. Registration error of the liver ct using deformable image registration of mim maestro and velocity ai. *BMC medical imaging* 2017;17:30.

210. Keall PJ, Mageras GS, Balter JM, et al. The management of respiratory motion in radiation oncology report of aapm task group 76 a. *Medical physics* 2006;33:3874-3900.

211. Seppenwoolde Y, Shirato H, Kitamura K, et al. Precise and real-time measurement of 3d tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int J Radiat Oncol Biol Phys* 2002;53:822-34.

212. Wang Y, Bao Y, Zhang L, et al. Assessment of respiration-induced motion and its impact on treatment outcome for lung cancer. *BioMed research international* 2013;2013.

213. Shahid Z, Burns B. Anatomy, abdomen and pelvis, diaphragm Statpearls [internet]: StatPearls Publishing; 2019.

214. Welgemoed C, Rogers J, McNaught P, et al. Deep inspirational breath hold to reduce cardiac dose in left-sided breast radiotherapy. *Journal of Radiotherapy in Practice* 2017;16:251-257.

215. Betgen A, Alderliesten T, Sonke J-J, et al. Assessment of set-up variability during deep inspiration breath hold radiotherapy for breast cancer patients by 3d-surface imaging. *Radiotherapy and Oncology* 2013;106:225-230.

216. Bartlett FR, Colgan RM, Carr K, et al. The uk heartspare study: Randomised evaluation of voluntary deep-inspiratory breath-hold in women undergoing breast radiotherapy. *Radiotherapy and Oncology* 2013;108:242-247.

217. Lutz CM, Poulsen PR, Fledelius W, et al. Setup error and motion during deep inspiration breath-hold breast radiotherapy measured with continuous portal imaging. *Acta Oncologica* 2016;55:193-200.

218. Brock KK, Mutic S, McNutt TR, et al. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the aapm radiation therapy committee task group no. 132. *Medical physics* 2017;44:e43-e76.

219. Zhang TW, Snir J, Boldt RG, et al. Is the importance of heart dose overstated in the treatment of non-small cell lung cancer? A systematic review of the literature. *International Journal of Radiation Oncology* Biology* Physics* 2019;104:582-589.

220. Gore E, Hu C, Ad VB, et al. Impact of incidental cardiac radiation on cardiopulmonary toxicity and survival for locally advanced non-small cell lung cancer: Reanalysis of nrg oncology/rtog 0617 with centrally contoured cardiac structures 2016;96:S129-30.

221. Wennstig A-K. Long-term side effects of radiotherapy in breast cancer: Studies in ischemic heart disease and lung cancer. In: Editor, editor^editors. Book Long-term side effects of radiotherapy in breast cancer: Studies in ischemic heart disease and lung cancer: Umeå Universitet; 2020.

222. Noel CE, Parikh PJ, Spencer CR, et al. Comparison of onboard low-field magnetic resonance imaging versus onboard computed tomography for anatomy visualization in radiotherapy. *Acta Oncol* 2015;54:1474-82.

223. Fischer-Valuck BW, Henke L, Green O, et al. Two-and-a-half-year clinical experience with the world's first magnetic resonance image guided radiation therapy system. *Adv Radiat Oncol* 2017;2:485-493.

224. Henke LE, Olsen JR, Contreras JA, et al. Stereotactic mr-guided online adaptive radiation therapy (smart) for ultracentral thorax malignancies: Results of a phase 1 trial. *Advances in radiation oncology* 2019;4:201-209.

225. Alderliesten T, Betgen A, Elkhuizen PH, et al. Estimation of heart-position variability in 3d-surface-image-guided deep-inspiration breath-hold radiation therapy for left-sided breast cancer. *Radiotherapy and Oncology* 2013;109:442-447.

226. Comsa D, Zhang B, Mosely D, et al. Poster—thur eve—26: Interfraction reproducibility of heart position during breast irradiation using active breathing control. *Medical physics* 2012;39:4629-4629.

227. Fuchs F, Laub G, Othomo K. Truefisp—technical considerations and cardiovascular applications. *European journal of radiology* 2003;46:28-32.

