UNIVERSITY OF CALIFORNIA

Los Angeles

The Clinical Development of Prostate Magnetic Resonance Imaging-Only Simulation for Radiation Therapy

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy in

Physics and Biology in Medicine

by

Kamal Singhrao

2020

ABSTRACT OF THE DISSERTATION

The Clinical Development of Prostate Magnetic Resonance Imaging-Only Simulation for Radiation Therapy

by

Kamal Singhrao

Doctor of Philosophy in Physics and Biology in Medicine

University of California Los Angeles, 2020

Professor Dan Ruan, Co-Chair

Professor Nzhde Agazaryan, Co-Chair

Magnetic resonance imaging-only (MRI-only) simulation for external beam radiation therapy treatment planning of prostate cancer has seen increased clinical use. The use of a single imaging modality for simulation imaging brings benefits to radiation therapy workflows such as the elimination of systematic positional errors associated with multimodal image registration during treatment planning. However, several challenges remain for the widespread clinical adoption of MRI-only simulation imaging for radiation therapy such as the lack of robust pre-treatment alignment methods and dedicated quality assurance testing equipment.

In the MRI-only simulation imaging workflow, synthetic computed tomography (CT) images are created for a variety of uses including providing tissue electron density information for dose calculations. Synthetic CT image generation algorithms are typically trained using

patient data and are highly sensitive to human tissue contrast and geometry. Most institutions that treat patients with MRI-only simulation images cannot use commercially available phantoms to quality assurance test processes such as synthetic CT image generation. This is because most commercially available phantoms do not mimic human tissue geometry and tissue imaging characteristics for both MRI/CT modalities. The absence of MRI/CT compatible end-to-end quality assurance testing instruments could potentially lead to systematic errors in treatments using MRI-only simulation imaging because of the lack of imaging and dosimetric benchmarking standards.

Studies on the commissioning of MRI-only simulation imaging for radiation therapy of prostate cancers have recommended the use of intraprostatic fiducial markers for pre-treatment patient positioning and alignment. However, fiducial markers appear as dark signal voids in MRI and are challenging to manually localize without the aid of CT imaging. Other intraprostatic objects such as calcifications produce similar signal voids to fiducial markers in MRI images. There is currently no consensus on the optimal fiducial marker or MRI sequence to detect fiducial markers with a high level of sensitivity and specificity in MRI-only simulation images. Additionally, there are no clinically available automatic marker detection workflows available to aid in the clinical transition to MRI-only simulation imaging.

This thesis presents work undertaken to meet the challenges of the clinical development of MRI-only simulation imaging for radiation therapy of prostate cancers. In the presented work, the author describes the development of a novel system of multimodal tissue mimicking materials for MRI and CT imaging. The aforementioned system of materials was adapted into a novel 3D-printed anthropomorphic phantom for quality assurance testing of MRI-only simulation procedures. To address the issues with patient positioning and alignment, a human and phantom study was conducted to quantitatively evaluate the optimal

fiducial marker and MRI sequence for patients receiving MRI-only radiation therapy simulation imaging. Finally, an automatic deep-learning based fiducial marker detection algorithm is presented to aid with the clinical transition of CT-based to MRI-only radiation therapy simulation workflow.

The thesis of Kamal Singhrao is approved.

Amar U. Kishan

Peng Hu

John H. Lewis

Dan Ruan, Committee Co-Chair

Nzhde Agazaryan, Committee Co-Chair

University of California, Los Angeles

2020

This thesis is dedicated to:

the memory of my grandfather Gurbachan Singh Singhrao

whose sacrifices allowed me to be where I am today

and

the memory of Dr. Jean. Pouliot who motivated me to pursue

graduate studies in this field.

Table of Contents

* A version of this chapter was published as a manuscript in Physics in Medicine & Biology in July 2020.
† A version of this chapter was published as a manuscript in Medical Physics in January 2020
‡ A version of this chapter was published as a manuscript in Physics in Medicine & Biology in February 2020
§ A version of this chapter was provisionally accepted for publication as a manuscript in Medical Physics in July 2020

List of Figures

List of Tables

List of Abbreviations

Anterior-Posterior	AP
Apparent Marker Volume	AMV
Calcium (II) Carbonate	$CaCO_3$
Carrageenan-Agarose-Gadolinium-NaCl	CAGN
Clinical Target Volume	CTV
Computed Tomography	CT
Cone Beam CT	CBCT
Convolutional Neural Network	CNN
Cranial-Caudal	CC
Dose Volume Histogram	DVH
Echo Time	TE
External Beam Radiation Therapy	EBRT
False Detection Rate	FDR
False Negative	FN
False Negative Rate	FNR
False Positive	FP
Fast Spin Echo	FSE
Fast-Spoiled-Gradient-Echo	FSPGR
Fiducial Marker	FM
Gadolinium	Gd
Generative Adversarial Network	GAN
Glass Microsphere	GM

Gradient Echo	GRE
Graphics Processing Unit	GPU
Gross Tumor Volume	GTV
Ground Truth	GT
Ground Truth Calcification	GTCalc
Ground Truth Fiducial Marker	GTFM
Hounsfield Unit	HU
Intensity Modulated Radiation Therapy	IMRT
Inversion Time	TI
Least Squares Generative Adversarial Network	LSGAN
Left-Right	LR
Linear Accelerator	linac
Local Contrast-to-Noise-Ratio	lCNR
Magnetic Resonance Imaging	MRI
Mean Absolute Error	MAE
Model Predicted Fiducial Marker	MPFM
Multi-Leaf Collimator	MLC
Multiple Echo Gradient Echo	MEGRE
Organ at Risk	OAR
Planning Target Volume	PTV
Poly-Lactic Acid	PLA
Positron Emission Tomography	PET
Quality Assurance	QA
Region of Interest	ROI

Repetition Time	TR
Sodium Fluoride	NaF
Statistical Decomposition Algorithm	SDA
Stereotactic Body Radiation Therapy	SBRT
Superior-Inferior	SI
Synthetic Computed Tomography	sCT
Target Registration Error	TRE
True Negative	TN
True Positive	TP
Turbo Spin Echo	TSE
Volume Interpolated Breath-hold Examination	VIBE
Volumetric Arc Therapy	VMAT

Acknowledgements

I would like to take this opportunity to thank my thesis advisor Dr. John H. Lewis. John has been instrumental in my development as a scientist and clinical researcher. Through John, I have learned to tackle research problems with tenacity while ensuring that the solution is oriented to have a clear clinical impact. John gave me enough creative and conceptual freedom to pursue projects and sufficient motivation to push through when the projects became challenging. I am extremely lucky to have been able to work with both a brilliant clinical scientist but also a humble and kind person. I entered this process looking for a mentor but have completed it with both a mentor and a friend.

I would also like to thank my committee Dr. Amar U. Kishan, Dr. Dan Ruan, Dr. Nzhde Agazaryan and Dr. Peng Hu. I sincerely appreciate the lessons and invaluable feedback from each of you. Additionally, I would like to acknowledge other faculty members of the Radiation Oncology department including Dr. Daniel A. Low, Dr. Anand Santhanam, Dr. Minsong Cao, Dr. Robert Chin and Dr. Yingli Yang. All of these people were instrumental in my understanding of medical physics and radiation oncology and am extremely grateful for the wisdom and advice that they imparted on to me. Additionally, I would like to thank the Physics and Biology in Medicine (PBM) program director Dr. Michael McNitt-Gray, the program staff Reth Im and Alondra Correa-Bautista, and the Department of Molecular and Medical Pharmacology student affairs officer Emily Fitch for their assistance with navigating the administrative aspects of graduate school.

I would not be the person who I am today without my former UCSF colleagues and friends including Dr. Neil Kirby and Dr. Atchar Sudhyadhom. They have provided unbelievable support and encouragement since my first exposure to the medical physics field. Through Atchar, I learned that alcohol is the key to academic and professional success. And through Neil, I learned that the

sporadic use of unrestrained rage is the key to completing research. They showed me that medical physicists can be cool and served as a bastion to that philosophy.

The PhD. journey would not have been the same without my friends and peers including Geraldine Chee, Michael V. Lauria, Bradley J. Stiehl, John S. Ginn, Adam Gold, Fadil Ali and Tyler E. Cork. I am truly lucky to have had such a supportive group of friends. Experiencing graduate school with these people has truly been a pleasure. Thanks to these people we accomplished the creation and disbandment of a drinking club, performed numerous pranks at the PBM holiday parties and popularized a new form of slang at UCLA.

I could not have embarked on this journey without the help and support of my family. My father Harjit, mother Amarjit, sister Ruby and brother Navid have been an endless source of support. Thank you for all of your sacrifices that have allowed me to be where I am today. We have moved from different continents and faced many challenges over the past 15 years. I hope the completion of this work honors the sacrifices that were made to allow me to do this work.

Finally, I would like to thank my life partner, Danielle H. Morrow. Danielle has been an endless source of love, inspiration and support for me through my graduate studies at UCLA. I am extremely lucky to be spending my life with a beautiful soul and a brilliant scientist. I cannot wait for our post-graduate school journey.

Vita

Education

M.S. in Physics and Biology in Medicine (Therapeutic Medical Physics track) – Oct. 2017
University of California Los Angeles, Los Angeles, California, USA

M.Sc. in Physics – Sept. 2015
University of Alberta, Edmonton, Alberta, Canada

B.S. in Physics (Astronomy and Planetary Sciences minor) – March 2012
University of California Santa Barbara, Santa Barbara, California, USA

Honors

John R. Cameron Young Investigator Competition Finalist - Jul. 2019
American Association of Physicists in Medicine Annual Meeting, San Antonio, USA

Physics and Biology in Medicine Interdepartmental Travel Award - May 2019.
Physics and Biology in Medicine, Los Angeles, California, USA

Norman A. Bailey Best Abstract Award - May 2017
Southern California Chapter of the American Association of Physicists in Medicine Annual Meeting.

Eugene V. Cota-Robles Fellowship - Sept. 2015
UCLA Graduate Division, University of California Los Angeles, Los Angeles, California, USA

Mount Zion Health Fund Travel Grant – Apr. 2013
UCSF Medical Center, University of California San Francisco, San Francisco, California, USA

Physics Departmental Honors - Mar. 2012
University of California Santa Barbara, Santa Barbara, California, USA

Patents

Thermoplastic 3-d phantom. Patent Filed - Oct. 2014.
Inventors: **K. Singhrao**, N. Kirby, and J. Pouliot

Multimodality Phantom Apparatus. Patent Filed - Feb. 2020
Inventors: **K. Singhrao** and John H. Lewis

Certifications

Therapeutic Medical Physics Part 1, The American Board of Radiology.

First Authored Publications

K. Singhrao, N. Kirby, and J. Pouliot, A three-dimensional head-and-neck phan- tom for validation of multimodality deformable image registration for adaptive radiotherapy, Medical Physics, vol. 41, no. 12, p. 121709, Dec. 2014.

K. Singhrao, D. Ruan, J. Fu, Y. Gao, G. Chee, Y. Yang, C. King, P. Hu, A.U. Kishan, J.H. Lewis, Quantification of Fiducial Marker Visibility for MRI-Only Prostate Radiotherapy Simulation., Physics in Medicine and Biology, Dec 2019

K. Singhrao, J. Fu, H.H. Wu, P. Hu, A.U. Kishan, R.K.Chin, J.H. Lewis, A Novel Anthropomorphic Multimodality Phantom for MRI-Based Radiotherapy Quality Assurance Testing., Medical Physics, Jan 2020

K. Singhrao, J Fu, Y Gao, H.H. Wu, Y Yang, P Hu, J.H. Lewis, A Generalized System of Tissue-Mimicking Materials for Computed Tomography (CT) and Magnetic Resonance Imaging (MRI)., Physics in Medicine and Biology, April 2020

K. Singhrao, J Fu, N.R. Parikh, A.G. Mikaeilian, D Ruan, A.U. Kishan, J.H. Lewis, A Generative Adversarial Network-Based (GAN-Based) Architecture for Automatic Fiducial Marker Detection in Prostate MRI-Only Radiotherapy Images, Medical Physics, (provisionally accepted for publication, July 2020)

Selected Conference Proceedings

K. Singhrao, J. Fu, A.U. Kishan, J.H. Lewis. A Generative Adversarial Network (GAN)-based Approach for Automatic Fiducial Marker Detection in MR Radiotherapy Simulation Images, American Society of Radiation Oncology (ASTRO) Annual Meeting, 2019

K. Singhrao, J. Fu, S. Tenn, H. Wu, P. Hu, M. Cao, A.U. Kishan, R. Chin, J.H. Lewis, A Novel Multimodal Anthropomorphic Physical Pelvis Phantom for MRI-Based Radiotherapy, American Association of Physicists in Medicine Annual Meeting, 2019. **Finalist in the John R. Cameron Young Investigators Symposium.**

K. Singhrao, J. Fu, Y. Gao, P. Hu, Y. Yang, J.H. Lewis., A system of materials capable of mimicking soft tissues and bone with both CT and MR imaging. European Society of Radiotherapy and Oncology Annual Meeting (ESTRO 38), 2019

K. Singhrao, Y. Gao, J. Fu, M. Guo, Y. Yang, P. Hu, A.U. Kishan, C. King, J.H. Lewis, An Evaluation of Fiducial Marker Visibility for MR-Guided Radiotherapy in a 3D-Printed Anthropomorphic Multimodality Phantom, American Association of Physicists in Medicine Annual Meeting, 2018

K. Singhrao, Y. Yang, T. Wong, G. Chee, J. Fu, D.A. Low, and J.H. Lewis, A Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) compatible anthropomorphic pelvic phantom for MRI-guided radiotherapy, International Society for Magnetic Resonance in Medicine (ISMRM) Annual Meeting, 2018

K. Singhrao, D. Ruan, Y. Yang, G. Chee and J.H. Lewis, An Evaluation of the Visibility and Detectability of Intra-prostatic Gold Markers in MR Images for Treatment Planning, American Association of Physicists in Medicine Annual Meeting, Jun. 2017

K. Singhrao, M. Bostani, A. Hardy, C. Cagnon, and M. McNitt-Gray, The Effects of Various CT Tube Current Modulation (TCM) Schemes On Dose. Do You Really Know What Your TCM Scheme Is Doing?, American Association of Physicists in Medicine Annual Meeting, 2016

K. Singhrao, N. Kirby, and J. Pouliot, The Development of a Set of Deformable Thermoplastic Materials That Mimic Tissue for Kilovoltage and Megavoltage Computed Tomography, American Association of Physicists in Medicine Annual Meeting, 2013

CHAPTER 1 – Introduction to the Dissertation

1.1. Motivation

Prostate cancer is one of the most commonly diagnosed cancers with 191,930 new diagnoses and 33,330 deaths predicted in the United States in 2020 [1]. Since the mid-1970s, 5-year survival rates for men with low risk prostate cancer have increased from 69% to over 95% today [2–4]. This substantial improvement in survival rate was due to the development of improved surgical, radiation and hormone therapies [5]. While there have been improvements in reducing disease morbidity, off target effects from these therapies can have deleterious effects on quality of life [6,7]. For example, external beam radiation therapy (EBRT) is a common treatment option for men with localized prostate cancer but it can lead to urinary and gastrointestinal complications [6,7]. These complications can occur because the prostate gland is located close to a number of radiosensitive structures including the bladder, urethra and rectum [8,9].

The current standard of care for patients receiving radiation therapy involves creating radiation therapy treatment plans using computed tomography (CT) images for tumor delineation [10]. Interobserver studies have demonstrated that using only CT images for tumor delineation can lead to inadequate tumor volume delineation [11,12]. Inadequate tumor delineation due to multimodal image registration in the treatment planning process can result in the propagation of dosimetric and positioning errors through the radiation therapy workflow [13]. These errors can result in radiosensitive structures positioned in high radiation dose regions and potentially lead to treatment induced side effects [14–16]. These errors can be reduced by using a single imaging modality for treatment planning. Magnetic resonance imaging (MRI) produces excellent soft tissue visualization and studies have shown that prostate organ delineations using MRI images have less

interobserver variability compared to delineations using CT images alone [11].

Utilizing a dedicated MRI simulator for radiation therapy treatment planning (MRI-only treatment planning) can allow for the elimination of treatment errors due to multimodal image registration from the treatment planning stages [13,17]. There are many challenges to the clinical implementation of MRI-only radiation therapy simulation imaging including the lack of specialized equipment to perform quality assurance (QA) testing for MRI-only simulation imaging processes or the lack of consensus-driven approaches for patient positioning and alignment at radiation therapy machines using MRI-only simulation images.

1.2. Background

This section will contain an overview of the radiation therapy workflow for patients diagnosed with localized prostate cancer and an introduction to the rationale of MRI-only radiation therapy planning. Prostate cancer diagnosis includes tests to evaluate disease stage and aggressiveness. The radiation therapy treatment workflow for prostate cancer consists of treatment preparation, simulation imaging and treatment planning, patient-specific QA measurements, and treatment delivery. Pretreatment procedures involve patient consultations with radiation oncologists and can include fiducial marker implantation. Simulation imaging and treatment planning involve acquiring images in treatment position and creating simulated radiation therapy plans. QA measurements can include phantom-based verification measurements to ensure that a delivered dose agrees with the planned dose. Treatment delivery consists of the delivery of radiation for therapeutic purposes.

1.2.A. Prostate Cancer Diagnosis

Prostate cancer is a slow-growing disease and at early stages doesn't typically produce apparent symptoms [18]. The disease can manifest itself in similar form to benign prostatic hyperplasia, where the prostate organ appears enlarged or inflamed [19,20]. Symptoms of prostate cancer can include increased urination frequency and dysuria (painful urination) [21,22]. The exact cause of prostate cancer is currently unknown however risk factors include age, obesity and family history [23]. If left untreated, prostate cancer can spread to nearby lymph nodes and metastasize [24,25]. Metastasis occurs when prostate cancer cells break away from their primary tumor location and spread through the lymphatic system to other areas of the body [26]. Metastasis from prostate cancer can occur in the bones, lymph nodes, lungs, liver and brain [27–29].

Prostate cancer diagnosis involves disease detection and staging [30]. Initial disease detection is done via a digital rectal examination (DRE) and/or a prostate specific antigen (PSA) test [31–34]. PSA is an enzyme created by the prostate gland and is detectable in blood plasma [35]. If abnormalities are detected from DRE examinations or elevated blood PSA levels are detected, further testing may be done for disease diagnosis and staging [36]. Further testing can include a prostate biopsy examination [30,36]. Disease spread is evaluated by diagnostic tests including prostate-specific membrane antigen (PSMA) level measurement to evaluate lymph node spread and imaging tests such as ultrasound, bone scans, CT, MRI and positron emission tomography (PET) imaging to evaluate distal spread [30]. Disease severity is quantified using the TNM (tumor-node-metastasis) scale [23,39]. A metric to determine cell differentiation and by proxy disease aggressiveness is the Gleason score [37]. The Gleason score is based on prostate biopsy results and most scores for cancer patients run between 6 to 10 [38]. Using the TNM score and Gleason scores the cancer can then be staged and the patient can be referred for the appropriate treatment [40].

1.2.B. Radiation Therapy for Prostate Cancer

Radiation therapy is a common treatment for prostate cancer and can be delivered in the form of brachytherapy or EBRT. Brachytherapy treatments involve the placement of a radioactive source at or near the tumor location [41,42]. There are two subtypes of brachytherapy treatment, low-dose-rate and high-dose-rate brachytherapy. Low-dose-rate brachytherapy involves the implantation of permanent radioactive seeds into the tumor site so that radiation is delivered to the tumor for an extended period of time [42]. High-dose-rate brachytherapy involves the implantation of interstitial catheters in the tumor site [41,43]. A high dose radioactive source is moved through the catheter and placed at one or more dwell positions within the catheter until the prescribed radiation dose is delivered. EBRT involves noninvasive delivery of radiation through the body [44]. Radiation therapy plans are designed to deliver the prescribed dose to the tumor while minimizing dose to normal tissues. Most standard EBRT treatments for localized prostate cancers are X-ray photon-based treatments [46].

Selection of treatment modality is dependent on multiple factors including disease severity and patient considerations. EBRT or brachytherapy can be used as monotherapies or in combination for localized low-to-medium risk prostate cancers. For high risk or metastatic prostate cancers, EBRT can be used in combination with other therapies such as hormone therapy or radical prostatectomy (surgical removal of the prostate) [43–45]. This thesis will be focused on EBRT for prostate cancer and will not focus on EBRT treatments that involve radical prostatectomy.

1.2.C. Simulation Imaging and Treatment Planning

If a patient is selected to receive EBRT, a radiation therapy plan needs to be created and tested prior to treatment. At some clinical sites, such as the radiation oncology department at the

University of California Los Angeles (UCLA), patients have intraprostatic fiducial markers inserted directly into the prostate gland for pretreatment positioning [47,48]. Fiducial markers are implanted for pretreatment patient positioning and alignment [47,48]. Simulation imaging involves the acquisition of a 3D medical image, typically via CT imaging, of the patient in treatment position for radiation therapy planning. Patients are placed in treatment position and immobilized using vacuum immobilization cushions [49,50]. Patients normally receive tattoos to allow for reproducible laser positioning [51,52].

Radiation therapy plans are typically created by contouring the gross tumor volume (GTV), clinical target volume (CTV) and organs at risks (OARs) [10]. The current standard of care involves treatment of the entire prostate gland as the CTV [53]. CTV structures are contoured by expert radiation oncologists. OARs are contoured by medical dosimetrists. For localized prostate cancer, typical OAR structures include the bladder, rectum, seminal vesicles, penile bulb, femoral heads and bowels [54,55]. The CTV is expanded by 2-10 mm to create a planning target volume (PTV) [56]. The PTV expansion is an error margin designed to factor for treatment induced errors including (but not limited to) patient setup imperfections and radiation beam inhomogeneities [57].

The treatment planning software that is commonly used for prostate cancer planning is Eclipse (Varian Medical Systems, Palo Alto, CA). Common treatment modalities for low risk prostate cancer include Intensity Modulated Radiation Therapy (IMRT) and Volumetric Arc Therapy (VMAT) [58]. At the radiation oncology department at UCLA, a common treatment method that is employed is Stereotactic Body Radiation Therapy (SBRT). SBRT treatments are commonly prescribed for patients with stage I/II prostate cancer and are typically delivered in 5 fractions (treatment days) totaling 40Gy to the PTV [59,60]. After beam arrangement, selection and multi-leaf collimator (MLC) positions are optimized followed by dose calculation. The Eclipse AAA

algorithm is typically used for dose calculation [61]. Based on the calculated dose deposition, plan quality reports (PQRs) are generated to evaluate if OAR/PTV dose volume histogram (DVH) constraints are met [62]. After an evaluation of the DVH values, plans are then reviewed by the attending radiation oncologist.

Phantom based QA measurements are a physical check to ensure that the planned doses are correctly delivered [63,64]. Using a phantom such as ArcCheck (Sun Nuclear, Melbourne, FL), film and ionization chamber measurements can be taken to verify that the planned dose matches the delivered dose using the gamma index criteria [65,66]. The gamma index is a tool for dose distribution comparison which combines planned vs delivered dose difference and distance to agreement into a single quantity [66,67]. If the plan passes these checks and other institution specific QA checks then it is ready for delivery.

1.2.D. Radiation Therapy Treatments

Prior to the delivery of radiation therapy, pretreatment steps are performed, which involve patient setup, pretreatment imaging and couch shifts for positioning correction. At the radiation oncology department in UCLA, vacuum sealed immobilization cushions are used for patient setup [68,69]. Patients are aligned using in room lasers based on skin tattoos received prior to radiation therapy simulation imaging [51]. Prior to treatment, alignment images are acquired in the form of a cone beam CT (CBCT) or anterior-posterior/left-right (AP/LR) X-ray projection images [70]. If fiducial markers are implanted, they are used as a reference to calculate the shifts between simulation and pretreatment images [47,48]. DRR images (generated at the treatment planning step) are used if AP and LR X-ray images were acquired [71]. Treatment can begin after couch shifts are calculated and checked by an attending radiation oncologist.

Photon EBRT treatments are typically delivered using a clinical linear accelerator (linac). Clinical linacs typically deliver photon energies between 1 to 25 MeV [72]. *Figure 1.2.1* shows an image of a typical clinical linac. The major components of a linac consist of a power supply, modulator, electron gun, radiofrequency amplifier, waveguide, accelerator tube and a treatment head [10]. The power supply and modulator are designed to produce a DC current to power the modulator and electron gun. The electron gun consists of a hot cathode where electrons are liberated via thermionic emission and are focused and accelerated towards an anode. The electron beam is then amplified by the modulator and waveguides. The function of the modulator is to provide a high voltage supply to the electron beam amplifier. Most linacs manufactured by Varian Medical Systems use klystrons as radiofrequency amplifiers. In a klystron an electron beam drifts between resonant microwave cavities, and is amplified in energy by absorbing the resonant microwave energy [10]. The klystron outputs a high energy radiofrequency signal that is directed to the waveguides and used to accelerate the main electron beam. The main electron beam is directed using waveguides towards the treatment head. Electron beams interact with a target in clinical linac head to produce a photon beam via bremsstrahlung. After a photon beam is created, it is collimated by a set of tungsten jaws in the linac head and flattened using either flattening filters or flattening-filter free beams. Prior to reaching the patient skin surface additional beam collimation and shaping can be done by tungsten MLCs [10]. The treatment head is typically contained inside the linac gantry which can rotate around the patient as shown in *Figure 1.2.1*.

EBRT treatment involves directing radiation fields and depositing energy at the tumor site. Most clinical linac treatment deliveries involve photon energy transfer interactions such as the Compton effect [10]. Radiation induced tumor cell death occurs by the induction of chromosomal abnormalities [73]. This can result in preprogrammed cell death or interphase cell death by

mechanisms including apoptosis triggered by DNA damage, cell membrane damage and induction of G_1/S phase cell cycle arrest [74–76].

Figure 1.2.1. Image of the Varian Truebeam® clinical linac in the radiation oncology department at UCLA.

1.2.E. Limitations to the Current Treatment Planning Workflow

There are limitations to the standard EBRT treatment planning workflow including challenges to accurate CTV definition and the introduction of tumor positioning errors due to multimodal image registration. These challenges can cause systematic errors that propagate through the treatment planning workflow, and can potentially lead to treatment-induced side effects and compromised tumor control [13].

Many institutions utilize MRI in addition to CT for tumor definition [77,78]. Accurate tumor

definition is necessary for radiation therapy treatments because misidentification of the target boundary could potentially compromise tumor control and overdose nearby critical structures [79–81]. Each voxel in a CT image is assigned a CT number quantified by Hounsfield Units (HU). CT numbers describe the relative ability for electromagnetic radiation to pass through a material. The CT number is closely related to the relative number of electrons per unit volume (electron density) [10]. The prostate gland is surrounded by intrapelvic muscles that have similar electron densities which makes it challenging to define the prostate organ boundary [82,83]. Several studies have reported a high degree of inter- and intra-observer variability in prostate CTV delineation [11,12]. Additionally, several potentially-important OAR structures such as the penile bulb cannot be identified using only CT images because the organ does not produce sufficient contrast relative to its surroundings [11,84–88]. Studies have shown that using both MRI and CT images for target delineation decreases variation in inter-observer organ definition and significantly reduced prostate CTVs [11,89,90]. One study investigated the long-term side effects of prostate cancer treatments using CTV delineation on MRI (73 patients) compared to CTV delineation on CT-images (72 patients) [90]. This study demonstrated that statistically, prostate organ delineation using MRI led to lower urinary toxicity scores 36 months after treatment [90].

Another source of treatment errors results from MRI/CT image registration, which can potentially lead to propagatable tumor definition errors [13]. Diagnostic MRI images are typically fused to CT simulation images using rigid registration for GTV delineation [91]. Image registration is typically performed by matching bone structures in both image modalities for prostate cancer [92]. Organ positioning errors between CT simulation and diagnostic MRI images can occur due to differences in diagnostic MRI scanning conditions (curved couch, no laser positioning), MRI image acquisition differences (typically courser resolutions, image acquisitions in non-axial

planes) and patient considerations (empty bladder protocol for MRI imaging, organ motion and deformation) [93]. Additional organ boundary definition errors can occur because the prostate organ appears systematically larger on CT than MRI [90]. Uncertainties associated with CTV organ definition as a result of multimodal image registration can result in propagation of systematic positioning errors through the treatment workflow [13,17]. Multimodal image registration errors can be eliminated by using only MRI images for radiation therapy simulation (MRI-only radiation therapy simulation).

1.2.F. MRI-Only Radiation Therapy Simulation Imaging

The clinical use of prostate MRI-only simulation imaging for radiation therapy has several advantages over the current simulation imaging standard. The primary advantage, as discussed in the previous section, is that MRI-only simulation images provide excellent soft tissue visualization, allowing improved CTV definition and eliminating multimodal image registration errors associated with MRI/CT image registration. Additional advantages include the ability of MRI to provide functional/multiparametric imaging and a reduction in delivered radiation dose by the removal of CT simulation imaging [13,17]. MRI also allows for biologically-driven imaging such as spectroscopic imaging (imaging directed by the chemical and metabolic changes in tissues), diffusion weighted imaging (measurement of Brownian motion of water molecules in tissues), perfusion imaging (measurement of fluid passage through the tissues) and elastography imaging (measurement of tissue elastic properties) [94–100]. This potentially allows MRI images to be utilized for pretreatment biological and functional tumor evaluation.

1.2.G. Clinical Implementation Challenges of MRI-Only Simulation Imaging

Most modern treatment planning workflows were designed for CT images to have a variety of functionalities including providing anatomical information for tumor and critical structure delineation and tissue electron density information for dose calculations [87,101]. Patient positioning and alignment methods such as DRR generation and fiducial marker implantation were developed or optimized for the CT-based simulation workflow [87,101]. Additionally, most clinically used treatment QA steps are designed to be used with CT images for radiation therapy planning. MRI-only simulation images do not provide tissue electron densities, require clinical solutions for patient positioning and alignment and require specialized QA steps [102].

Clinically available solutions have been developed to provide electron density information for prostate MRI-only simulation images. Synthetic CT (sCT) images are used to provide MRI images with tissue electron densities [103]. Correct assignment of tissue electron densities is necessary because errors can cause inaccuracies in the treatment planning dose calculation steps [103]. A variety of sCT generation methods have been developed including voxel-based assignment methods and atlas-based methods. Voxel-based assignment methods involve converting voxels containing MRI intensities into CT numbers [104,105]. Atlas-based assignment methods involve deforming or matching the tissue morphology of MRI images to an atlas of aligned MRI and CT images [104,105]. Electron density assignment can be done using atlas-based methods by assigning a single CT number to tissue groups (such as bone, musculature and adipose tissue) or by deforming and smoothing an atlas CT image [13,106–108]. An example of two commercially-available sCT generators for prostate MRI simulation are the Phillips MR-CAT and Spectronics MRIPlanner [107,108,109]. The MR-CAT software uses a 3D Dixon-method fast field echo sequence to generate sCT images using bulk density assignment for air, adipose, water, trabecular bone and cortical bone [110]. MRIPlanner converts T_2-weighted MRI images to sCTs using a statistical decomposition

algorithm (SDA) [111]. The SDA method uses a multimodal patient atlas to assign CT numbers to a deformed T_2-weighted MRI image [111]. Unlike MR-CAT, the MRIplanner method can be used with images acquired from MRI scanners from a variety of manufacturers. The mean absolute error (MAE) between generated sCT images and reference deformed CT images were 47 ± 5 HU and 36 ± 4 HU for the MR-CAT and MRIplanner methods [109,112]. sCT images generated using both methods have been used to create dosimetrically similar treatment plans to those using CT images with gamma pass rates of over 99% with a 2%/2-mm global gamma criteria [109,112]. Using both methods, DRR images can be generated and allow alignment accuracy to pre-treatment X-ray images to within 1.5mm [111,113].

Procedures for fiducial marker-based pretreatment patient positioning using MRI-only simulation images still require development. Fiducial markers provide a high contrast surrogate for prostate organ location in the CT-based treatment planning workflow for pretreatment positioning and alignment [47,48]. Fiducial markers are used in the MRI-only treatment planning workflow because the differences in the organ visibility in MRI and CBCT images and the high degree of interobserver variability of the prostate in CBCT images may cause challenges in pretreatment patient positioning [114,115]. Many commercially available fiducial markers for treatment planning are manufactured from high atomic number, *Z*, metals because they provide a distinct and easily identifiable artifact on CT. However, many commercially available fiducial markers appear as small signal voids in MRI. Intraprostatic calcifications occur in ~40% of prostate cancer patients and appear as small signal voids not unlike implanted fiducial markers [116]. Use of fiducial markers manufactured using materials such as polymers or gold doped with iron may allow for improved detectability of fiducial markers in MRI-only radiation therapy simulation images [117,118]. Additionally, a variety of clinical MRI sequences including gradient echo (GRE)

have been suggested for fiducial marker detection using MRI-only simulation images [119]. A quantitative comparison of the optimal commercially available fiducial marker and the optimal clinical MRI sequence for prostate MRI-only simulation images has yet to be conducted.

Approaches for the automatic detection of fiducial markers in prostate MRI-only radiation therapy images could potentially aid clinical transition to MRI-only workflows. Deep-learning based approaches have shown tremendous promise in small object detection in MRI images [120,121]. 2D and 3D deep learning architectures such as conditional generative adversarial networks (GANs) and DeepMedic have been used to classify small hypointense multiple sclerosis lesions in T_1-weighted MRI images [122–124]. Deep learning model-based classification approaches have substantially improved the detection and classification of objects in medical images such as lung nodules for diagnostic CT imaging [125–127]. Utilizing such tools for the automatic detection of fiducial markers can potentially aid radiation oncology clinics adopting MRI-only simulation imaging for the first time [128].

Another challenge to the clinical implementation of MRI-only radiation therapy simulation is the development of QA testing tools. MRI has a number of modality-specific steps that are not present in the conventional QA workflow including quantification of magnetic field inhomogeneities and validation of sCT images [102]. Commercially available phantoms do exist for magnetic inhomogeneity characterization and several MRI-only studies have used these phantoms to quantify scanner-specific magnetic field inhomogeneities [129–131]. Studies have also been done to quantify patient-induced susceptibility artifacts in the male pelvis, and have shown <1mm distortion [132,133]. Software methods have been developed for sCT QA and conversion accuracy including using statistical approaches to evaluate the significance of errors when using sCT images in the planning workflow compared to CT and evaluating sCT dose accuracy by recalculating sCT-

based plans using kV-CBCT images [134,135]. The American Association of Physicists in Medicine (AAPM) practice guidelines on treatment planning using EBRT (along with other AAPM task groups) and have recommended the use of physical phantoms for treatment plan QA and end-to-end testing of new workflows [136(p158),137,138,139(p101)].

Development of physical phantoms for MRI-only simulation QA is challenging because it requires adaptations of phantoms that were developed and optimized for the CT-based treatment planning workflow or requires the development of novel phantoms for the MRI-only workflow. The AAPM Medical Physics Practice Guidelines on the commissioning and QA of treatment planning dose calculations for megavoltage photon and electron beams recommends the use of water-equivalent phantoms with radiation measurement tools such as ionization chambers and radiochromic film for end-to-end testing [140]. Phantom-based QA testing for MRI-based radiation therapy processes is challenging because it requires phantoms to produce similar MRI contrast and morphology to human pelvic tissues for sCT QA and similar electron densities to human pelvic tissues for dose calculation and DRR validation [141]. Commercially available phantoms for CT-based workflows cannot be utilized to validate or check sCT results because they are not designed to produce tissue-like contrast on MRI [130,141,142]. To create phantoms that can be utilized for end-to-end testing of MRI-only radiation therapy processes, a novel system of materials will need to be developed that quantitatively produces MRI contrast and electron densities for a range of pelvic tissues including trabecular bone, muscle and adipose tissue [142]. This novel system of materials will need to be adapted in a structure that conforms to AAPM medical physics practice guidelines and can be used to test MRI-only processes such as sCT generation.

1.3. Specific Aims

There are major challenges to the clinical development of prostate MRI-only simulation imaging including the lack of dedicated end-to-end testing equipment for MRI-only simulation imaging and consensus-driven approaches for fiducial marker identification and verification in MRI images. The goal of this dissertation is to address these challenges with the following specific aims (SAs):

Specific Aim 1: Develop instrumentation to perform end-to-end testing of treatments using MRI-only simulation images for radiation therapy.

SA 1.1. Develop a novel system of bone and soft tissue surrogate materials for MRI and CT suitable for use in a physical phantom.

SA 1.2. Develop a physical anthropomorphic phantom using bone and soft tissue surrogate materials to generate and test sCT image quality.

SA 1.3. Use the physical anthropomorphic phantom to perform end-to-end testing of prostate MRI-only simulation workflows with X-ray-based photon-based treatments.

Specific Aim 2: Quantify the visibility, distinguishability and alignment accuracy of clinically used and commercially available fiducial markers in a set of clinically available MRI sequences.

SA 2.1. Quantify the visibility of a variety of commercially available fiducial markers in phantom.

SA 2.2. Quantify the visibility, distinguishability and alignment accuracy of clinically used gold fiducial markers in a set of clinically used MRI sequences.

SA 2.3. Compare the quantitative visibility and distinguishability metrics of clinically used gold fiducial markers in a set of clinically used MRI sequences to expert human ratings.

Specific Aim 3: Develop an automatic gold fiducial marker detection workflow and compare its classification accuracy to human raters.

SA 3.1. Develop a deep learning-based workflow to automatically classify gold fiducial markers.

SA 3.2. Quantify the registration error from alignment of MRI-only simulation images to CBCT images using automatically classified fiducial markers.

SA 3.3. Compare the classification and alignment accuracy of the automatically detected markers to those defined by expert human raters.

Broader Implications: The specific aims addressed in this dissertation include the development of clinical solutions to improve QA processes for MRI-only radiation therapy simulation imaging and aid the positioning and alignment of cancer patients receiving EBRT treatments with MRI-only simulation images by optimizing fiducial marker detection techniques.

1.4. Dissertation Chapter Summary

The following chapters are written to address these aims. The work presented in Chapter 2 describes the development of a generalizable system of tissue mimicking materials for MRI and CT imaging. Chapter 3 presents the development of a multimodality anthropomorphic phantom for MRI-based QA measurements using the tissue surrogate materials described in Chapter 2. Chapter 4 presents a quantitative evaluation of the visibility and distinguishability of fiducial markers tested across a variety of clinically used MRI-sequences. Marker visibility is quantified in phantom and in patients. Using the presented quantitative metrics, a recommendation is made on the optimal fiducial marker for MRI-only simulation and MRI-sequence for clinically used fiducial markers. Chapter 5 describes the development of an automatic deep-learning-based fiducial marker detection workflow using the optimal MRI sequence identified from Chapter 4. Chapter 6 contains a summary of the major findings from each chapter and an outline for future studies that can be conducted based on this work.