228. Jung BA, Hennig J, Scheffler K. Single-breathhold 3d-truefisp cine cardiac imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* 2002;48:921-925.

229. Landberg T, Chavaudra J, Dobbs J, et al. Report 62. *Journal of the ICRU* 1999:NP-NP.

230. McKenzie A, van Herk M, Mijnheer B. Margins for geometric uncertainty around organs at risk in radiotherapy. *Radiotherapy and Oncology* 2002;62:299-307.

231. Topolnjak R, Borst GR, Nijkamp J, et al. Image-guided radiotherapy for left-sided breast cancer patients: Geometrical uncertainty of the heart. *International Journal of Radiation Oncology* Biology* Physics* 2012;82:e647-e655.

232. Klemm HU, Franzen O, Ventura R, et al. Catheter based simultaneous mapping of cardiac activation and motion: A review. *Indian pacing and electrophysiology journal* 2007;7:148.

233. Aldridge K, Morris E, Ghanem AI, et al. Potential for sensitive cardiac substructure sparing using mr-guided radiation therapy. *International Journal of Radiation Oncology• Biology• Physics* 2019;105:E728.

234. Ferris MJ, Martin KS, Switchenko JM, et al. Sparing cardiac substructures with optimized volumetric modulated arc therapy and intensity modulated proton therapy in thoracic radiation for locally advanced non-small cell lung cancer. *Practical Radiation Oncology* 2019;9:e473-e481.

235. Gore E, Hu C, Ad VB, et al. Impact of incidental cardiac radiation on cardiopulmonary toxicity and survival for locally advanced non-small cell lung cancer: Reanalysis of nrg oncology/rtog 0617 with centrally contoured cardiac structures. *International Journal of Radiation Oncology* Biology* Physics* 2016;96:S129-30.

236. Hoppe B, Bates J, Mendenhall N, et al. The relationship of mean heart dose and cardiac substructure dose over evolving radiation techniques in mediastinal lymphoma. *International Journal of Radiation Oncology* Biology* Physics* 2018;102:S87.

237. Jacob S, Camilleri J, Derreumaux S, et al. Is mean heart dose a relevant surrogate parameter of left ventricle and coronary arteries exposure during breast cancer radiotherapy: A dosimetric evaluation based on individually-determined radiation dose (baccarat study) 2019;14:29.

238. Klüter S. Technical design and concept of a 0.35 t mr-linac. *Clinical Translational Radiation Oncology* 2019;18:98-101.

239. Wen N, Kim J, Doemer A, et al. Evaluation of a magnetic resonance guided linear accelerator for stereotactic radiosurgery treatment. *Radiotherapy and Oncology* 2018;127:460-466.

240. Benedict SH, Yenice KM, Followill D, et al. Stereotactic body radiation therapy: The report of aapm task group 101. *Medical physics* 2010;37:4078-4101.

241. Marks LB, Yorke ED, Jackson A, et al. Use of normal tissue complication probability models in the clinic. *International Journal of Radiation Oncology* Biology* Physics* 2010;76:S10-S19.

242. Badiyan SN, Robinson CG, Bradley JD. Radiation toxicity in lung cancer patients: The heart of the problem? *International Journal of Radiation Oncology* Biology* Physics* 2019;104:590-592.

243. Poitevin-Chacón A, Chávez-Nogueda J, Prudencio RR, et al. Dosimetry of the left anterior descending coronary artery in left breast cancer patients treated with postoperative external radiotherapy. *Reports of Practical Oncology Radiotherapy* 2018;23:91-96.

244. Lester SC, Taparra K, Petersen M, et al. Electrocardiogram-gated computed tomography with coronary angiography for cardiac substructure delineation and

sparing in patients with mediastinal lymphomas treated with radiotherapy. *Practical Radiation Oncology* 2019.

245. Hoppe BS, Bates JE, Mendenhall NP, et al. The meaningless meaning of mean heart dose in mediastinal lymphoma in the modern radiotherapy era. *Practical radiation oncology* 2019.

246. Jacob S, Camilleri J, Derreumaux S, et al. Is mean heart dose a relevant surrogate parameter of left ventricle and coronary arteries exposure during breast cancer radiotherapy: A dosimetric evaluation based on individually-determined radiation dose (baccarat study). *Radiation Oncology* 2019;14:29.