Bibliography

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70(1):7-30. doi:10.3322/caac.21590

2. Kim MM, Hoffman KE, Levy LB, et al. Improvement in prostate cancer survival over time: a 20-year analysis. *Cancer J*. 2012;18(1):1-8. doi:10.1097/PPO.0b013e3182467419

3. SEER Cancer Statistics Review 1975-2007 - Previous Version - SEER Cancer Statistics. SEER. Accessed June 3, 2020. https://seer.cancer.gov/archive/csr/1975_2007/index.html

4. Galper SL, Chen M-H, Catalona WJ, Roehl KA, Richie JP, D'Amico AV. Evidence to support a continued stage migration and decrease in prostate cancer specific mortality. *J Urol*. 2006;175(3 Pt 1):907-912. doi:10.1016/S0022-5347(05)00419-2

5. Cooperberg MR, Lubeck DP, Meng MV, Mehta SS, Carroll PR. The changing face of low-risk prostate cancer: trends in clinical presentation and primary management. *J Clin Oncol*. 2004;22(11):2141-2149. doi:10.1200/JCO.2004.10.062

6. Crook J, Esche B, Futter N. Effect of pelvic radiotherapy for prostate cancer on bowel, bladder, and sexual function: the patient's perspective. *Urology*. 1996;47(3):387-394. doi:10.1016/S0090-4295(99)80458-0

7. Michaelson MD, Cotter SE, Gargollo PC, Zietman AL, Dahl DM, Smith MR. Management of Complications of Prostate Cancer Treatment. *CA Cancer J Clin*. 2008;58(4):196-213. doi:10.3322/CA.2008.0002

8. Wu SY, Boreta L, Wu A, et al. Improved rectal dosimetry with the use of SpaceOAR during high-dose-rate brachytherapy. *Brachytherapy*. 2018;17(2):259-264. doi:10.1016/j.brachy.2017.10.014

9. Moiseenko V, Liu M, Kristensen S, Gelowitz G, Berthelet E. Effect of bladder filling on doses to prostate and organs at risk: a treatment planning study. *J Appl Clin Med Phys*. 2007;8(1):55-68. doi:10.1120/jacmp.v8i1.2286

10. Khan FM, Gibbons (Jr.) JP. *Khan's The Physics of Radiation Therapy*. Lippincott Williams & Wilkins; 2014.

11. Rasch C, Barillot I, Remeijer P, Touw A, van Herk M, Lebesque JV. Definition of the prostate in CT and MRI: a multi-observer study. *Int J Radiat Oncol Biol Phys*. 1999;43(1):57-66.

12. Rasch C, Steenbakkers R, van Herk M. Target definition in prostate, head, and neck. *Semin Radiat Oncol*. 2005;15(3):136-145. doi:10.1016/j.semradonc.2005.01.005

13. Owrangi AM, Greer PB, Glide-Hurst CK. MRI-only treatment planning: benefits and challenges. *Phys Med Biol*. 2018;63(5):05TR01. doi:10.1088/1361-6560/aaaca4

14. Wu J, Haycocks T, Alasti H, et al. Positioning errors and prostate motion during conformal prostate radiotherapy using on-line isocentre set-up verification and implanted prostate markers. *Radiother Oncol*. 2001;61(2):127-133.

15. Wang G, Wang W-L, Liu Y-Q, Dong H-M, Hu Y-X. Positioning error and expanding margins of planning target volume with kilovoltage cone beam computed tomography for prostate cancer radiotherapy. *Onco Targets Ther*. 2018;11:1981-1988. doi:10.2147/OTT.S152915

16. Patel AK, Houser C, Benoit R, Smith RP, Beriwal S. Acute patient-reported bowel quality of life and rectal bleeding with the combination of prostate external beam radiation, low-dose-rate brachytherapy boost, and SpaceOAR. *Brachytherapy*. Published online April 21, 2020. doi:10.1016/j.brachy.2020.03.006

17. Schmidt MA, Payne GS. Radiotherapy planning using MRI. *Phys Med Biol*. 2015;60(22):R323. doi:10.1088/0031-9155/60/22/R323

18. Merriel SWD, Funston G, Hamilton W. Prostate Cancer in Primary Care. *Adv Ther*. 2018;35(9):1285-1294. doi:10.1007/s12325-018-0766-1

19. Miah S, Catto J. BPH and prostate cancer risk. *Indian J Urol*. 2014;30(2):214-218. doi:10.4103/0970-1591.126909

20. Dai X, Fang X, Ma Y, Xianyu J. Benign Prostatic Hyperplasia and the Risk of Prostate Cancer and Bladder Cancer. *Medicine (Baltimore)*. 2016;95(18). doi:10.1097/MD.0000000000003493

21. Just J, Osgun F, Knight C. Lower urinary tract symptoms and prostate cancer: is PSA testing in men with symptoms wise? *Br J Gen Pract*. 2018;68(676):541-542. doi:10.3399/bjgp18X699689

22. Hamilton W, Sharp DJ, Peters TJ, Round AP. Clinical features of prostate cancer before diagnosis: a population-based, case-control study. *Br J Gen Pract*. 2006;56(531):756-762.

23. Information NC for B, Pike USNL of M 8600 R, MD B, Usa 20894. *Localized Prostate Cancer: Overview*. Institute for Quality and Efficiency in Health Care (IQWiG); 2020. Accessed June 3, 2020. https://www.ncbi.nlm.nih.gov/books/NBK284958/

24. Ziaee S, Chu GC-Y, Huang J-M, Sieh S, Chung LWK. Prostate cancer metastasis: roles of recruitment and reprogramming, cell signal network and three-dimensional growth characteristics. *Transl Androl Urol*. 2015;4(4):438-454. doi:10.3978/j.issn.2223-4683.2015.04.10

25. Sartor O, de Bono JS. Metastatic Prostate Cancer. *N Engl J Med*. 2018;378(7):645-657. doi:10.1056/NEJMra1701695

26. Rycaj K, Tang DG. Molecular determinants of prostate cancer metastasis. *Oncotarget*. 2017;8(50):88211-88231. doi:10.18632/oncotarget.21085

27. Bubendorf L, Schöpfer A, Wagner U, et al. Metastatic patterns of prostate cancer: an autopsy study of 1,589 patients. *Hum Pathol*. 2000;31(5):578-583. doi:10.1053/hp.2000.6698

28. Wong SK, Mohamad N-V, Giaze TR, Chin K-Y, Mohamed N, Ima-Nirwana S. Prostate Cancer and Bone Metastases: The Underlying Mechanisms. *Int J Mol Sci*. 2019;20(10). doi:10.3390/ijms20102587

29. Van Hemelrijck M, Feller A, Garmo H, et al. Incidence of second malignancies for prostate cancer. *PLoS ONE*. 2014;9(7):e102596. doi:10.1371/journal.pone.0102596

30. Litwin MS, Tan H-J. The Diagnosis and Treatment of Prostate Cancer: A Review. *JAMA*. 2017;317(24):2532-2542. doi:10.1001/jama.2017.7248

31. Descotes J-L. Diagnosis of prostate cancer. *Asian J Urol*. 2019;6(2):129-136. doi:10.1016/j.ajur.2018.11.007

32. Adhyam M, Gupta AK. A Review on the Clinical Utility of PSA in Cancer Prostate. *Indian J Surg Oncol*. 2012;3(2):120-129. doi:10.1007/s13193-012-0142-6

33. Walsh AL, Considine SW, Thomas AZ, Lynch TH, Manecksha RP. Digital rectal examination in primary care is important for early detection of prostate cancer: a retrospective cohort analysis study. *Br J Gen Pract*. 2014;64(629):e783-787. doi:10.3399/bjgp14X682861

34. Candas B, Cusan L, Gomez JL, et al. Evaluation of prostatic specific antigen and digital rectal examination as screening tests for prostate cancer. *Prostate*. 2000;45(1):19-35. doi:10.1002/1097-0045(20000915)45:1<19::aid-pros3>3.0.co;2-m

35. Stenman UH, Leinonen J, Zhang WM, Finne P. Prostate-specific antigen. *Semin Cancer Biol*. 1999;9(2):83-93. doi:10.1006/scbi.1998.0086

36. Pospihalj B. Staging of prostate cancer: a review with reference for further refinement. *Anal Quant Cytopathol Histpathol*. 2015;37(1):69-74.

37. Chen N, Zhou Q. The evolving Gleason grading system. *Chin J Cancer Res*. 2016;28(1):58-64. doi:10.3978/j.issn.1000-9604.2016.02.04

38. Munjal A, Leslie SW. Gleason Score. In: *StatPearls*. StatPearls Publishing; 2020. Accessed June 3, 2020. http://www.ncbi.nlm.nih.gov/books/NBK553178/

39. Cheng L, Montironi R, Bostwick DG, Lopez-Beltran A, Berney DM. Staging of prostate cancer. *Histopathology*. 2012;60(1):87-117. doi:10.1111/j.1365-2559.2011.04025.x

40. PDQ Adult Treatment Editorial Board. Prostate Cancer Treatment (PDQ®): Patient Version. In: *PDQ Cancer Information Summaries*. National Cancer Institute (US); 2002. Accessed June 3, 2020. http://www.ncbi.nlm.nih.gov/books/NBK65915/

41. Crook J, Marbán M, Batchelar D. HDR Prostate Brachytherapy. *Semin Radiat Oncol*. 2020;30(1):49-60. doi:10.1016/j.semradonc.2019.08.003

42. Stish BJ, Davis BJ, Mynderse LA, McLaren RH, Deufel CL, Choo R. Low dose rate prostate brachytherapy. *Transl Androl Urol*. 2018;7(3):341-356. doi:10.21037/tau.2017.12.15

43. Stromberg JS, Martinez AA, Horwitz EM, et al. Conformal high dose rate iridium-192 boost brachytherapy in locally advanced prostate cancer: superior prostate-specific antigen response compared with external beam treatment. *Cancer J Sci Am*. 1997;3(6):346-352.

44. Horwitz EM, Hanks GE. External beam radiation therapy for prostate cancer. *CA: A Cancer Journal for Clinicians*. 2000;50(6):349-375. doi:10.3322/canjclin.50.6.349

45. Goy BW, Burchette R, Soper MS, Chang T, Cosmatos HA. Ten-Year Treatment Outcomes of Radical Prostatectomy Vs External Beam Radiation Therapy Vs Brachytherapy for 1503 Patients With Intermediate-risk Prostate Cancer. *Urology*. 2020;136:180-189. doi:10.1016/j.urology.2019.09.040

46. Kamran SC, Light JO, Efstathiou JA. Proton versus photon-based radiation therapy for prostate cancer: emerging evidence and considerations in the era of value-based cancer care. *Prostate Cancer and Prostatic Diseases*. 2019;22(4):509-521. doi:10.1038/s41391-019-0140-7

47. O'Neill AGM, Jain S, Hounsell AR, O'Sullivan JM. Fiducial marker guided prostate radiotherapy: a review. *Br J Radiol*. 89(1068). doi:10.1259/bjr.20160296

48. Alexander SE, Kinsella J, McNair HA, Tree AC. National survey of fiducial marker insertion for prostate image guided radiotherapy. *Radiography (Lond)*. 2018;24(4):275-282. doi:10.1016/j.radi.2018.06.003

49. Jacky J. 3-D radiation therapy treatment planning: overview and assessment. *Am J Clin Oncol*. 1990;13(4):331-343. doi:10.1097/00000421-199008000-00014

50. Nicolaou N. Radiation therapy treatment planning and delivery. *Semin Oncol Nurs*. 1999;15(4):260-269. doi:10.1016/s0749-2081(99)80055-2

51. Rathod S, Munshi A, Agarwal J. Skin markings methods and guidelines: A reality in image guidance radiotherapy era. *South Asian J Cancer*. 2012;1(1):27-29. doi:10.4103/2278-330X.96502

52. Greer PB, Dahl K, Ebert MA, Wratten C, White M, Denham JW. Comparison of prostate set-up accuracy and margins with off-line bony anatomy corrections and online implanted fiducial-based corrections. *J Med Imaging Radiat Oncol*. 2008;52(5):511-516. doi:10.1111/j.1440-1673.2008.02005.x

53. Leer JWH. What the clinician wants to know: radiation oncology perspective. *Cancer Imaging*. 2005;5(Spec No A):S1-S2. doi:10.1102/1470-7330.2005.0027

54. Gay HA, Barthold HJ, O'Meara E, et al. Pelvic normal tissue contouring guidelines for radiation therapy: a Radiation Therapy Oncology Group consensus panel atlas. *Int J Radiat Oncol Biol Phys*. 2012;83(3):e353-362. doi:10.1016/j.ijrobp.2012.01.023

55. Taylor T, Richmond N. A UK wide study of current prostate planning practice. *Br J Radiol*. Published online April 24, 2020:20200142. doi:10.1259/bjr.20200142

56. Røthe Arnesen M, Eilertsen K, Malinen E. Optimal treatment margins for radiotherapy of prostate cancer based on interfraction imaging. *Acta Oncol*. 2008;47(7):1373-1381. doi:10.1080/02841860802244190

57. van Herk M. Errors and margins in radiotherapy. *Semin Radiat Oncol*. 2004;14(1):52-64. doi:10.1053/j.semradonc.2003.10.003

58. Murray LJ, Cosgrove V, Lilley J, et al. Developing a class solution for Prostate Stereotactic Ablative Body Radiotherapy (SABR) using Volumetric Modulated Arc Therapy (VMAT). *Radiother Oncol*. 2014;110(2):298-302. doi:10.1016/j.radonc.2013.10.036

59. Jackson WC, Silva J, Hartman HE, et al. Stereotactic Body Radiation Therapy for Localized Prostate Cancer: A Systematic Review and Meta-Analysis of Over 6,000 Patients Treated On Prospective Studies. *Int J Radiat Oncol Biol Phys*. 2019;104(4):778-789. doi:10.1016/j.ijrobp.2019.03.051

60. Cihan Y. The role and importance of SBRT in prostate cancer. *Int Braz J Urol*. 2018;44(6):1272-1274. doi:10.1590/S1677-5538.IBJU.2018.0484

61. Sharpe MB, Battista JJ. Dose calculations using convolution and superposition principles: the orientation of dose spread kernels in divergent x-ray beams. *Med Phys*. 1993;20(6):1685-1694. doi:10.1118/1.596955

62. Clarke S, Goodworth J, Westhuyzen J, et al. Software-based evaluation of a class solution for prostate IMRT planning. *Rep Pract Oncol Radiother*. 2017;22(6):441-449. doi:10.1016/j.rpor.2017.08.001

63. McKenzie EM, Balter PA, Stingo FC, Jones J, Followill DS, Kry SF. Toward optimizing patient-specific IMRT QA techniques in the accurate detection of dosimetrically acceptable and unacceptable patient plans. *Med Phys*. 2014;41(12):121702. doi:10.1118/1.4899177

64. Defoor DL, Stathakis S, Roring JE, et al. Investigation of error detection capabilities of phantom, EPID and MLC log file based IMRT QA methods. *J Appl Clin Med Phys*. 2017;18(4):172-179. doi:10.1002/acm2.12114

65. Liang B, Liu B, Zhou F, Yin F, Wu Q. Comparisons of volumetric modulated arc therapy (VMAT) quality assurance (QA) systems: sensitivity analysis to machine errors. *Radiation Oncology*. 2016;11(1):146. doi:10.1186/s13014-016-0725-4

66. Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys*. 1998;25(5):656-661. doi:10.1118/1.598248

67. Chen M, Lu W, Chen Q, Ruchala K, Olivera G. Efficient gamma index calculation using fast Euclidean distance transform. *Phys Med Biol*. 2009;54(7):2037–2047. doi:10.1088/0031-9155/54/7/012

68. Teh BS, Woo SY, Mai W-Y, et al. Clinical experience with intensity-modulated radiation therapy (IMRT) for prostate cancer with the use of rectal balloon for prostate immobilization. *Med Dosim*. 2002;27(2):105-113. doi:10.1016/s0958-3947(02)00092-4

69. McGary JE, Grant W. A clinical evaluation of setup errors for a prostate immobilization system. *J Appl Clin Med Phys*. 2000;1(4):138-147. doi:10.1120/jacmp.v1i4.2635

70. Ariyaratne H, Chesham H, Pettingell J, Alonzi R. Image-guided radiotherapy for prostate cancer with cone beam CT: dosimetric effects of imaging frequency and PTV margin. *Radiother Oncol*. 2016;121(1):103-108. doi:10.1016/j.radonc.2016.07.018

71. Galvin JM, Sims C, Dominiak G, Cooper JS. The use of digitally reconstructed radiographs for three-dimensional treatment planning and CT-simulation. *Int J Radiat Oncol Biol Phys*. 1995;31(4):935-942. doi:10.1016/0360-3016(94)00503-6

72. Gerbi BJ, Antolak JA, Deibel FC, et al. Recommendations for clinical electron beam dosimetry: supplement to the recommendations of Task Group 25. *Med Phys*. 2009;36(7):3239-3279. doi:10.1118/1.3125820

73. Little JB. Principal Cellular and Tissue Effects of Radiation. *Holland-Frei Cancer Medicine 6th edition*. Published online 2003. Accessed June 3, 2020. https://www.ncbi.nlm.nih.gov/books/NBK12344/

74. Polyak K, Xia Y, Zweier JL, Kinzler KW, Vogelstein B. A model for p53-induced apoptosis. *Nature*. 1997;389(6648):300-305. doi:10.1038/38525

75. Carmeliet P, Jain RK. Molecular mechanisms and clinical applications of angiogenesis. *Nature*. 2011;473(7347):298-307. doi:10.1038/nature10144

76. Little JB, Nagasawa H. Effect of confluent holding on potentially lethal damage repair, cell cycle progression, and chromosomal aberrations in human normal and ataxia-telangiectasia fibroblasts. *Radiat Res*. 1985;101(1):81-93.

77. Anderson CM, Sun W, Buatti JM, et al. Interobserver and intermodality variability in GTV delineation on simulation CT, FDG-PET, and MR Images of Head and Neck Cancer. *Jacobs J Radiat Oncol*. 2014;1(1):006.

78. Dean CJ, Sykes JR, Cooper RA, et al. An evaluation of four CT–MRI co-registration techniques for radiotherapy treatment planning of prone rectal cancer patients. *Br J Radiol*. 2012;85(1009):61-68. doi:10.1259/bjr/11855927

79. Al-Mamgani A, van Putten WLJ, Heemsbergen WD, et al. Update of Dutch multicenter dose-escalation trial of radiotherapy for localized prostate cancer. *Int J Radiat Oncol Biol Phys*. 2008;72(4):980-988. doi:10.1016/j.ijrobp.2008.02.073

80. Dearnaley DP, Sydes MR, Graham JD, et al. Escalated-dose versus standard-dose conformal radiotherapy in prostate cancer: first results from the MRC RT01 randomised controlled trial. *Lancet Oncol*. 2007;8(6):475-487. doi:10.1016/S1470-2045(07)70143-2

81. Kuban DA, Tucker SL, Dong L, et al. Long-Term Results of the M. D. Anderson Randomized Dose-Escalation Trial for Prostate Cancer. *International Journal of Radiation Oncology, Biology, Physics*. 2008;70(1):67-74. doi:10.1016/j.ijrobp.2007.06.054

82. Betrouni N, Iancu A, Puech P, Mordon S, Makni N. ProstAtlas: A digital morphologic atlas of the prostate. *European Journal of Radiology*. 2012;81(9):1969-1975. doi:10.1016/j.ejrad.2011.05.001

83. Luk L, Taffel MT. Cross-sectional anatomy of the male pelvis. *Abdom Radiol (NY)*. Published online January 9, 2020. doi:10.1007/s00261-019-02369-6

84. Steenbakkers RJHM, Deurloo KEI, Nowak PJCM, Lebesque JV, van Herk M, Rasch CRN. Reduction of dose delivered to the rectum and bulb of the penis using MRI delineation for radiotherapy of the prostate. *International Journal of Radiation Oncology*Biology*Physics*. 2003;57(5):1269-1279. doi:10.1016/S0360-3016(03)01446-9

85. Debois M, Oyen R, Maes F, et al. The contribution of magnetic resonance imaging to the three-dimensional treatment planning of localized prostate cancer. *Int J Radiat Oncol Biol Phys*. 1999;45(4):857-865.

86. 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. *Radiother Oncol*. 2001;60(1):49-59.

87. Hentschel B, Oehler W, Strauß D, Ulrich A, Malich A. Definition of the CTV Prostate in CT and MRI by Using CT–MRI Image Fusion in IMRT Planning for Prostate Cancer. *Strahlenther Onkol*. 2011;187(3):183-190. doi:10.1007/s00066-010-2179-1

88. Nguyen P, Aizer A, Assimos D, et al. ACR Appropriateness Criteria® Definitive External-Beam Irradiation in Stage T1 and T2 Prostate Cancer. *American Journal of Clinical Oncology*. 2014;37(3):278-288. doi:10.1097/COC.0000000000000049

89. Salembier C, Villeirs G, Bari BD, et al. ESTRO ACROP consensus guideline on CT- and MRI-based target volume delineation for primary radiation therapy of localized prostate cancer. *Radiotherapy and Oncology*. 2018;127(1):49-61. doi:10.1016/j.radonc.2018.01.014

90. Sander L, Langkilde NC, Holmberg M, Carl J. MRI target delineation may reduce long-term toxicity after prostate radiotherapy. *Acta Oncol*. 2014;53(6):809-814. doi:10.3109/0284186X.2013.865077

91. Giesel FL, Mehndiratta A, Locklin J, et al. IMAGE FUSION USING CT, MRI AND PET FOR TREATMENT PLANNING, NAVIGATION AND FOLLOW UP IN PERCUTANEOUS RFA. *Exp Oncol*. 2009;31(2):106-114.

92. Sarkar A, Santiago RJ, Smith R, Kassaee A. Comparison of manual vs. automated multimodality (CT-MRI) image registration for brain tumors. *Med Dosim*. 2005;30(1):20-24. doi:10.1016/j.meddos.2004.10.004

93. Brock KK, Mutic S, McNutt TR, Li H, Kessler ML. 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(7):e43-e76. doi:10.1002/mp.12256

94. Harry VN, Semple SI, Parkin DE, Gilbert FJ. Use of new imaging techniques to predict tumour response to therapy. *The Lancet Oncology*. 2010;11(1):92-102. doi:10.1016/S1470-2045(09)70190-1

95. Glunde K, Bhujwalla ZM, Ronen SM. Choline metabolism in malignant transformation. *Nature Reviews Cancer*. 2011;11(12):835-848. doi:10.1038/nrc3162

96. Pinker K, Stadlbauer A, Bogner W, Gruber S, Helbich TH. Molecular imaging of cancer: MR spectroscopy and beyond. *European Journal of Radiology*. 2012;81(3):566-577. doi:10.1016/j.ejrad.2010.04.028

97. Arias-Mendoza F, Payne GS, Zakian K, et al. Noninvasive Phosphorus Magnetic Resonance Spectroscopic Imaging Predicts Outcome to First-line Chemotherapy in Newly Diagnosed Patients with Diffuse Large B-Cell Lymphoma. *Academic Radiology*. 2013;20(9):1122-1129. doi:10.1016/j.acra.2013.04.013

98. Padhani AR, Liu G, Mu-Koh D, et al. Diffusion-Weighted Magnetic Resonance Imaging as a Cancer Biomarker: Consensus and Recommendations. *Neoplasia*. 2009;11(2):102-125. doi:10.1593/neo.81328

99. Tsien C, Cao Y, Chenevert T. Clinical Applications for Diffusion Magnetic Resonance Imaging in Radiotherapy. *Seminars in Radiation Oncology*. 2014;24(3):218-226. doi:10.1016/j.semradonc.2014.02.004

100. Pepin KM, Chen J, Glaser KJ, et al. MR elastography derived shear stiffness—a new imaging biomarker for the assessment of early tumor response to chemotherapy. *Magnetic Resonance in Medicine*. 2014;71(5):1834-1840. doi:10.1002/mrm.24825

101. Doemer A, Chetty IJ, Glide-Hurst C, et al. Evaluating organ delineation, dose calculation and daily localization in an open-MRI simulation workflow for prostate cancer patients. *Radiation Oncology*. 2015;10:37. doi:10.1186/s13014-014-0309-0

102. Bird D, Henry AM, Sebag-Montefiore D, Buckley DL, Al-Qaisieh B, Speight R. A Systematic Review of the Clinical Implementation of Pelvic Magnetic Resonance Imaging–Only Planning for External Beam Radiation Therapy. *International Journal of Radiation Oncology, Biology, Physics*. 2019;105(3):479-492. doi:10.1016/j.ijrobp.2019.06.2530

103. Johnstone E, Wyatt JJ, Henry AM, et al. Systematic Review of Synthetic Computed Tomography Generation Methodologies for Use in Magnetic Resonance Imaging-Only Radiation Therapy. *Int J Radiat Oncol Biol Phys*. 2018;100(1):199-217. doi:10.1016/j.ijrobp.2017.08.043

104. Edmund JM, Nyholm T. A review of substitute CT generation for MRI-only radiation therapy. *Radiat Oncol*. 2017;12. doi:10.1186/s13014-016-0747-y

105. Andreasen D, Van Leemput K, Hansen RH, Andersen JAL, Edmund JM. Patch-based generation of a pseudo CT from conventional MRI sequences for MRI-only radiotherapy of the brain. *Med Phys*. 2015;42(4):1596-1605. doi:10.1118/1.4914158

106. Chin AL, Lin A, Anamalayil S, Teo B-KK. Feasibility and limitations of bulk density assignment in MRI for head and neck IMRT treatment planning. *J Appl Clin Med Phys*. 2014;15(5):4851. doi:10.1120/jacmp.v15i5.4851

107. Prior P, Chen X, Gore E, Johnstone C, Li XA. Technical Note: Is bulk electron density assignment appropriate for MRI-only based treatment planning for lung cancer? *Med Phys*. 2017;44(7):3437-3443. doi:10.1002/mp.12267

108. Kim J, Garbarino K, Schultz L, et al. Dosimetric evaluation of synthetic CT relative to bulk density assignment-based magnetic resonance-only approaches for prostate radiotherapy. *Radiation Oncology*. 2015;10:239. doi:10.1186/s13014-015-0549-7

109. Persson E, Gustafsson C, Nordström F, et al. MR-OPERA: A Multicenter/Multivendor Validation of Magnetic Resonance Imaging-Only Prostate Treatment Planning Using Synthetic Computed Tomography Images. *Int J Radiat Oncol Biol Phys*. 2017;99(3):692-700. doi:10.1016/j.ijrobp.2017.06.006

110. Tyagi N, Fontenla S, Zelefsky M, et al. Clinical workflow for MR-only simulation and planning in prostate. *Radiation Oncology*. 2017;12:119. doi:10.1186/s13014-017-0854-4

111. Siversson C, Nordström F, Nilsson T, et al. Technical Note: MRI only prostate radiotherapy planning using the statistical decomposition algorithm. *Med Phys*. 2015;42(10):6090-6097. doi:10.1118/1.4931417

112. Farjam R, Tyagi N, Deasy JO, Hunt MA. Dosimetric evaluation of an atlas-based synthetic CT generation approach for MR-only radiotherapy of pelvis anatomy. *Journal of Applied Clinical Medical Physics*. 2019;20(1):101-109. doi:10.1002/acm2.12501

113. Kemppainen R, Vaara T, Joensuu T, Kiljunen T. Accuracy and precision of patient positioning for pelvic MR-only radiation therapy using digitally reconstructed radiographs. *Phys Med Biol*. 2018;63(5):055009. doi:10.1088/1361-6560/aaad21

114. Choi HJ, Kim YS, Lee SH, et al. Inter- and intra-observer variability in contouring of the prostate gland on planning computed tomography and cone beam computed tomography. *Acta Oncol*. 2011;50(4):539-546. doi:10.3109/0284186X.2011.562916

115. McNair HA, Harris EJ, Hansen VN, et al. Magnitude of observer error using cone beam CT for prostate interfraction motion estimation: effect of reducing scan length or increasing exposure. *Br J Radiol*. 2015;88(1054). doi:10.1259/bjr.20150208

116. Singhrao K, Ruan D, Fu J, et al. Quantification of fiducial marker visibility for MRI-only prostate radiotherapy simulation. *Phys Med Biol*. 2020;65(3):035015. doi:10.1088/1361-6560/ab65db

117. Tanaka O, Nishigaki Y, Hayashi H, et al. The advantage of iron-containing fiducial markers placed with a thin needle for radiotherapy of liver cancer in terms of visualization on MRI: an initial experience of Gold Anchor. *Radiol Case Rep*. 2017;12(2):416-421. doi:10.1016/j.radcr.2017.03.014

118. Chan MF, Cohen GN, Deasy JO. Qualitative Evaluation of Fiducial Markers for Radiotherapy Imaging. *Technol Cancer Res Treat*. 2015;14(3):298-304. doi:10.1177/1533034614547447

119. Gustafsson Christian, Korhonen Juha, Persson Emilia, Gunnlaugsson Adalsteinn, Nyholm Tufve, Olsson Lars E. Registration free automatic identification of gold fiducial markers in MRI target delineation images for prostate radiotherapy. *Medical Physics*. 2017;44(11):5563-5574. doi:10.1002/mp.12516

120. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*. 2018;321:321-331. doi:10.1016/j.neucom.2018.09.013

121. Xue Y, Xu T, Zhang H, Long LR, Huang X. SegAN: Adversarial Network with Multi-scale L1 Loss for Medical Image Segmentation. *Neuroinform*. 2018;16(3):383-392. doi:10.1007/s12021-018-9377-x

122. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, eds. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Lecture Notes in Computer Science. Springer International Publishing; 2015:234-241. doi:10.1007/978-3-319-24574-4_28

123. Peng C, Zhang X, Yu G, Luo G, Sun J. Large Kernel Matters -- Improve Semantic Segmentation by Global Convolutional Network. In: ; 2017:4353-4361. Accessed March 13, 2020. http://openaccess.thecvf.com/content_cvpr_2017/html/Peng_Large_Kernel_Matters_CVPR_2017_paper.html

124. Kamnitsas K, Ledig C, Newcombe VFJ, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*. 2017;36:61-78. doi:10.1016/j.media.2016.10.004

125. Nasrullah N, Sang J, Alam MS, Mateen M, Cai B, Hu H. Automated Lung Nodule Detection and Classification Using Deep Learning Combined with Multiple Strategies. *Sensors (Basel)*. 2019;19(17). doi:10.3390/s19173722

126. Li D, Mikela Vilmun B, Frederik Carlsen J, et al. The Performance of Deep Learning Algorithms on Automatic Pulmonary Nodule Detection and Classification Tested on

Different Datasets That Are Not Derived from LIDC-IDRI: A Systematic Review. *Diagnostics (Basel)*. 2019;9(4). doi:10.3390/diagnostics9040207

127. Shen W, Zhou M, Yang F, Yang C, Tian J. Multi-scale Convolutional Neural Networks for Lung Nodule Classification. *Inf Process Med Imaging*. 2015;24:588-599. doi:10.1007/978-3-319-19992-4_46

128. Maspero M, Berg CAT van den, Zijlstra F, et al. Evaluation of an automatic MR-based gold fiducial marker localisation method for MR-only prostate radiotherapy. *Phys Med Biol*. 2017;62(20):7981. doi:10.1088/1361-6560/aa875f

129. Price RG, Kadbi M, Kim J, Balter J, Chetty IJ, Glide-Hurst CK. Technical Note: Characterization and correction of gradient nonlinearity induced distortion on a 1.0 T open bore MR-SIM. *Medical Physics*. 2015;42(10):5955-5960. doi:10.1118/1.4930245

130. Sun J, Dowling J, Pichler P, et al. MRI simulation: end-to-end testing for prostate radiation therapy using geometric pelvic MRI phantoms. *Phys Med Biol*. 2015;60(8):3097–3109. doi:10.1088/0031-9155/60/8/3097

131. Walker A, Liney G, Holloway L, Dowling J, Rivest-Henault D, Metcalfe P. Continuous table acquisition MRI for radiotherapy treatment planning: Distortion assessment with a new extended 3D volumetric phantom. *Medical Physics*. 2015;42(4):1982-1991. doi:10.1118/1.4915920

132. Stanescu T, Wachowicz K, Jaffray DA. Characterization of tissue magnetic susceptibility-induced distortions for MRIgRT. *Med Phys*. 2012;39(12):7185-7193. doi:10.1118/1.4764481

133. Glide-Hurst C, Nejad-Davarani S, Weiss S, Zheng W, Chetty IJ, Renisch S. Per-organ assessment of subject-induced susceptibility distortion for MR-only male pelvis treatment planning. *Radiat Oncol*. 2018;13(1):149. doi:10.1186/s13014-018-1090-2

134. Korsholm ME, Waring LW, Edmund JM. A criterion for the reliable use of MRI-only radiotherapy. *Radiat Oncol*. 2014;9:16. doi:10.1186/1748-717X-9-16

135. Palmér E, Persson E, Ambolt P, Gustafsson C, Gunnlaugsson A, Olsson LE. Cone beam CT for QA of synthetic CT in MRI only for prostate patients. *J Appl Clin Med Phys*. 2018;19(6):44-52. doi:10.1002/acm2.12429

136. AAPM TG 158: Measurement and calculation of doses outside the treated volume from external-beam radiation therapy - Kry - 2017 - Medical Physics - Wiley Online Library. Accessed June 3, 2020. https://aapm.onlinelibrary.wiley.com/doi/full/10.1002/mp.12462

137. Quality assurance of U.S.-guided external beam radiotherapy for prostate cancer: Report of AAPM Task Group 154 - Molloy - 2011 - Medical Physics - Wiley Online Library. Accessed June 3, 2020. https://aapm.onlinelibrary.wiley.com/doi/full/10.1118/1.3531674

138. Klein EE, Hanley J, Bayouth J, et al. Task Group 142 report: quality assurance of medical accelerators. *Med Phys*. 2009;36(9):4197-4212. doi:10.1118/1.3190392

139. Stereotactic body radiation therapy: The report of AAPM Task Group 101 - Benedict - 2010 - Medical Physics - Wiley Online Library. Accessed June 3, 2020. https://aapm.onlinelibrary.wiley.com/doi/full/10.1118/1.3438081

140. Smilowitz JB, Das IJ, Feygelman V, et al. AAPM Medical Physics Practice Guideline 5.a.: Commissioning and QA of Treatment Planning Dose Calculations — Megavoltage Photon and Electron Beams. *Journal of Applied Clinical Medical Physics*. 2015;16(5):14-34. doi:10.1120/jacmp.v16i5.5768

141. Singhrao K, Fu J, Wu HH, et al. A novel anthropomorphic multimodality phantom for MRI-based radiotherapy quality assurance testing. *Medical Physics*. n/a(n/a). doi:10.1002/mp.14027

142. Singhrao K, Fu J, Gao Y, et al. A generalized system of tissue-mimicking materials for computed tomography (CT) and magnetic resonance imaging (MRI). *Phys Med Biol*. Published online April 6, 2020. doi:10.1088/1361-6560/ab86d4

CHAPTER 2 - A Generalized System of Tissue-Mimicking Materials for CT and MRI

2.1. Abstract

Purpose: The advent of technologies such as MRI-guided radiation therapy has led to the need for phantom materials that are capable of producing tissue-like contrast on both MRI and CT imaging modalities. The purpose of this work is to develop a system of easily made and formed materials with adjustable T_1 and T_2 relaxation times, and X-ray attenuation properties, for mimicking soft tissue and bone with both MRI and CT imaging modalities.

Methods and Materials: The effects on T_1/T_2 relaxation times and CT numbers were quantified for a range of gadolinium (Gd) contrast (0–25 μmol/g), agarose (0–8% w/w), glass microspheres (GMs) (0–10% w/w) and $CaCO_3$ (0–50% w/w) concentrations in a carrageenan-based gel. 105 gel samples were prepared with the additives, carrageenan and water. Samples were imaged in a 3D-printed holding structure to find the attainable range of T_1/T_2 relaxation time and CT number combinations. T_1 and T_2 relaxation time maps were generated using voxel-wise inversion-recovery and spin-echo techniques, respectively. A multivariate linear regression model was generated to allow the materials system to be generalized to semi-arbitrary T_1/T_2 relaxation times and CT numbers. Nine diverse tissue types were mimicked for fit model validation.

Results: The achievable T_1/T_2 relaxation times and CT numbers for the additive concentrations tested in this study spanned from 82 to 2180 ms, 12 to 475 ms, and -117 to +914 Hounsfield Units (HU) respectively. The MAE between the fit model predicted and measured T_1/T_2 relaxation times and CT numbers for the nine tested tissue types was 113 ± 64 ms, 16 ± 26 ms and 11 ± 14 HU respectively.

Conclusions: We have created a system of materials capable of producing tissue-like contrast for 3.0T MRI and CT imaging modalities.

2.2. Introduction

MRI and CT are two of the most commonly used medical imaging modalities [1–5]. Many groups of patients typically receive both MRI and CT imaging for diagnostic and radiation treatment planning purposes [6–8]. Typically, MRI and CT machines are calibrated and QA-tested separately for their respective applications using specialized phantoms [9–12]. The advent of applications such as MRI-guided radiation therapy has led to the need for materials and phantoms that are capable of producing bone and soft tissue-like contrast on both MRI and CT imaging modalities [13,14]. The development of diverse tissue-mimicking materials for both MRI and CT is challenging because it requires matching both the X-ray attenuation and local magnetic properties of human tissue.

Previous studies have been done on the development of tissue surrogate materials for both MRI and CT imaging. D'Souza et al. developed a tissue-mimicking prostate phantom for multimodal imaging [15]. Prostate and muscle tissue contrast was developed using water, agarose, lipid particles, condensed milk, copper (II) sulfate ($CuSO_4$), EDTA and glass beads, with thermiserol as a preservative. Adipose tissue was mimicked using safflower oil in a polyurethane mesh. T_1 and T_2 relaxometry was performed on a 0.94T 40MHz Minispec relaxometer. These phantom materials were able to mimic soft tissue behavior well for prostate, skeletal muscle, and adipose tissue at 0.94T. This system of materials has limitations including the lack of MRI relaxometry measurements at higher magnetic fields and reproducibility issues associated with the use of animal-derived proteins. Niebuhr et al. created non-fat soft tissue surrogate materials using agarose gels, Gd and sodium fluoride (NaF) and olive oil to mimic adipose tissue [16]. MRI relaxometry was performed on 1.5T clinical MRI scanner and CT images were acquired on a dual source CT scanner. This approach allowed for the development of generalized fits to mimic a range of soft tissues. However, the use of large amounts of NaF salts resulted in severe artifacts in T_2-

weighted MRI images. Polyvinyl chloride (PVC)-based tissue surrogate materials have been developed for MRI and CT imaging [17,18]. Tissue surrogate materials were created by varying the PVC-softener ratios and the mass fractions of mineral oil and GMs in PVC. The CT numbers for this system of PVC-based materials varied from −10 to 110 HU. The measured T_1 and T_2 relaxation times were 206.81 ± 17.50 and 20.22 ± 5.74 ms, respectively. However, the PVC-based system of materials cannot mimic the T_1/T_2 relaxation times and CT numbers of adipose tissue and bone.