247. Daniel N, Lashmett H, Mazur T, et al. Cardiac motion tracked on cine mri in free-breathing and deep inspiration breath hold: Effect on dose delivered to the heart in breast cancer patients. *International Journal of Radiation Oncology• Biology• Physics* 2017;99:E50.

248. Levis M, De Luca V, Fiandra C, et al. Plan optimization for mediastinal radiotherapy: Estimation of coronary arteries motion with ecg-gated cardiac imaging and creation of compensatory expansion margins. *Radiotherapy Oncology* 2018;127:481-486.

249. van Herk M, Kooy HM. Automatic three-dimensional correlation of ct-ct, ct-mri, and ct-spect using chamfer matching. *Medical physics* 1994;21:1163-1178.

250. Dess RT, Sun Y, Muenz DG, et al. Cardiac dose in locally advanced lung cancer: Results from a statewide consortium. *Practical radiation oncology* 2019;10:27-36.

251. Nejad-Davarani SP, Kim JP, Du D, et al. Large field of view distortion assessment in a low-field mr-linac. *Medical physics* 2019;46:2347-2355.

252. Avendi M, Kheradvar A, Jafarkhani H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac mri. *Medical image analysis* 2016;30:108-119.

253. Radau P, Lu Y, Connelly K, et al. Evaluation framework for algorithms segmenting short axis cardiac mri. *The MIDAS Journal-Cardiac MR Left Ventricle Segmentation Challenge* 2009;49.

254. Brion E, Léger J, Javaid U, et al. Using planning cts to enhance cnn-based bladder segmentation on cone beam ct. Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling. International Society for Optics and Photonics. 2019. pp. 109511M.

255. Griswold MA, Jakob PM, Heidemann RM, et al. Generalized autocalibrating partially parallel acquisitions (grappa). *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* 2002;47:1202-1210.

256. Springenberg JT, Dosovitskiy A, Brox T, et al. Striving for simplicity: The all convolutional net. *arXiv preprint arXiv:14126806* 2014.

257. Wang P, Chung AC. Focal dice loss and image dilation for brain tumor segmentation Deep learning in medical image analysis and multimodal learning for clinical decision support: Springer; 2018. pp. 119-127.

258. Khened M, Kollerathu VA, Krishnamurthi G. Fully convolutional multi-scale residual densenets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. *Medical image analysis* 2019;51:21-45.

259. Lin T-Y, Goyal P, Girshick R, et al. Focal loss for dense object detection. Proceedings of the IEEE international conference on computer vision. 2017. pp. 2980-2988.

260. Chen M, Fang L, Liu H. Fr-net: Focal loss constrained deep residual networks for segmentation of cardiac mri. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). IEEE. 2019. pp. 764-767.

261. Maubon AJ, Ferru J-M, Berger V, et al. Effect of field strength on mr images: Comparison of the same subject at 0.5, 1.0, and 1.5 t. *Radiographics* 1999;19:1057-1067.

262. Cuculich PS, Schill MR, Kashani R, et al. Noninvasive cardiac radiation for ablation of ventricular tachycardia. *New England Journal of Medicine* 2017;377:2325-2336.

263. Gach HM, Green OL, Cuculich PS, et al. Lessons learned from the first human low-field mri guided radiation therapy of the heart in the presence of an implantable cardiac defibrillator. *Practical radiation oncology* 2019;9:274-279.

264. Mayinger M, Kovacs B, Tanadini-Lang S, et al. First magnetic resonance imaging-guided cardiac radioablation of sustained ventricular tachycardia. *Radiotherapy and Oncology* 2020.

265. Maguire PJ, Gardner E, Jack AB, et al. Cardiac radiosurgery (cyberheart™) for treatment of arrhythmia: Physiologic and histopathologic correlation in the porcine model. *Cureus* 2011;3.

266. Sharma A, Wong D, Weidlich G, et al. Noninvasive stereotactic radiosurgery (cyberheart) for creation of ablation lesions in the atrium 2010;7:802-810.

267. Blanck O, Bode F, Gebhard M, et al. Dose-escalation study for cardiac radiosurgery in a porcine model 2014;89:590-598.