Carrageenan-based materials have been developed to mimic a diverse array of soft tissues. Hattori et al. developed the Carrageenan-Agarose-Gadolinium-NaCl (CAGN) phantom for 3.0T MRI [19,20]. The CAGN phantom used carrageenan as a gelatinizer, and Gd and agarose to control T_1 and T_2 relaxometry. The CAGN phantom was able to mimic T_1 and T_2 relaxometry for most soft tissues such as muscle, gray matter and cervix at 3.0T fields. This method uses base material with long T_1 and T_2 relaxation times, which allows the MRI characteristics to be easily tuned using T_1 and T_2 relaxation time modifying additives. CT functionality can potentially be added using electron density modifying additives. Singhrao et al used GMs to decrease, and $CaCO_3$ powder to increase the CT number of a cast polyurethane mold [21]. We hypothesized that combining this method with the approach developed by Hattori et al. could allow the for the production of materials that could mimic a diverse range of tissues for both MRI and CT.

In this work, a system of carrageenan-based tissue-mimicking materials for MRI and CT was developed. We tested the achievable T_1/T_2 relaxation times and CT numbers of this system of materials and created a generalizable fit model to mimic nine diverse tissue types including human pelvic bone marrow, adipose tissue and skeletal muscle.

2.3. Methods and Materials

The development of a system of multimodal carrageenan-based tissue surrogate materials involved testing the ranges of achievable T_1/T_2 relaxation times and CT numbers, and creating a multivariate linear fit model to predict the required additive concentrations to create specific tissue surrogate materials. Testing the ranges of achievable T_1/T_2 relaxation times and CT numbers involved creating carrageenan-based gels with varying concentrations of four additives and performing T_1/T_2 relaxometry and CT number measurements. A multivariate linear fit model was developed to attain the required additive concentrations to formulate a tissue-mimicking material. The model was validated by mimicking the T_1/T_2 relaxation times and CT numbers of nine diverse tissue types.

2.3.A. Sample Preparation

The range of achievable MRI relaxometry values and CT numbers was quantified by creating 50g gel samples contained in syringes. Samples were prepared using carrageenan (Food Grade Kappa Carrageenan: Modernist Pantry, Eliot, ME, USA) as a gelatinizer; gadofosveset trisodium, (Lantheus Medical Imaging, Inc, North Billlerica, MA, USA) as a T_1 modifier; A1700 Agarose LE powder (Benchmark Scientific, Sayreville, NJ, USA) as a T_2 modifier; $CaCO_3$ (LD Carlson Co., Kent, OH, USA) and GMs (Alumilite, Kalamazoo, MI, USA) as CT number modifiers, and deionized water. Carrageenan was selected as a gelatinizer because of its insignificant effects on T_1/T_2 relaxation times [19]. Each sample was prepared to weigh 50g and have a fixed concentration of carrageenan at 3% w/w. 105 samples were made with different combinations of T_1/T_2 relaxation time, and CT number modifying additives. Concentrations of Gd contrast were tested at 0, 0.25, 2.5 12.5 and 25 μmol/g increments. Agarose concentrations were tested at 0, 2, 5 and 8% w/w

increments. $CaCO_3$ concentrations at 0, 5, 10, 20, 30 and 50% w/w increments. GMs concentrations were tested at 0, 5 and 10% w/w increments. The range of Gd contrast and carrageenan concentrations were fixed based on the observations from prior studies [19,21,22]. The working range of additive concentrations were established based on prior literature and exploratory tests [19,21,22]. The exploratory tests involved selecting additive concentrations that did not cause the carrageenan gel to prematurely congeal. For samples containing $CaCO_3$ concentrations at 20, 30 and 50% w/w increments, the agarose concentration was tested at 0% w/w. For samples containing GMs concentrations at 5 and 10% w/w, the agarose concentration was tested at 0, 2 and 5% w/w increments. These constraints were placed because exploratory tests demonstrated that samples containing high agarose concentrations and high $CaCO_3$/GMs concentrations congealed before they could be funneled into the sample syringes.

Sample preparation involved creating a base mix for each batch followed by the addition of Gd contrast and agarose. A 500 ml initial base mix batch containing water, $CaCO_3$ (when required), and GMs (when required) was created. The base mix was then poured into 100 ml beakers where Gd contrast was added if required. The final mixing step involved adding the gelling agents, carrageenan and (if necessary) agarose. The gelling agents were added last to prevent premature solidification. Each sample mixture was maintained at 120°C to prevent solidification during construction. Samples were completed by funneling 50g of the hot base mix into a 60 cc Leur Lock syringe and capped. Completed samples were cooled to 21.5°C and stored at that temperature.

Sample syringes were ordered by concentration and placed in a 3D-printed case as shown in *Figure 2.3.1*. The case structure was designed in Inventor (Autodesk, San Rafael, CA, USA) to

hold 20 samples in a 4 x 5 grid. The case structure was 3D-printed on a MakerGear M3 3D-printer (MakerGear, Beachwood, OH, USA).

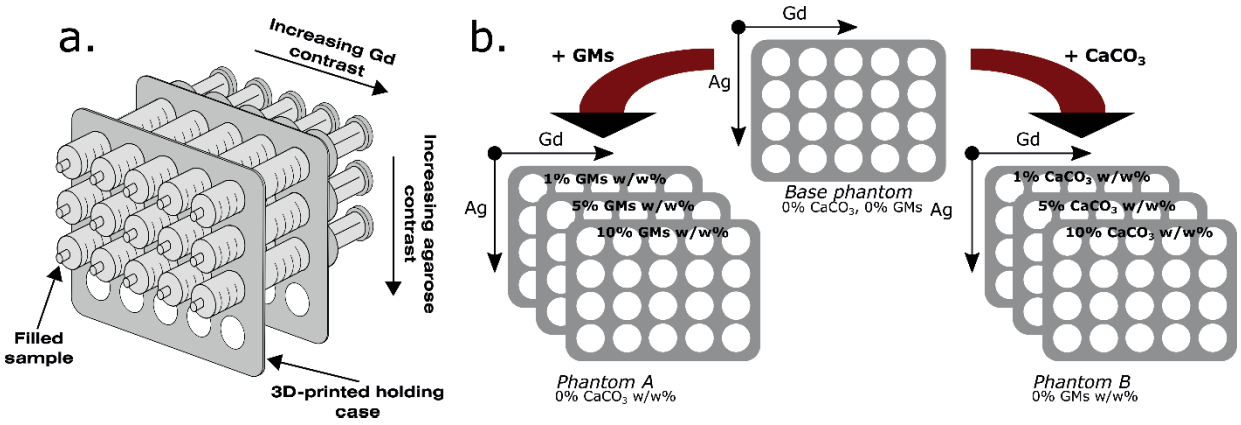


Figure 2.3.1. Graphic of samples contained in holding structure (a). Diagram of sample positioning in holding structure (b). Each phantom was based on a phantom containing 20 samples with a range of Gd contrast and agarose concentrations. Additional sets of phantoms were created with the same concentration variation of Ag and Gd as the base phantom but with increasing concentrations of $CaCO_3$ (Phantom set B) or GMs concentrations (Phantom set A).

2.3.B. Quantification of Imaging Properties

The imaging properties of each sample were quantified using T_1/T_2 relaxometry and CT number measurements (*Figure 2.3.2*). A 3µmol/g $NiCl_2$ hexahydrate solution (Sigma Chemical Corp., St. Louis, MO) was used as a control. CT images were acquired with scanner settings at 120 kVp and 400 mAs. CT number was used as a surrogate for X-ray attenuation. T_1/T_2 relaxometry measurements were performed on a 3.0T MRI scanner (MAGNETOM Skyra, Siemens, Erlangen, Germany). CT images were acquired on a 20-slice single source CT (Sensation Open, Siemens, Erlangen, Germany). All measurements were performed between 20°C to 22°C.

T_1/T_2 relaxation times and CT numbers for each sample were calculated within a 20 mm diameter region of interest (ROI) inside each sample. T_1 measurements were acquired using an inversion-recovery turbo spin-echo (TSE) sequence with 180-degree refocusing pulses. The echo

time (TE) and repetition time (TR) were fixed at 12 ms and 15,000 ms respectively, and the inversion time (TI) was varied from 25 ms to 2000 ms in 50 ms steps. Images were acquired with an echo-train-length of 22, no signal averaging, and without parallel imaging. A 128 x 128 matrix, single 8-mm slice with a readout bandwidth of 130 Hz/px was acquired. A T_1 map was generated by performing a voxel-wise fit to the observed signal intensity according to the equation $M_{z,TI_i} = M_0\left(1 - 2e^{-TI/T_1}\right)$ in MATLAB (MathWorks, Natick, MA) where M_0 is the initial magnetization vector magnitude. T_2 measurements were acquired using spin-echo images using a TSE sequence with a constant TR at 4000 ms and TE values at 25, 50, 62, 75, 87, 107, 167, and 262 ms. Images were acquired with an echo-train-length of 25, 2 signal averages and without parallel imaging. T_2 maps were generated by performing a voxel-wise fit according to the equation $M_{TE_i} = M_0\left(e^{-\frac{TE_i}{T_2}}\right) + c$ where c is a noise offset variable.

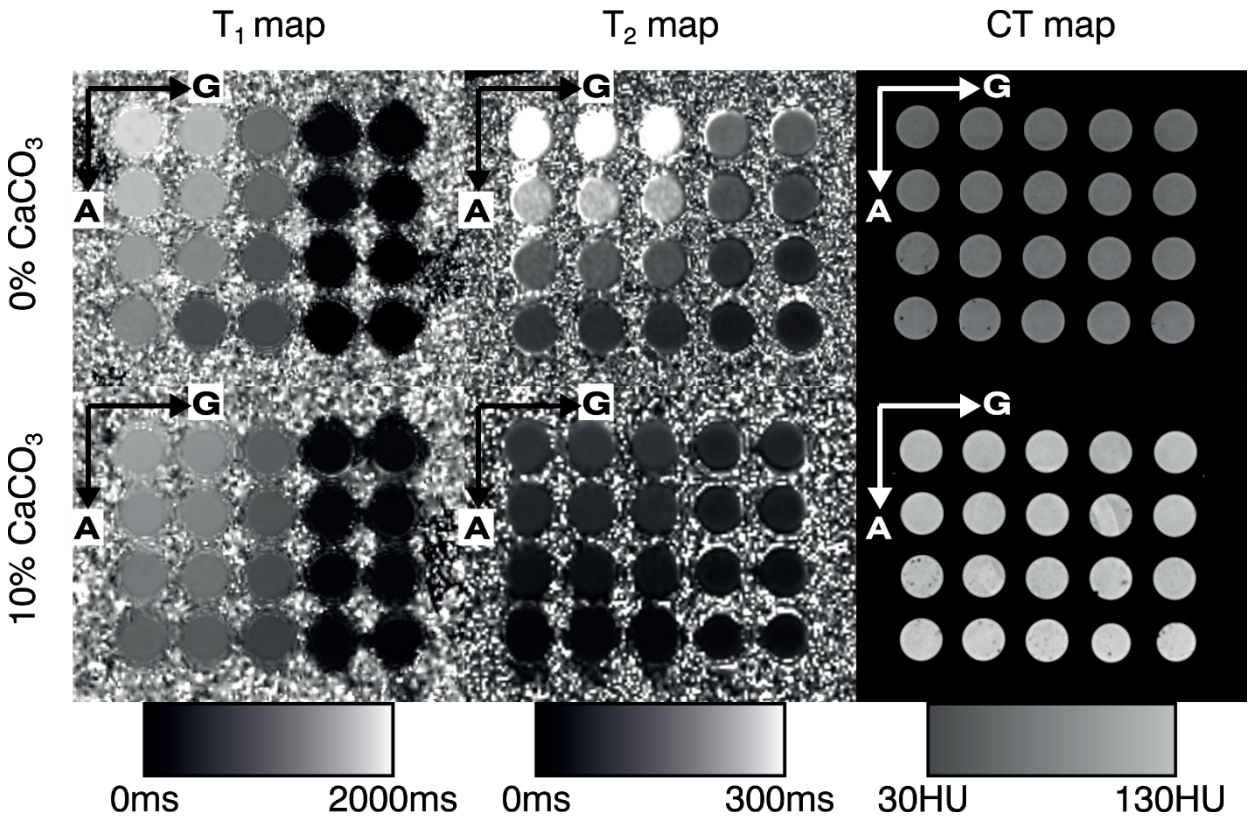


Figure 2.3.2. T_1 and T_2 maps and CT images of the phantom samples containing 0% and 10% $CaCO_3$ by weight. For reference, the concentration axes of Gd contrast (0, 0.25, 2.5 12.5 and 25 μmol/g) and agarose (0, 2, 5 and 8 w/w%) are indicted where the Gd concentration increases from left to right and the agarose concentration increases from top to bottom.

2.3.C. Multivariate Fit Model Development

Multivariate linear models between $1/T_1$, $1/T_2$ and CT number with respect to the four additives were developed using the Python-based Bayesian Ridge regression model package in the sklearn package [23]. A single predictive model for T_1/T_2 relaxation times and CT numbers (CTn) was developed by combining each multivariate linear regression:

$$\begin{pmatrix} 1/T_1 \\ 1/T_2 \\ CTn \end{pmatrix} = \begin{pmatrix} \alpha_0 & \alpha_{Gd} & \alpha_{Ag} & \alpha_{CaCO_3} & \alpha_{GM} \\ \beta_0 & \beta_{Gd} & \beta_{Ag} & \beta_{CaCO_3} & \beta_{GM} \\ \gamma_0 & \gamma_{Gd} & \gamma_{Ag} & \gamma_{CaCO_3} & \gamma_{GM} \end{pmatrix} \begin{pmatrix} 1 \\ c_{Gd} \\ c_{Ag} \\ c_{CaCO_3} \\ c_{GM} \end{pmatrix} \quad (1)$$

where $(\alpha, \beta, \gamma)_0$ are fit intercepts, $(\alpha, \beta, \gamma)_{Gd,Ag,CaCO_3,GM}$ are the fit parameters for Gd contrast, agarose, $CaCO_3$ or GMs respectively, and $c_{Gd,Ag,CaCO_3,GM}$ are the input concentrations of each additive. The fit parameters in equation (1) were established using the T_1/T_2 relaxometry and CT number measurements for all samples. The linear programming package CVX was used to select additive concentrations to produce a desired set of T_1/T_2 relaxation times and CT numbers, [24,25].

The fit model accuracy was tested by creating nine target tissue-surrogate materials: muscle, white matter, gray matter, liver, prostate, glandular breast, adipose tissue, pelvic bone marrow, adipose tissue and cortical bone. Target T_1/T_2 relaxation times and CT numbers were selected based on *in vivo* literature reports [26–48]. Samples representing each tissue type were created using the fit-specified concentrations and imaged under the same conditions described above. Measured T_1/T_2 relaxation times and CT numbers were compared to their corresponding target values for each tissue type. Each tissue-mimicking sample was produced three times.

2.4. Results

2.4.A. Quantification of Imaging Properties

Figure 2.4.1. shows all measured T_1/T_2 relaxation times and CT numbers for each sample. Plots of T_1 relaxation times vs. T_2 relaxation times, T_1 relaxation times vs. CT numbers, and CT numbers vs. T_2 relaxation times are included. Shaded regions in each plot represents the expected T_1/T_2 relaxation times, and CT number values for a variety of human tissues.

The T_1 vs T_2 relaxation time results (*Figure 2.4.1a*) show that the concentrations tested for this system of materials can span a range of T_1 relaxation times from 82 ms to 2180 ms and T_2 relaxation times from 12 ms to 475 ms. The T_1 relaxation time vs CT number and T_2 relaxation time vs CT number (*Figure 2.4.1b, c*) results also show that the range of achievable CT numbers

is -117 HU to 914 HU by the range of concentrations that we tested. Plots showing all T_1/T_2 relaxation time, and CT number measurements as a function of additive concentrations are provided in the *Appendix I-III*.

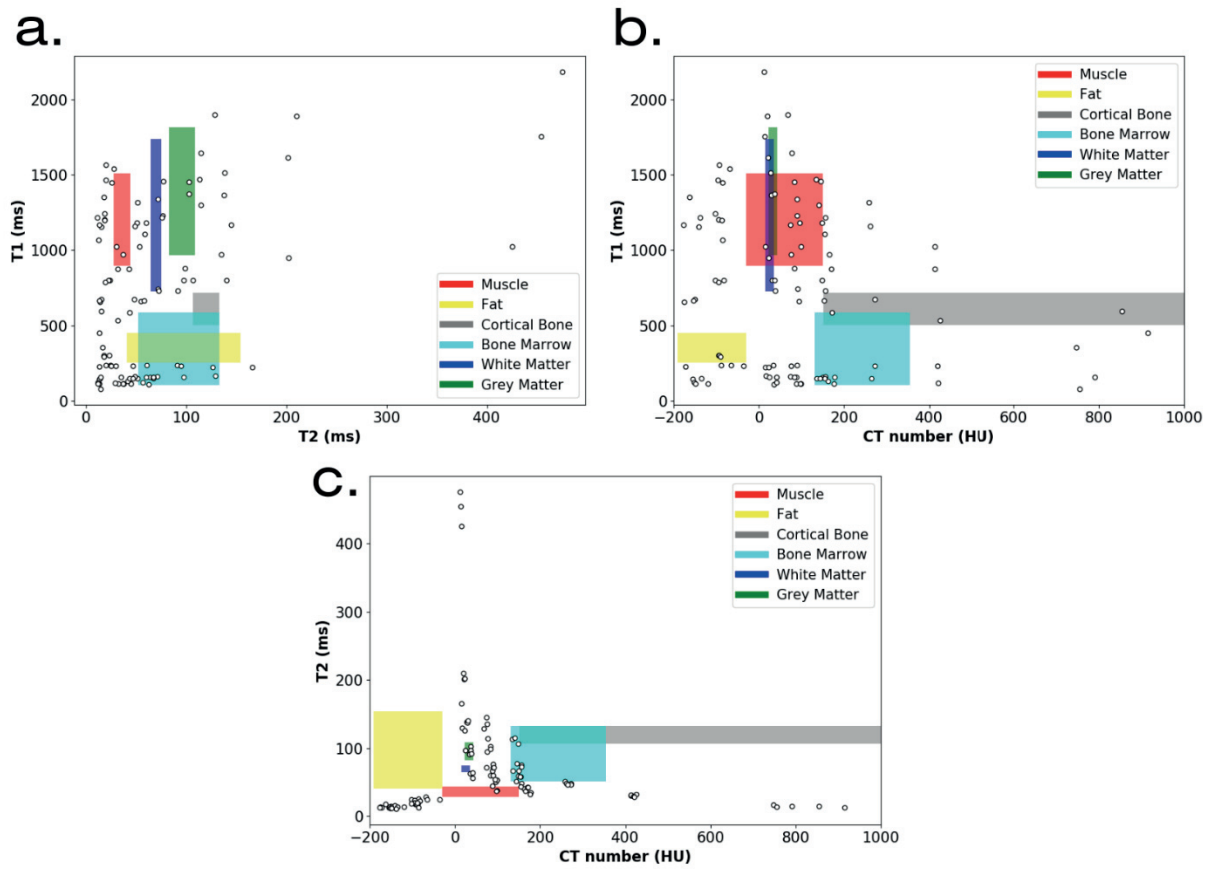


Figure 2.4.1 Scatter plots of T_1 vs T_2 relaxation times (a), T_1 relaxation times vs CT numbers (b) and T_2 relaxation times vs CT numbers (c) for carrageenan-based samples containing different combinations of Gd contrast, agarose, GMs, and $CaCO_3$. Estimated expected ranges for a variety of tissue types are shown in the shaded regions. The shaded T_1 and T_2 relaxation time regions were selected based on reported measurements from a review of in vivo 3.0T MRI T_1 and T_2 maps by Bojorquez et al[36]. The shaded CT number regions were derived from a review of multiple publications [28,31,35,39,40,43,44].

2.4.B. Multivariate Fit Model Parameter Estimation

The multivariate fit coefficients $\alpha_{0,Gd,Ag,CaCO_3,GM}$, calculated using sample measurements were:

$$\begin{matrix} \alpha_0 = 6.1\text{E}-05 \\ \alpha_{Gd} = 1.3\text{E}-02 \\ \alpha_{Ag} = 3.2\text{E}-04 \\ \alpha_{CaCO_3} = 1.4\text{E}-04 \\ \alpha_{GM} = 1.7\text{E}-04 \end{matrix} .$$

Sample measurements and corresponding fits of $1/T_1$ vs Gd contrast concentrations are shown in Appendix *I*.

The multivariate fit coefficients $\beta_{0,Gd,Ag,CaCO_3,GM}$, calculated using sample measurements were:

$$\begin{matrix} \beta_0 = 3.4\text{E}-03 \\ \beta_{Gd} = 1.1\text{E}-02 \\ \beta_{Ag} = 2.9\text{E}-03 \\ \beta_{CaCO_3} = 1.6\text{E}-03 \\ \beta_{GM} = 1.4\text{E}-02 \end{matrix} .$$

Sample measurements and corresponding fits of $1/T_2$ vs agarose concentrations are shown in Appendix *II*.

The multivariate fit coefficients $\gamma_{0,Gd,Ag,CaCO_3,GM}$, calculated using sample measurements were:

$$\begin{matrix} \gamma_0 = 6.4\text{E}+00 \\ \gamma_{Gd} = 1.0\text{E}+01 \\ \gamma_{Ag} = 6.7\text{E}+00 \\ \gamma_{CaCO_3} = 2.7\text{E}+01 \\ \gamma_{GM} = 3.3\text{E}+01 \end{matrix} .$$

Sample measurements and corresponding fits of CT number vs $CaCO_3$ and GMs concentrations are shown in Appendix *III*.

2.4.C. Multivariate Fit Model Validation

The fit model predicted additive concentrations are shown in *Table 2.4.I*. *Table 2.4.II* shows the measured T_1/T_2 relaxation times and CT numbers for samples generated with fit-predicted additive

concentrations, along with the target values for each of the nine tissue types tested. The standard deviations for the tissue-mimicking materials are based on the standard deviation of the mean ROI measurements from three repeated samples.

Tissue type	Gd contrast (µmol/g)	Agarose (% w/w)	$CaCO_3$ (% w/w)	GMs (% w/w)	Carrageenan (% w/w)	Water (% w/w)
Muscle	0.6	0.7	4.7	2.1	3.0	89.5
White Matter	1.8	0.9	3.3	1.2	3.0	91.6
Gray Matter	1.4	0.7	3.3	0.8	3.0	92.2
Liver	2.1	1.2	4.5	2.9	3.0	88.4
Prostate	1.0	0.6	3.5	0.9	3.0	92
Glandular breast	0.8	0.6	2.6	1.9	3.0	91.9
Adipose tissue	11.0	0.0	0.0	5.0	3.0	92
Pelvic Bone Marrow	7.5	1.3	6.1	1.1	3.0	88.6
Cortical Bone	2.0	1.8	22	0.0	3.0	73.2

Table 2.4.I. Fit model predicted concentrations of Gd contrast, agarose, $CaCO_3$, and GMs required to create 100g of each tissue surrogate.

Tissue type	In vivo	Tissue Surrogate	In vivo	Tissue Surrogat	In vivo	ue rogate
White Matter	1110 ± 40 [29]	1023 ± 9	65 ± 6 [30]	61 ± 7	29 (25,34) [31]	: 3
Liver	809 ± 71 [34]	917 ± 68	34 ± 4 [34]	37 ± 6	24.9 (16.7,37.2) [35]	: 1
Glandular breast	1680 ± 180 [37]	1440 ± 153	54 ± 9 [38]	60 ± 6	8 ± 22 [39]	: 2
Pelvic Bone Marrow	381 ± 8 [41]	445 ± 6	52 ± 1 [41]	58 ± 4	77 ± 75 [43]	: 8

Table 2.4.II. Literature reported in vivo tissue measurements and tissue surrogate T_1/T_2 relaxation times and CT numbers for the nine tested tissue types. Some in vivo CT numbers only reported measured ranges and not standard deviations. The reported measured ranges are indicated in parenthesis.

2.5. Discussion

The goal of this work was to create a generalizable model using carrageenan-based tissue-surrogate materials for MRI and CT imaging. The ranges of attainable T_1/T_2 relaxation times and CT numbers were quantified and demonstrated to span a range of T_1 relaxation times between 82 to 2180 ms, T_2 relaxation times between 12 to 475 ms and CT numbers between -117 to 914 HU. A generalizable fit model was developed and the T_1/T_2 relaxation times and CT numbers of nine diverse tissue types were mimicked. For the nine tested tissue types, the multivariate fit model yielded a MAE between the fit model predicted and measured T_1/T_2 relaxation times, and CT numbers were 113 ± 64 ms,16 ± 26 ms and 10 ± 14 HU respectively.

The temporal stability and physical characteristics of this system of materials were not directly investigated in this study. The materials were tested for a period of 12 weeks without exhibiting visible physical degradation. Previous studies have investigated the stability of carrageenan-based materials over time. In et al. quantified the mechanical stability and T_1/T_2 relaxation time stability of carrageenan-based materials with agar and Gd contrast additives [49]. The density and compressive strength changes over 6 weeks were relatively stable for samples with 3% carrageenan as used in our study. The 3.0T T_1 relaxation times of 8 tested samples generally increased by 1.25 times compared to baseline after 5 weeks. The 3.0T T_2 relaxation times of eight tested samples generally decreased 0.75 to 0.95 times compared to baseline after 5 weeks [49]. Hattori et al. reported that the T_1 values of carrageenan gels containing Gd contrast and agarose additives did not change over the course of a 4-month period [19]. They reported that T_2 values of carrageenan gels showed increases in T_2 values of as much as 15%. Previous studies have utilized preservatives such as sodium azide or thimerosal to preserve carrageenan-based or agarose gel-

based materials [15,19]. In future studies, these preservatives could be incorporated into our material formulation by adding them during the mixing steps.

The development of the fit model was based on the assumption that the additives can quasi-linearly control the imaging properties within the range of concentrations studied. Gd contrast and agarose additives were selected because they have been observed to control the T_1 and T_2 relaxation times of carrageenan-based gels [16,22]. Gd contrast agents are paramagnetic and improve the efficiency of T_1 and T_2 relaxation [50]. Typically, the T_2 shortening effect of Gd contrast is not as pronounced as the T_1 shortening effects. Agarose has been shown to shorten T_1 relaxation times of carrageenan gels however its shortening effect are more pronounced on the T_2 relaxation times [19]. Niebuhr et al. demonstrated that the concentration of agarose in gel form has a quasi-linear relationship with electron density [16]. Hollow GMs have been used in previous studies to decrease the apparent CT number of agarose and polyurethane materials [15,21]. The apparent decrease in CT number due to the presence of hollow GMs may be due to voxel averaging between the X-ray attenuation behavior of the hollow GMs and the surrounding carrageenan-based gel. $CaCO_3$ has been used in several previous studies to increase the apparent electron density of phantom molding materials [21,51,52]. All of these previous studies showed no explicit relationship between the concentration of $CaCO_3$ and its effect on electron density. In this study, we established a linear relationship between the $CaCO_3$ concentration and the CT number (which served as a proxy for electron density). Since $CaCO_3$ is insoluble in water, we conjecture that the increase in CT number is due to voxel averaging of $CaCO_3$ particulates and the surrounding carrageenan gel. However, further microscopy studies may elucidate the microscopic effects of $CaCO_3$ on carrageenan gels.

The assumption of linearity in the fit model is a potential source of error because of the observed deviations between the fit-specified concentrations and the measured T_1/T_2 relaxation

times and CT numbers. The fit model underestimated the T_1 relaxation times for tissue-mimicking materials greater than 1200 ms by up to -14%. This could be attributed to higher T_1 relaxation target times being subject to a higher fit error because the T_1 relaxation time error is proportional to T_1^2. In this study we observed that the addition of GMs severely degraded the T_2 relaxation times of carrageenan-based gels. This made development of adipose tissue mimicking materials challenging because it required relatively high concentrations of GMs to match the low CT numbers of in vivo tissues. Developing a model-based approach for these types of tissue required a tradeoff between selecting an optimal CT number and T_2 relaxation time. This could explain why the fit model could not perfectly mimic the CT numbers of glandular breast and adipose tissue because those tissue types required the addition of a high concentration of GMs. To create materials with CT numbers above 200 HU, such as bone surrogate material, more $CaCO_3$ is required. However, the addition of large amounts of $CaCO_3$ also results in suppression of T_2 relaxation times of bone tissue-surrogates, possibly because of the decrease in water volume in the bone surrogate gels.

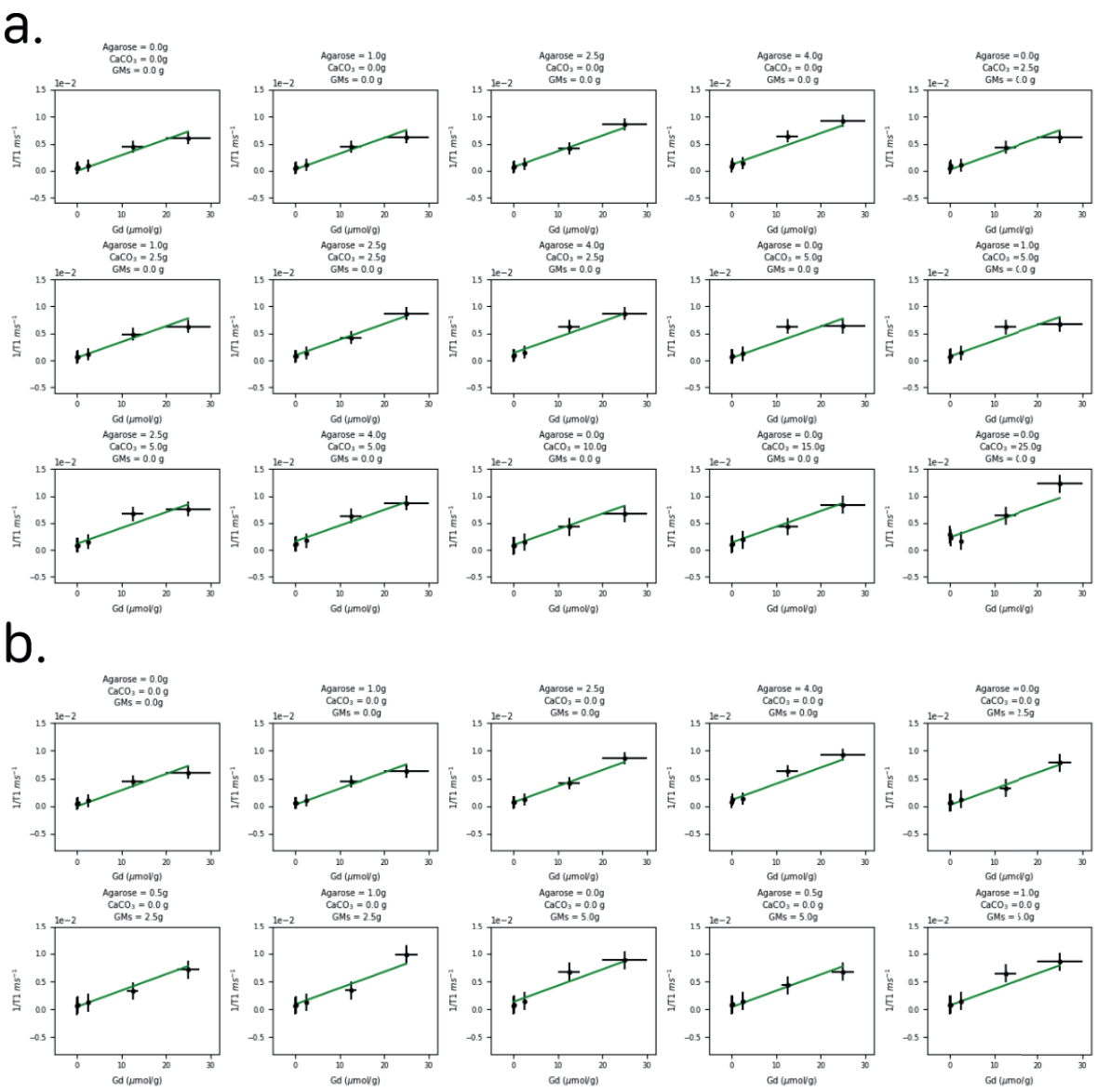
This system of materials could potentially be adapted in physical phantoms for MRI-based radiation therapy treatment QA testing. We plan to adapt this system of materials to be used in a 3D printed structure for the end-to-end testing for MRI-guided QA. Other potential applications include using this system of materials to mimic different types of adipose tissue such as brown adipose tissue and tumors. Further experiments will involve testing a larger range of additive concentrations and testing this system of materials. This experiment was conducted at 3.0T because that is currently the practical upper limit for most MRI simulators for radiation therapy however, future experiments will involve testing these materials at lower magnetic field strengths.

Testing at different field strengths may allow this system of materials to be used in QA instrumentation for a wider range of MRI-only and onboard-MRI-guided treatments.

2.6. Conclusions

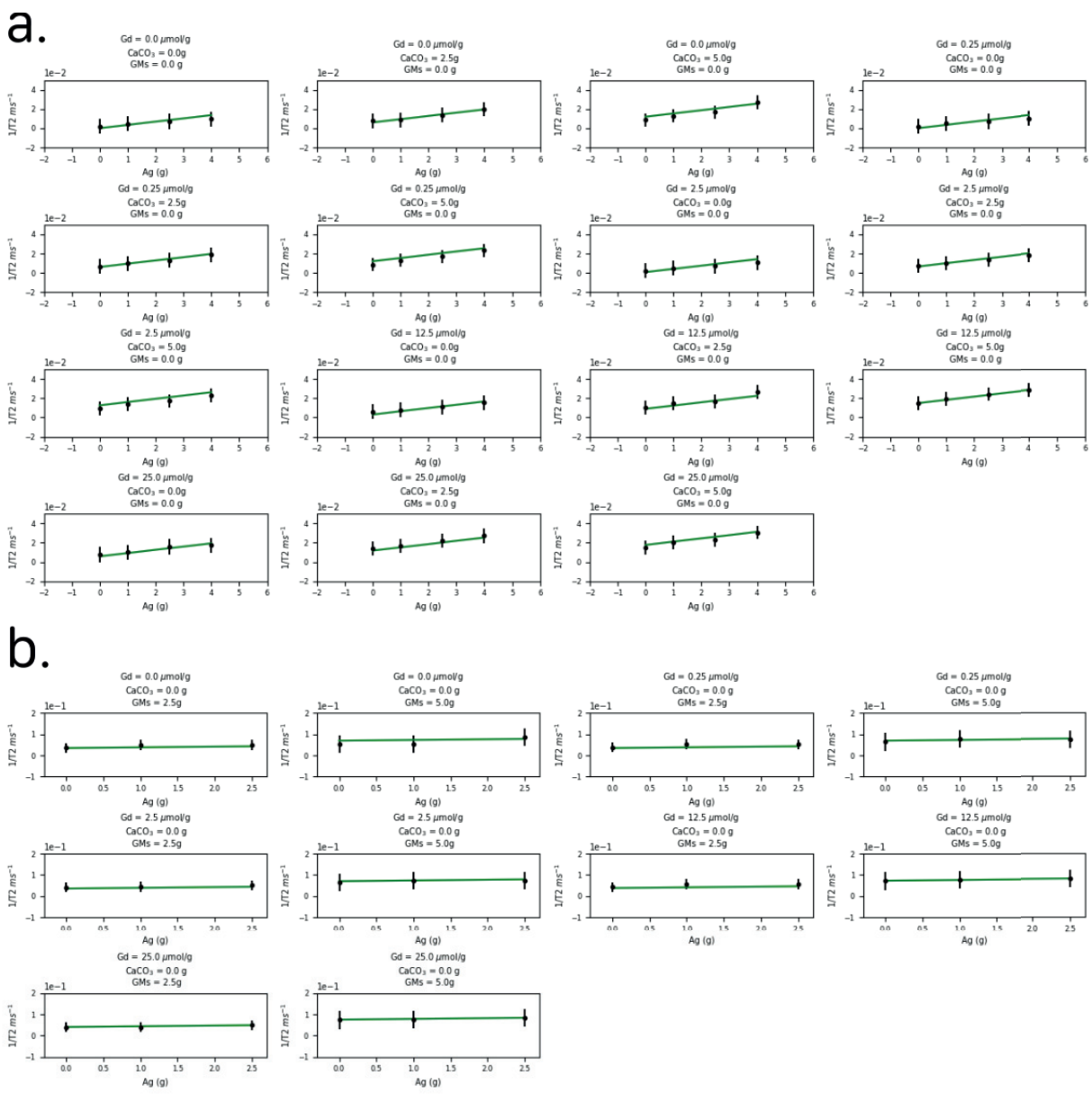
We have created a system of tissue-mimicking carrageenan-based materials for both MRI and CT imaging. The tissue-surrogates produced using this system of materials have good T_1/T_2 relaxation time and CT number agreement compared with published *in vivo* measurements. This system of materials can mimic MRI and CT characteristics for bone and a wide range of soft tissues such as adipose tissue and skeletal muscle.