268. Gardner EA, Sumanaweera TS, Blanck O, et al. In vivo dose measurement using tlds and mosfet dosimeters for cardiac radiosurgery 2012;13:190-203.

269. Teo B, Dieterich S, Blanck O, et al. Su-ff-t-559: Effect of cardiac motion on the cyberknife synchrony tracking system for radiosurgical cardiac ablation 2009;36:2653-2653.

270. Luo Y, Xu Y, Liao Z, et al. Automatic segmentation of cardiac substructures from noncontrast ct images: Accurate enough for dosimetric analysis? *Acta Oncologica* 2019;58:81-87.

271. Monteiro M, Figueiredo MA, Oliveira AL. Conditional random fields as recurrent neural networks for 3d medical imaging segmentation. *arXiv preprint arXiv:180707464* 2018.

272. Sahiner B, Pezeshk A, Hadjiiski LM, et al. Deep learning in medical imaging and radiation therapy. *Medical physics* 2019;46:e1-e36.

273. Liu M-Y, Breuel T, Kautz J. Unsupervised image-to-image translation networks. In: Editor, editor^editors. Book Unsupervised image-to-image translation networks; 2017. pp. 700-708.

274. Russo P, Carlucci FM, Tommasi T, et al. From source to target and back: Symmetric bi-directional adaptive gan. In: Editor, editor^editors. Book From source to target and back: Symmetric bi-directional adaptive gan; 2018. pp. 8099-8108.

275. Nie D, Trullo R, Lian J, et al. Medical image synthesis with context-aware generative adversarial networks. In: Descoteaux M, Maier-Hein L, Franz A, Jannin P, Collins DL, Duchesne S, editors. Medical image computing and computer-assisted intervention − miccai 2017: 20th international conference, quebec city, qc, canada, september 11-13, 2017, proceedings, part iii. Cham: Springer International Publishing; 2017. pp. 417-425.

276. won Cha J, Farhangi MM, Dunlap N, et al. 4d lung tumor segmentation via shape prior and motion cues. 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE. 2016. pp. 1284-1287.

277. Cvek J, Neuwirth R, Knybel L, et al. Cardiac radiosurgery for malignant ventricular tachycardia. *Cureus* 2014;6:e190.

ABSTRACT

INCORPORATING CARDIAC SUBSTRUCTURES INTO RADIATION THERAPY FOR IMPROVED CARDIAC SPARING

by

ERIC DANIEL MORRIS

August 2020

Advisor: Dr. Carri K. Glide-Hurst

Major: Medical Physics

Degree: Doctor of Philosophy

Growing evidence suggests that radiation therapy (RT) doses to the heart and cardiac substructures (CS) are strongly linked to cardiotoxicities, though only the whole heart is considered clinically. This work aimed to utilize the superior soft tissue contrast of magnetic resonance (MR) to segment CS on non-contrast enhanced computed tomography (CT) and low-field MR, quantify uncertainties in their position, and assess their effect on treatment planning.

Automatic segmentation of 12 CS was completed using a novel hybrid MR/CT atlas method and was improved upon using a 3-dimensional neural network (U-Net) from deep learning (DL). Intra-fraction motion from respiration was then quantified. Inter-fraction setup uncertainties were also assessed to derive planning organ at risk volumes (PRVs) for substructures utilizing a novel MR linear accelerator (MR-linac). Treatment planning comparisons were performed with and without CS and methods to reduce radiation dose were evaluated. Lastly, the 3D U-Net was translated to a low-field MR-linac and a segmentation pipeline was generated.

The hybrid MR/CT atlas was able to generate accurate segmentations for the chambers and great vessels (Dice similarity coefficient (DSC) ≥0.70) but coronary artery segmentations were unsuccessful (DSC<0.27). After implementing DL on CT and low-field MRI, DSC for the chambers and great vessels improved to ≥0.83 along with a coronary artery enhancement (DSC>0.4). Regarding mean distance to agreement, DL segmentations were within 2mm (hybrid MR/CT) and 3mm (low-field MR). Auto-segmentations required ~10 minutes/patient where DL only required 14-20 seconds/patient. Intra-fraction motion caused mean dose variations over 5 Gy for CS in respiration. PRVs from 3-6mm were determined from inter-fraction motion yielding anisotropic substructure-specific margins. The inclusion of CS in treatment planning did not yield statistically significant changes in plan complexity, planning target volume dose, or organ at risk dose.