2.7. Appendix I



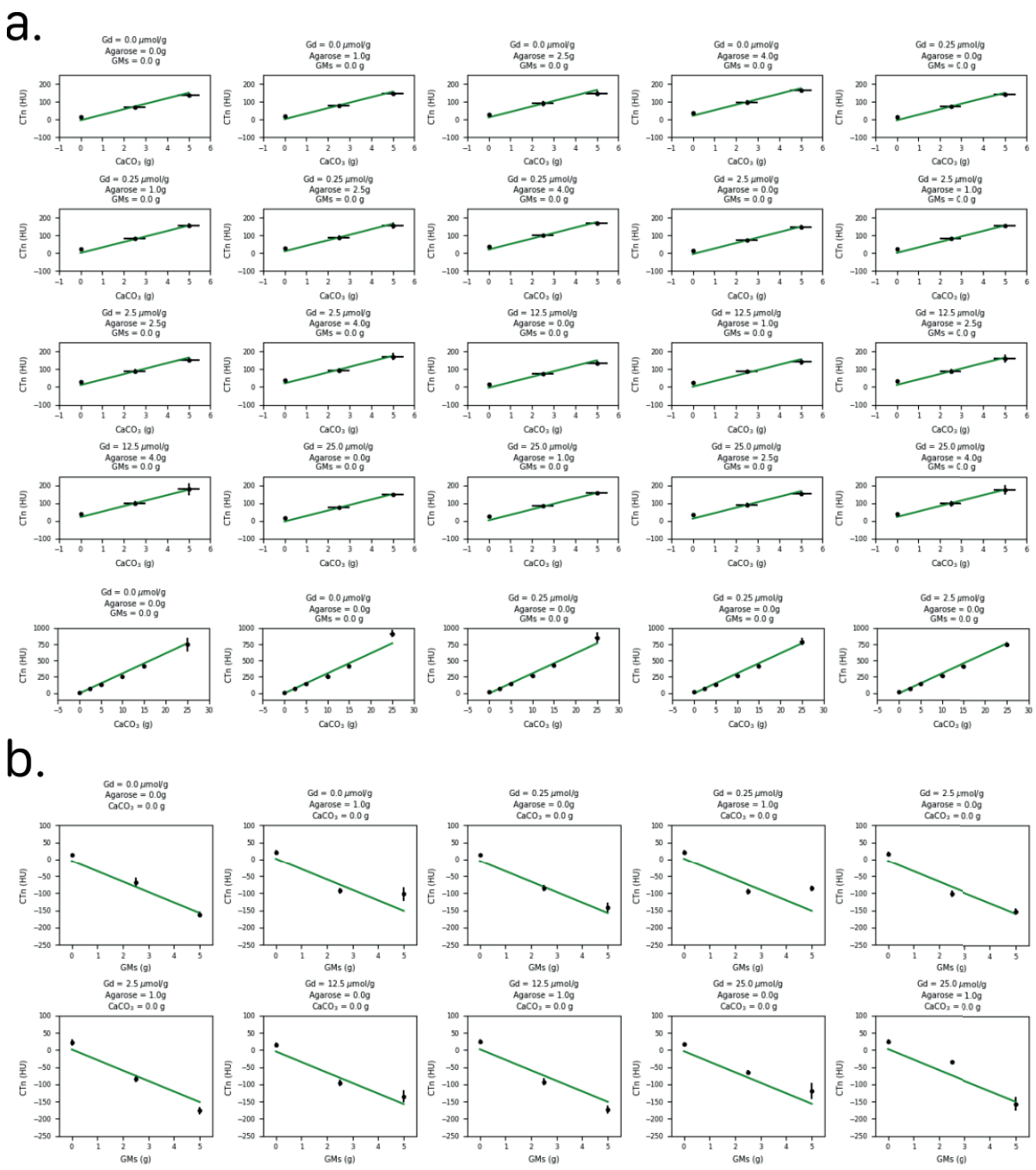
Appendix I. $1/T_1$ measurements for increasing Gd contrast concentrations (black dots), and the corresponding multivariate linear regression fits. Each point represents data generated from a 50g gel sample. Each plot in (a) represents a different combination of agarose and $CaCO_3$ concentrations. Each plot in figure (b) represents a different combination of agarose and GM concentrations. Each point represents data generated from a 50g gel sample.

2.8. Appendix II



Appendix II. $1/T_2$ measurements for increasing agarose concentrations (black dots), and the corresponding multivariate linear regression fits. Each point represents data generated from a 50g gel sample. Each plot in (a) represents a different combination of Gd contrast and $CaCO_3$ concentrations. Each plot in figure (b) represents a different combination of Gd contrast and GM concentrations.

2.9. Appendix III



Appendix III. (a) CT number measurements for increasing $CaCO_3$ concentrations (black dots), and the corresponding multivariate linear regression fits. Each point represents data generated from a 50g gel sample. Each plot in (a) represents a different combination of Gd contrast and agarose concentrations. (b) CT number measurements for increasing GM concentrations (black dots), and the corresponding multivariate linear regression fits. Each plot in (b) represents a different combination of Gd contrast and agarose concentrations.

2.10. Bibliography

1. Kocher KE, Meurer WJ, Fazel R, Scott PA, Krumholz HM, Nallamothu BK. National trends in use of computed tomography in the emergency department. *Ann Emerg Med*. 2011;58(5):452-462.e3. doi:10.1016/j.annemergmed.2011.05.020

2. Larson DB, Johnson LW, Schnell BM, Salisbury SR, Forman HP. National Trends in CT Use in the Emergency Department: 1995–2007. *Radiology*. 2011;258(1):164-173. doi:10.1148/radiol.10100640

3. Quaday KA, Salzman JG, Gordon BD. Magnetic resonance imaging and computed tomography utilization trends in an academic ED. *Am J Emerg Med*. 2014;32(6):524-528. doi:10.1016/j.ajem.2014.01.054

4. Scheinfeld MH, Moon J-Y, Fagan MJ, Davoudzadeh R, Wang D, Taragin BH. MRI usage in a pediatric emergency department: an analysis of usage and usage trends over 5 years. *Pediatr Radiol*. 2017;47(3):327-332. doi:10.1007/s00247-016-3764-y

5. Wachtel RE, Dexter F, Dow AJ. Growth rates in pediatric diagnostic imaging and sedation. *Anesth Analg*. 2009;108(5):1616-1621. doi:10.1213/ane.0b013e3181981f96

6. Gondi V, Pugh SL, Tome WA, et al. Preservation of memory with conformal avoidance of the hippocampal neural stem-cell compartment during whole-brain radiotherapy for brain metastases (RTOG 0933): a phase II multi-institutional trial. *J Clin Oncol Off J Am Soc Clin Oncol*. 2014;32(34):3810-3816. doi:10.1200/JCO.2014.57.2909

7. Bruzzone MG, D'Incerti L, Farina LL, Cuccarini V, Finocchiaro G. CT and MRI of brain tumors. *Q J Nucl Med Mol Imaging Off Publ Ital Assoc Nucl Med AIMN Int Assoc Radiopharmacol IAR Sect Soc Of*. 2012;56(2):112-137.

8. Zaorsky NG, Showalter TN, Ezzell GA, et al. ACR Appropriateness Criteria for external beam radiation therapy treatment planning for clinically localized prostate cancer, part II of II. *Adv Radiat Oncol*. 2017;2(3):437-454. doi:10.1016/j.adro.2017.03.003

9. Iball GR, Moore AC, Crawford EJ. A routine quality assurance test for CT automatic exposure control systems. *J Appl Clin Med Phys*. 2016;17(4):291-306. doi:10.1120/jacmp.v17i4.6165

10. Droege RT. A quality assurance protocol for CT scanners. *Radiology*. 1983;146(1):244-246. doi:10.1148/radiology.146.1.6849059

11. Chen C-C, Wan Y-L, Wai Y-Y, Liu H-L. Quality assurance of clinical MRI scanners using ACR MRI phantom: preliminary results. *J Digit Imaging*. 2004;17(4):279-284. doi:10.1007/s10278-004-1023-5

12. Firbank MJ, Harrison RM, Williams ED, Coulthard A. Quality assurance for MRI: practical experience. *Br J Radiol*. 2000;73(868):376-383. doi:10.1259/bjr.73.868.10844863

13. Paulson ES, Erickson B, Schultz C, Allen Li X. Comprehensive MRI simulation methodology using a dedicated MRI scanner in radiation oncology for external beam radiation treatment planning. *Med Phys*. 2015;42(1):28-39. doi:10.1118/1.4896096

14. Papanikolaou N, Saenz D. Quality assurance in MRI for radiotherapy planning. *Phys Med*. 2016;32:172. doi:10.1016/j.ejmp.2016.07.271

15. D'Souza WD, Madsen EL, Unal O, Vigen KK, Frank GR, Thomadsen BR. Tissue mimicking materials for a multi-imaging modality prostate phantom. *Med Phys*. 2001;28(4):688-700. doi:10.1118/1.1354998

16. Niebuhr NI, Johnen W, Güldaglar T, et al. Technical Note: Radiological properties of tissue surrogates used in a multimodality deformable pelvic phantom for MR-guided radiotherapy. *Med Phys*. 2016;43(2):908-916. doi:10.1118/1.4939874

17. He Y, Qin S, Dyer BA, et al. Characterizing mechanical and medical imaging properties of polyvinyl chloride-based tissue-mimicking materials. *J Appl Clin Med Phys*. 2019;20(7):176-183. doi:10.1002/acm2.12661

18. Li W, Belmont B, Greve JM, et al. Polyvinyl chloride as a multimodal tissue-mimicking material with tuned mechanical and medical imaging properties. *Med Phys*. 2016;43(10):5577-5592. doi:10.1118/1.4962649

19. Hattori K, Ikemoto Y, Takao W, et al. Development of MRI phantom equivalent to human tissues for 3.0-T MRI. *Med Phys*. 2013;40(3):032303. doi:10.1118/1.4790023

20. Kato H, Kuroda M, Yoshimura K, et al. Composition of MRI phantom equivalent to human tissues. *Med Phys*. 2005;32(10):3199-3208. doi:10.1118/1.2047807

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

22. Yoshimura K, Kato H, Kuroda M, et al. Development of a tissue-equivalent MRI phantom using carrageenan gel. *Magn Reson Med*. 2003;50(5):1011-1017. doi:10.1002/mrm.10619

23. Pedregosa F, et al. Scikit-learn: Machine Learning in Python. *J Mach Learn Res*. 2011;12:2825−2830.

24. CVX Research Inc. *CVX: Matlab Software for Disciplined Convex Programming, Version 2.0*.; 2012. http://cvxr.com/cvx.

25. Grant M, Boyd S. Graph implementations for nonsmooth convex programs. In: Blondel V, Boyd S, Kimura H, eds. *Recent Advances in Learning and Control*. Lecture Notes in Control and Information Sciences. Springer-Verlag Limited; 2008:95–110.

26. Gold GE, Han E, Stainsby J, Wright G, Brittain J, Beaulieu C. Musculoskeletal MRI at 3.0 T: Relaxation Times and Image Contrast. *Am J Roentgenol*. 2004;183(2):343-351. doi:10.2214/ajr.183.2.1830343

27. Chen Y, Jiang Y, Pahwa S, et al. MR Fingerprinting for Rapid Quantitative Abdominal Imaging. *Radiology*. 2016;279(1):278-286. doi:10.1148/radiol.2016152037

28. van der Werf A, Dekker IM, Meijerink MR, Wierdsma NJ, de van der Schueren MAE, Langius JAE. Skeletal muscle analyses: agreement between non-contrast and contrast CT scan measurements of skeletal muscle area and mean muscle attenuation. *Clin Physiol Funct Imaging*. 2018;38(3):366-372. doi:10.1111/cpf.12422

29. Ethofer T, Mader I, Seeger U, et al. Comparison of longitudinal metabolite relaxation times in different regions of the human brain at 1.5 and 3 Tesla. *Magn Reson Med*. 2003;50(6):1296-1301. doi:10.1002/mrm.10640

30. Jiang Y, Ma D, Seiberlich N, Gulani V, Griswold MA. MR fingerprinting using fast imaging with steady state precession (FISP) with spiral readout. *Magn Reson Med*. 2015;74(6):1621-1631. doi:10.1002/mrm.25559

31. Weinstein MA, Duchesneau PM, MacIntyre WJ. White and gray matter of the brain differentiated by computed tomography. *Radiology*. 1977;122(3):699-702. doi:10.1148/122.3.699

32. Preibisch C, Deichmann R. T1 mapping using spoiled FLASH-EPI hybrid sequences and varying flip angles. *Magn Reson Med*. 2009;62(1):240-246. doi:10.1002/mrm.21969

33. Lu H, Nagae-Poetscher LM, Golay X, Lin D, Pomper M, van Zijl PCM. Routine clinical brain MRI sequences for use at 3.0 Tesla. *J Magn Reson Imaging JMRI*. 2005;22(1):13-22. doi:10.1002/jmri.20356

34. de Bazelaire CMJ, Duhamel GD, Rofsky NM, Alsop DC. MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: preliminary results. *Radiology*. 2004;230(3):652-659. doi:10.1148/radiol.2303021331

35. Jadvar H, Ye W, Groshen S, Conti PS. [F-18]-fluorodeoxyglucose PET-CT of the normal prostate gland. *Ann Nucl Med*. 2008;22(9):787-793. doi:10.1007/s12149-008-0177-5

36. Bojorquez JZ, Bricq S, Acquitter C, Brunotte F, Walker PM, Lalande A. What are normal relaxation times of tissues at 3 T? *Magn Reson Imaging*. 2017;35(Supplement C):69-80. doi:10.1016/j.mri.2016.08.021

37. Edden RAE, Smith SA, Barker PB. Longitudinal and multi-echo transverse relaxation times of normal breast tissue at 3 Tesla. *J Magn Reson Imaging*. 2010;32(4):982-987. doi:10.1002/jmri.22306

38. Rakow-Penner R, Daniel B, Yu H, Sawyer-Glover A, Glover GH. Relaxation times of breast tissue at 1.5T and 3T measured using IDEAL. *J Magn Reson Imaging JMRI*. 2006;23(1):87-91. doi:10.1002/jmri.20469

39. Urata M, Kijima Y, Hirata M, et al. Computed tomography Hounsfield units can predict breast cancer metastasis to axillary lymph nodes. *BMC Cancer*. 2014;14:730. doi:10.1186/1471-2407-14-730

40. Kim WH, Kim CG, Kim D-W. Optimal CT Number Range for Adipose Tissue When Determining Lean Body Mass in Whole-Body F-18 FDG PET/CT Studies. *Nucl Med Mol Imaging*. 2012;46(4):294-299. doi:10.1007/s13139-012-0175-3

41. Jordan CD, Saranathan M, Bangerter NK, Hargreaves BA, Gold GE. Musculoskeletal MRI at 3.0 T and 7.0 T: a comparison of relaxation times and image contrast. *Eur J Radiol*. 2013;82(5):734-739. doi:10.1016/j.ejrad.2011.09.021

42. Elias EJ, Liao JH, Jara H, et al. Quantitative MRI analysis of craniofacial bone marrow in patients with sickle cell disease. *Am J Neuroradiol*. 2013;34(3):622-627. doi:10.3174/ajnr.A3240

43. Lee AY, Thomas S, Liauw S. Using Pelvic Bone Marrow Hounsfield Units to Predict Cytopenia During Anal Cancer Chemoradiation. *Int J Radiat Oncol • Biol • Phys*. 2016;96(2):E214-E215. doi:10.1016/j.ijrobp.2016.06.1129

44. Patrick S, Birur NP, Gurushanth K, Raghavan AS, Gurudath S. Comparison of gray values of cone-beam computed tomography with hounsfield units of multislice computed tomography: An in vitro study. *Indian J Dent Res*. 2017;28(1):66. doi:10.4103/ijdr.IJDR_415_16

45. Mitsiopoulos N, Baumgartner RN, Heymsfield SB, Lyons W, Gallagher D, Ross R. Cadaver validation of skeletal muscle measurement by magnetic resonance imaging and computerized tomography. *J Appl Physiol Bethesda Md 1985*. 1998;85(1):115-122. doi:10.1152/jappl.1998.85.1.115

46. Vivanco JF, Burgers TA, García-Rodríguez S, et al. Estimating the density of femoral head trabecular bone from hip fracture patients using computed tomography scan data. *Proc Inst Mech Eng [H]*. 2014;228(6):616-626. doi:10.1177/0954411914540285

47. Aubrey J, Esfandiari N, Baracos VE, et al. Measurement of skeletal muscle radiation attenuation and basis of its biological variation. *Acta Physiol*. 2014;210(3):489-497. doi:10.1111/apha.12224

48. Boris P, Bundgaard F, Olsen A. The CT (Hounsfield unit) number of brain tissue in healthy infants. A new reliable method for detection of possible degenerative disease. *Childs Nerv Syst ChNS Off J Int Soc Pediatr Neurosurg*. 1987;3(3):175-177.

49. In E, Naguib HE, Haider M. Mechanical stability analysis of carrageenan-based polymer gel for magnetic resonance imaging liver phantom with lesion particles. *J Med Imaging*. 2014;1(3):035502. doi:10.1117/1.JMI.1.3.035502

50. Ibrahim MA, Dublin AB. Magnetic Resonance Imaging (MRI) Gadolinium. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing; 2019. http://www.ncbi.nlm.nih.gov/books/NBK482487/. Accessed December 18, 2019.

51. Rethy A, Sæternes JO, Halgunset J, et al. Anthropomorphic liver phantom with flow for multimodal image-guided liver therapy research and training. *Int J Comput Assist Radiol Surg*. 2018;13(1):61-72. doi:10.1007/s11548-017-1669-3

52. SHIROTANI T. Realistic Torso Phantom for Calibration of in-vivo Transuranic-Nuclide Counting Facilities. *J Nucl Sci Technol*. 1988;25(11):875-883. doi:10.1080/18811248.1988.9735941

CHAPTER 3 - A Novel Anthropomorphic Multimodality Phantom for MRI-Based Radiation Therapy QA Testing

3.1. Abstract

Purpose: Increased utilization of MRI in radiation therapy has caused a growing need for phantoms that provide tissue-like contrast in both CT and MRI images. Such phantoms can be used to compare MRI-based processes with CT-based clinical standards. Here, we develop and demonstrate the clinical utility of a 3D-printed anthropomorphic pelvis phantom containing materials capable of T_1, T_2 and electron density matching for a clinically relevant set of soft tissues and bone.

Methods and Materials: The phantom design was based on a male pelvic anatomy template with thin boundaries separating tissue types. Slots were included to allow insertion of various dosimeters. The phantom structure was created using a 3D-printer. The tissue compartments were filled with carrageenan-based materials designed to match the T_1 and T_2 relaxation times and electron densities of the corresponding tissues. CT and MRI images of the phantom were acquired and used to compare phantom T_1 and T_2 relaxation times and electron densities to literature-reported values for human tissue. To demonstrate clinical utility, the phantom was used for end-to-end testing of an MRI-only treatment simulation and planning workflow. Based on a T_2-weighted MRI image, sCTs were created using a statistical decomposition algorithm (MRIPlanner, Spectronic Research AB, Sweden) and used for dose calculation during treatment planning of VMAT and 7-field IMRT prostate plans. The plans were delivered on a Truebeam STX (Varian

Medical Systems, Palo Alto, CA), with film and a 0.3 cc ion chamber used to measure the delivered dose. Doses calculated on the CT and sCTs were compared using common DVH metrics.

Results: T_1 and T_2 relaxation time and electron density measurements for muscle, prostate and bone agreed well with literature-reported *in vivo* measurements. Film analysis resulted in a 99.7% gamma-pass-rate (3.0%, 3.0mm) for both plans. The ion chamber-measured dose discrepancies at the isocenter were 0.36% and 1.67% for the IMRT and VMAT plans respectively. The differences in PTV D98% and D95% between plans calculated on the CT and 1.5T/3.0T-derived sCT images were under 3%.

Conclusions: The developed phantom provides tissue-like contrast on MRI and CT and can be used to validate MRI-based processes through comparison with standard CT-based processes.

3.2. Introduction

The use of Magnetic Resonance Imaging (MRI) is becoming increasingly prevalent in radiation therapy because it provides excellent soft tissue visualization [1–7]. Examples of the utilization of MRI in clinical practice include registration of planning CT and MRI for target delineation, treatments using MRI-guided treatment units, and MRI-only simulation/treatment planning [8–13].