Segmentation results from DL pose major efficiency and accuracy gains offering high potential for rapid implementation into RT planning for improved cardiac sparing. Introducing CS into treatment planning presented an opportunity for more effective sparing with limited increase in plan complexity. This work may be applied to a prospective clinical trial to determine the effect the dosimetric sparing of CS has on clinical outcomes.

AUTOBIOGRAPHICAL STATEMENT

ERIC DANIEL MORRIS

EDUCATION

2015-2020 Ph.D. Medical Physics, Wayne State University (WSU), Detroit, Michigan
2013-2015 B.S. Physics, Michigan Technological University, Houghton, Michigan
2011-2013 Honors A.S. Delta College, University Center, Michigan

CERTIFICATION

2018 Certificate in Radiomics and Deep Learning
2016 American Board of Radiology Part 1

PROFESSIONAL APPOINTMENTS

2016-2020 Graduate Research Assistant, Henry Ford Health System
Department of Radiation Oncology (Dept. of RO), Detroit, MI
2016-2017 Graduate Teaching Assistant, Wayne State University
2015-2016 Volunteer Research Assistant, Henry Ford Health System & WSU
Department of Radiation Oncology, Detroit, MI
2014-2015 Research Assistant/Laboratory Instructor, Michigan Technological Univ.
Department of Biomedical Engineering, Houghton, Michigan

AWARDS & COMMITTEE APPOINTMENTS

2020 Henry Ford Cancer Institute Outstanding Clinical Manuscript of the Year
2019 23rd Annual WSU School of Medicine Graduate Research Oral Presentation Competition 2nd Place
2019 AAPM John R. Cameron Young Investigators Symposium 3rd Place
2019 2018 Best general medical physics paper in JACMP
2018- AAPM Working Group for Student and Trainee Research: Member
2019- AAPM Student and Trainee Subcommittee: Member
2018 AAPM National Meeting Medical Physics SLAM competition 2nd Place
2018,19 AAPM Great Lakes Chapter (GLC) Medical Physics SLAM competition 1st Place
2018,19 WSU Graduate Student Research Award competition 1st Place
2018 AAPM Expanding Horizons Travel Grant Recipient
2018,19 AAPM GLC Young Investigators Symposium 1st Place (runner-up 2017)
2016,17 WSU Graduate Professional Scholarship Recipient
2016,17,18,19 WSU Graduate Student Professional Travel Award

PUBLICATIONS

1. **E. Morris**, A. Ghanem, M. Dong, M. Pantelic, E. Walker, C. Glide-Hurst. *"Cardiac substructure segmentation with deep learning for improved cardiac sparing."* Medical Physics 47, no. 6 (2020): 576-586.
2. C. Miller, D. Mittelstaedt, N. Black, P. Klahr, S. Nejad-Davarani, H. Schulz, L. Goshen, X. Han, A. I Ghanem, **E. Morris**, C. Glide-Hurst. *"Impact of CT reconstruction algorithm on auto-segmentation performance."* JACMP 20, no. 9 (2019): 95-103.
3. **E. Morris**, A. Ghanem, M. Pantelic, E. Walker, X. Han, and C. Glide-Hurst. *"Cardiac Substructure Segmentation and Dosimetry Using a Novel Hybrid MR/CT Cardiac Atlas."* IJROBP 103, no. 4 (2019): 985-993.
4. **E. Morris**, R. Price, J. Kim, L. Schultz, M. Siddiqui, I. Chetty, and C. Glide-Hurst. *"Using synthetic CT for partial brain radiation therapy: Impact on image guidance."* Practical radiation oncology 8, no. 5 (2018): 342-350.
5. **E. Morris**, J. Kim, P. Klahr, and C. Glide-Hurst. *"Impact of a novel exponential weighted 4DCT reconstruction algorithm."* JACMP 19, no. 6 (2018): 217-225.