Phantoms play a critical role in acceptance, commissioning, and periodic QA of linacs and imaging systems, and standardizing multisite clinical trial design [14-17] . Phantoms such as the American College of Radiology (ACR) large phantom are used for contrast, resolution and geometry tests of MRI-based treatments [13,14]. However, most commercially available phantoms do not mimic both MRI/CT tissue imaging characteristics for different human organs making it challenging to comprehensively test MRI-based processes such as sCT image generation [15–18].

Non-commercial multimodal MRI/CT phantoms have been developed to test MRI-based radiation therapy workflows. Niebuhr et al developed tissue surrogate materials for MRI and CT imaging and used them in an anthropomorphic phantom to test aspects of MRI-based workflows [19–21]. This phantom used agarose gels in combination with Gd-based contrast agents and NaF to simulate non-fat soft tissues. Vegetable oils were used to mimic adipose tissue. Pelvis bone was mimicked with petroleum jelly, K_2HPO_4, gypsum bandages, and a photopolymer-based 3D-printed structure. These materials were contained in a silicone-based deformable hollow organ structure. The organ structures were encased in a cylindrical PMMA case. However, the presence of NaF caused shadowing artifacts in T_2-weighted MRI images, and the phantom's design did not include features to facilitate ion chamber- or film-based dosimetry measurements. Sun et al developed a dedicated end-to-end QA phantom for MRI-only simulation[22]. The phantom contained two components, one for quantifying internal distortion and the second for dosimetry and

alignment evaluation. The dosimetry and alignment component consisted of a mineral oil-filled pelvic-shaped structure with simulated pelvic organs. Dosimetric and DRR-based alignment differences were compared between MRI-only and CT-based plans. This phantom used mineral oil to represent all soft tissues and could not mimic the different imaging characteristics of various pelvic tissue types such as adipose tissue and pelvic bone. Cunningham et al developed a pelvic end-to-end phantom for MRI-based radiation therapy applications [23]. Soft tissue organ structures were mimicked with silicone balloons filled with an $MnCl_2$ solution, and pelvic bone was mimicked with cast polyurethane structures. This phantom structure allowed for ion chamber and film measurements. This phantom did not use materials with T_2 relaxation times representative of human pelvic tissues.

This work reports on the development of an anthropomorphic phantom to test and validate MRI-based workflows. Descriptions of the design, fabrication and an example application of the phantom are included. The phantom was a 3D-printed structure imitating male pelvic geometry containing slots for isocentric dosimetry measurements. The phantom was designed to contain materials that produce tissue-like contrast for CT and 1.5T/3.0T MRI using a novel system of carrageenan-based materials to mimic soft tissues and bone. The MRI and CT tissue imaging characteristics were quantified using T_1 and T_2 mapping and CT number measurements. As an example application, the phantom was used for end-to-end testing of an MRI-only simulation workflow including evaluation of phantom-based sCT images, dosimetric accuracy, and alignment accuracy of SBRT VMAT and 7-field IMRT prostate plans.

3.3. Methods and Materials

3.3.A. Phantom Design and Fabrication

Phantom construction involved the creation of a 3D-printed structure imitating male pelvic geometry. The structure contained MRI/CT human tissue-mimicking materials (Figure *3.3.1*). The system of tissue-surrogate materials was based on a common carrageenan-based solution. The materials are solid at room temperature and liquify when heated above 45°C, allowing them to be formed into shapes. The phantom structure was 3D-printed using a computer-aided-design (CAD) model based on human pelvic anatomy. The structure contained rigidly separated compartments conforming to pelvic anatomy, allowing the tissue-mimicking materials to be poured in. The materials were poured into the phantom in 2cm layers allowing the previous layer to solidify.

The phantom structure design was based on human pelvic anatomy and modified into a 3D-printable CAD model with radiation therapy features. A typical prostate cancer patient CT image was used as a template for phantom design. The gross musculature, prostate, rectum, bladder, penile bulb, pelvic bones and adipose tissues were contoured on the patient image and used to define rigid tissue compartments in the phantom structure. Tissue segmentation was performed in MIM (MIM Software Inc., Beachwood, Ohio). The resultant contours were expanded to create 2mm thick walls creating the rigid tissue compartments separating tissue types in the 3D-printed structure (*Figure 3.3.1*). The rigid tissue compartment contours were imported as a mesh into Fusion360 CAD software (Autodesk, Inc, San Rafael, CA) in stereolithography (.stl) file format. The rigid tissue compartment mesh was smoothed using Laplacian smoothing without surface modification to remove self-intersecting vertices and simplify generation of the structure-defining triangular mesh[24]. The smoothed rigid tissue compartment mesh was imported into AutoCAD (Autodesk, Inc, San Rafael, CA) and converted into a solid structure. A slot was added to allow insertion of dosimeters. Two adapters were designed to allow for placement of a 0.3cc PTW Farmer ion chamber (PTW Freiburg GmbH, Germany), and radio-chromic film. Alignment

crosshair indentations and a flat base were added to facilitate reproducible phantom positioning (*Figure 3.3.1)*. The phantom structure was 3D-printed on a Raise3D Pro2 (Raise3D, Irvine, CA) using Polylactic Acid (PLA) filament. The dimensions of the printed phantom were 29 cm in the LR x 17cm in the AP x 6cm in the superior-inferior (SI) directions. The phantom was 3D-printed as a single piece. A closure lid was designed and 3D-printed to protect the internal components of the phantom. The phantom design was rescaled by a factor of 0.73 in the AP, LR and SI directions to accommodate the printer bed area of 31cm x 31cm.

A novel system of carrageenan-based tissue-mimicking materials for MRI and CT imaging was used in constructing the phantom [25]. Carrageenan was selected as a gelatinizer because it produces a solid-at-room-temperature gel and because of its minimal effects on T_1 and T_2 values [26,27]. Gd-contrast (gadofosveset trisodium, Lantheus Medical Imaging, Inc, North Billlerica, MA) and agarose (A1700 Agarose LE powder, Benchmark Scientific, Sayreville, NJ) were used as T_1 and T_2 relaxation time modifiers. $CaCO_3$ (LD Carlson Co., Kent, OH) and GMs (Alumilite, Kalamazoo, MI) were used as CT number modifiers [28]. Surrogate materials were used to mimic muscle, prostate, pelvic bone and adipose tissue (*Table 3.3.I)*. A carrageenan-based water surrogate material was developed for urinary bladder. The gross musculature and penile bulb structures were filled with the muscle-mimicking material. The prostate and rectum structures were filled with the prostate-mimicking material.

Mixtures to create 100g tissue mimicking material						
Tissue type	Gd solution (g)	Agarose (g	CaCO3 (g)	GMs (g)	arrageenan (g	Water (g)
			4.7	2.0	3.0	89.5
			3.5		3.0	92.0
Pelvic Bone	3.9×10^{-3}	1.8	22.0	0	3.0	73.2
Adipose Tissue	1.5×10^{-2}	0	0	8.0×10^{-1}	3.0	96.2
Urinary Bladder	0	0	0	0	2.0	98.0

Table 3.3.I. Summary of material mixtures required to create 100g mimics of skeletal muscle, prostate, trabecular bone, adipose tissue or urinary bladder.

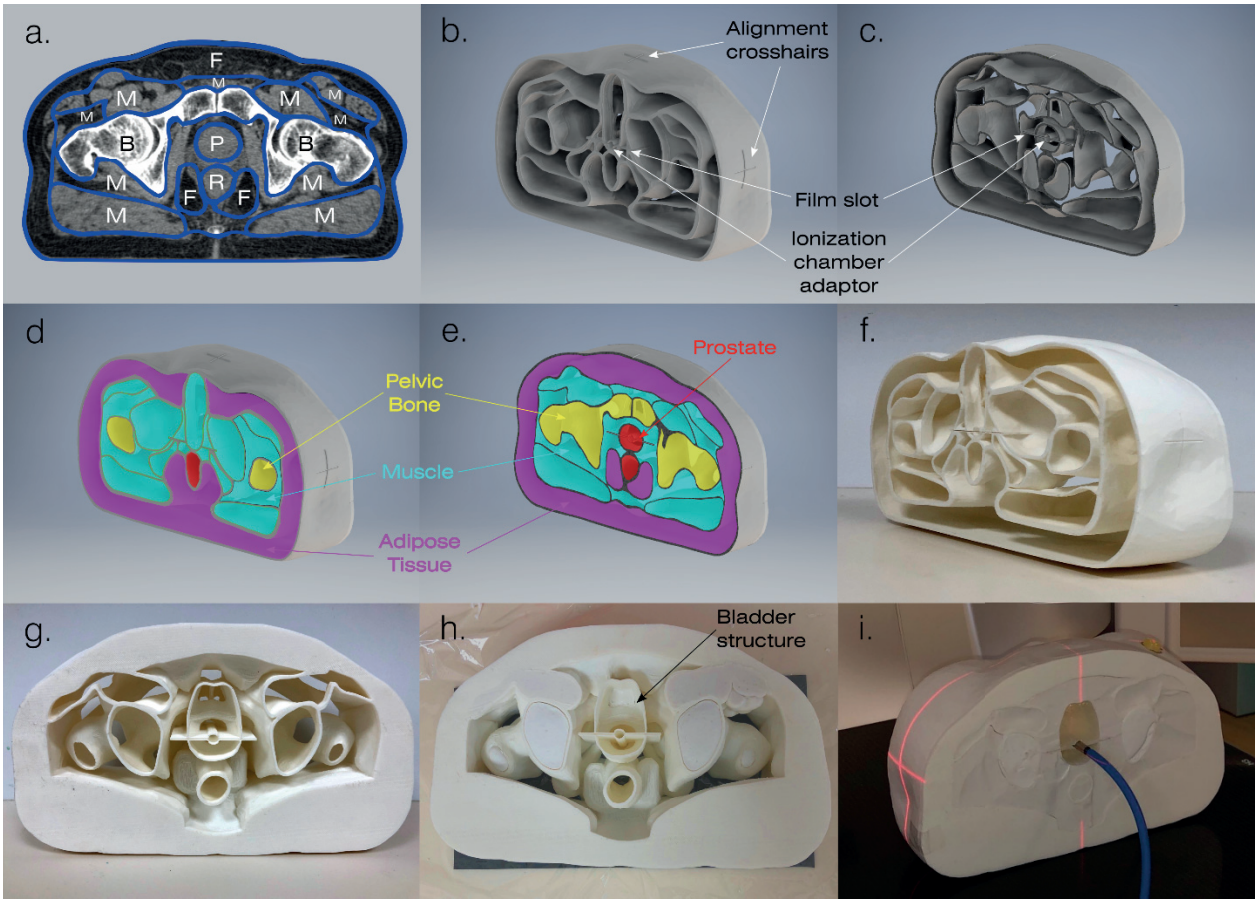


Figure 3.3.1. Major anatomical structures such as the gross musculature (M), adipose tissue (F), prostate (P), rectum (R) and pelvic bone (B) were segmented on a reference patient CT image (a). The front and midsection of the phantom CAD design are depicted in (b) and (c). The phantom design structure was filled with either adipose tissue, muscle, bone or prostate tissue-mimicking materials for MRI and CT imaging (d,e). The front and back of the completed 3D-printed phantom structure are shown in (f) and (g). The 3D-printed phantom structure partially filled with tissue-mimicking materials is shown in (h). The urinary bladder structure is labelled for reference. The completed phantom on a Truebeam STX couch with an ionization chamber inserted is shown in (i).

3.3.B. Validation of Phantom Imaging Characteristics

CT images were acquired on a 20-slice single source CT (Sensation Open, Siemens, Erlangen, Germany) with 120 kVp and 400 mA. T_1- and T_2-weighted TSE MRI images were acquired on 1.5T and 3.0T MRI systems (MAGNETOM Avanto and Skyra, Siemens, Erlangen, Germany). All

measurements were performed at 20 - 22° C. A 3D balanced steady-state free precession (bSSFP) MRI image was acquired on a real-time 0.35T MRI-guided radiation therapy system (MRIdian, Viewray Inc, Mountain View, CA, USA).

T_1 mapping measurements were done using an inversion-recovery TSE sequence with 180-degree refocusing pulses. The TEs and TRs were fixed at 12 ms and 15,000 ms respectively, and the TI was varied from 25 ms to 2,000 ms in 50 ms steps. Images were acquired with an echo-train-length of 5, no signal averaging, and no parallel imaging. A T_1 map was generated by performing a voxel-wise fit to the observed signal intensity according to the equation $M_{z,TI_i} = M_0\left(1 - 2e^{-TI/T_1}\right)$ where M_0 is the initial magnetization vector magnitude [29]. T_2 mapping measurements were acquired using a TSE sequence with a constant TR at 4,000 ms and effective TE values of 25, 50, 62, 75, 87 and 107 ms. Images were acquired with an echo-train-length of 5, no signal averaging, and no parallel imaging. T_2 maps were generated by performing a voxel-wise fit according to the equation $M_{TE_i} = M_0\left(e^{-\frac{TE_i}{T_2}}\right) + c$ where c is a noise offset parameter [29]. A 0.35T True Fast Imaging with Steady State Precession (TRUFI) image was acquired. The TRUFI sequence is a type of bSSFP sequence that yields images with T_2/T_1-weighted contrast. TRUFI images were acquired with a 60˚ flip angle, TR = 3.37 ms, TE = 1.45 ms, an echo train length of 1 and no signal averaging.

MRI/CT tissue imaging characteristics were quantified by measuring the T_1 and T_2 relaxation times and CT numbers of prostate, skeletal muscle, pelvic bone and adipose tissue mimics in the phantom via ROI analysis and compared to *in vivo* literature measurements. A 1.5 cm diameter circular ROI was used in the analysis. ROIs were placed in homogenous regions of

the urinary bladder, pelvic bone, adipose tissue, prostate and muscle mimicking materials. The same ROIs were used across all image sets.

3.3.C. Phantom Application: MRI-Only Simulation and Treatment Planning QA

End-to-end testing was performed for MRI-only simulation and treatment planning to demonstrate the utility of the phantom. MRI-only plans were created using phantom-generated sCT images. End-to-end tests were performed by comparing the planned and delivered dose differences between MRI-only and CT-based plans. Alignment differences between MRI-only and CT images were quantified and sCT image quality tests were performed. 1.5T and 3.0T-derived sCT images were generated from corresponding T_2 TSE MRI images. sCT images were generated using a well-validated statistical decomposition algorithm developed [30]. The mean-absolute-error (MAE) and the bone Dice Similarity Coefficient (DSC) were compared between the 1.5T/3.0T sCT and CT images.

sCT and CT plan dose differences were evaluated for a 6MV, 180 cGy, 7-field IMRT and 800 cGy SBRT VMAT plan. Plans were created in Eclipse version 15.6.05 (Varian Medical Systems, Palo Alto, CA). For treatment planning the femoral heads, rectum, and bladder were contoured as OARs, and the prostate was contoured as a target on the CT image. The PTV was defined by applying a 3 mm margin to the prostate. One IMRT and one VMAT plan was created using the CT image as a reference. The plans for both treatment types were normalized for 98% PTV coverage. The plans for both treatment types were copied and calculated on the 1.5T and 3.0T-derived sCT images. The MLC positions, monitor units (MU) and beam angles were preserved when the IMRT and VMAT plans were copied to the sCT images. The dose metric differences for plans calculated on the sCT and CT images are reported for both treatment types.

Alignment differences were quantified using sCT/CT-to-CBCT and DRR-to-X-ray pair alignments. Alignments were performed using translational bone-based shifts. The alignment difference, $\Delta P_{x,y,z}^{sCT,CT}$, between sCT/CT-to-CBCT registrations was evaluated using, $\Delta P_{x,y,z}^{sCT,CT} = |P_{x,y,z}^{sCT,CBCT} - P_{x,y,z}^{CT,CBCT}|$, where $P_{x,y,z}^{CT,CBCT}$ is the alignment resulting from registration between the CT and CBCT images and $P_{x,y,z}^{sCT,CBCT}$ is the alignment resulting from registration between the sCT and CBCT images. The alignment difference, $\Delta Q_{x,y,z}^{sCT,CT\ DRR}$, between DRR-to-X-ray pair alignments was evaluated using $\Delta Q_{x,y,z}^{sCT\ ,CT\ DRR} = |Q_{x,y,z}^{sCT\ DRR,X-ray} - Q_{x.y.z}^{CT\ DRR,X-ray}|$, where $Q_{x,y,z}^{sCT\ DRR,X-ray}$ and $Q_{x,y,z}^{CT\ DRR,X-ray}$ are the alignments resulting from registration of the orthogonal X-ray image pairs to sCT DRRs and CT DRRs respectively. The phantom was setup four times on the treatment couch. CBCT and X-ray image pairs were acquired for each setup.

IMRT and VMAT plans were delivered on a TrueBeam STX (Varian Medical Systems, Palo Alto, CA). During treatment delivery a 0.3 cc ionization chamber and radio-chromic film were placed in the phantom dosimeter slot and used to evaluate the absolute dose and spatial dose distribution, respectively. Gamma analysis was performed using the film measurements at 3%/3 mm and the corresponding pass rates (at a 30% dose threshold) are reported.

3.4. Results

3.4.A. Validation of Phantom Imaging Characteristics

Axial and sagittal CT, CBCT, 1.5T/3.0T TSE MRI and 1.5T/3.0T sCT images are presented in *Figure 3.4.1* A comparison between phantom and *in vivo* literature reported CT numbers for pelvic bone, adipose tissue, muscle and prostate tissues are reported in *Table 3.4.Ia* [31–34]. A comparison between the phantom and *in vivo* literature reported T_1 and T_2 relaxation times for pelvic bone,

adipose tissue, muscle and prostate tissues for measurements at 1.5T and 3.0T are reported in *Table 3.4.Ib* and *Table 3.4.Ic* [29,35–41]. The measured CT number for the urinary bladder-mimicking material was 11 ± 8.5 HU. The T_1 and T_2 relaxation times, measured at 1.5T, for the urinary bladder-mimicking material were 1888 ± 30 ms and 377 ± 43 ms, respectively. The T_1 and T_2 relaxation times, measured at 3.0T, for the urinary bladder mimicking material were 2114 ± 18 ms and 371 ± 39 ms, respectively. The maximum absolute differences in CT number between *in vivo* tissues and phantom tissue-mimicking materials was less than 2.8% for prostate, pelvic bone and muscle. The mean and maximum absolute differences in T_1- and T_2- relaxation times at 1.5T between *in vivo* tissues and phantom tissue-mimicking materials was 3.0% and 9.3%, respectively. The mean and maximum absolute differences in T_1- and T_2- relaxation times at 3.0T between *in vivo* tissues and phantom tissue-mimicking materials was 3.7% and 8.7%, respectively. Due to 3D-printing considerations, the phantom structure was truncated in the SI direction causing part of the bladder and rectum to be cut off. Therefore, contoured structures derived for these organs do not strictly conform to contouring guidelines [42].

ROI analysis was performed to assess MRI and CT contrast uniformity within the phantom. ROIs were drawn in the first and final slices in the phantom for skeletal muscle, trabecular bone and adipose tissues. The mean signal change for skeletal muscle, trabecular bone and adipose tissue was less than 1.9%, 2.2% and 0.5% in 3.0T T_1/T_2 TSE MRI images. The mean CT number difference for skeletal muscle, trabecular bone and adipose tissue was 0.1%, 0.1% and 1.0% respectively. To examine the reproducibility of the phantom materials, the phantom was built twice to evaluate the changes in T_1/T_2-relaxation times and CT numbers. The mean T_1- and T_2- relaxation times and CT number differences across all simulated tissue types was 46 ms, 12 ms and 11 HU respectively.

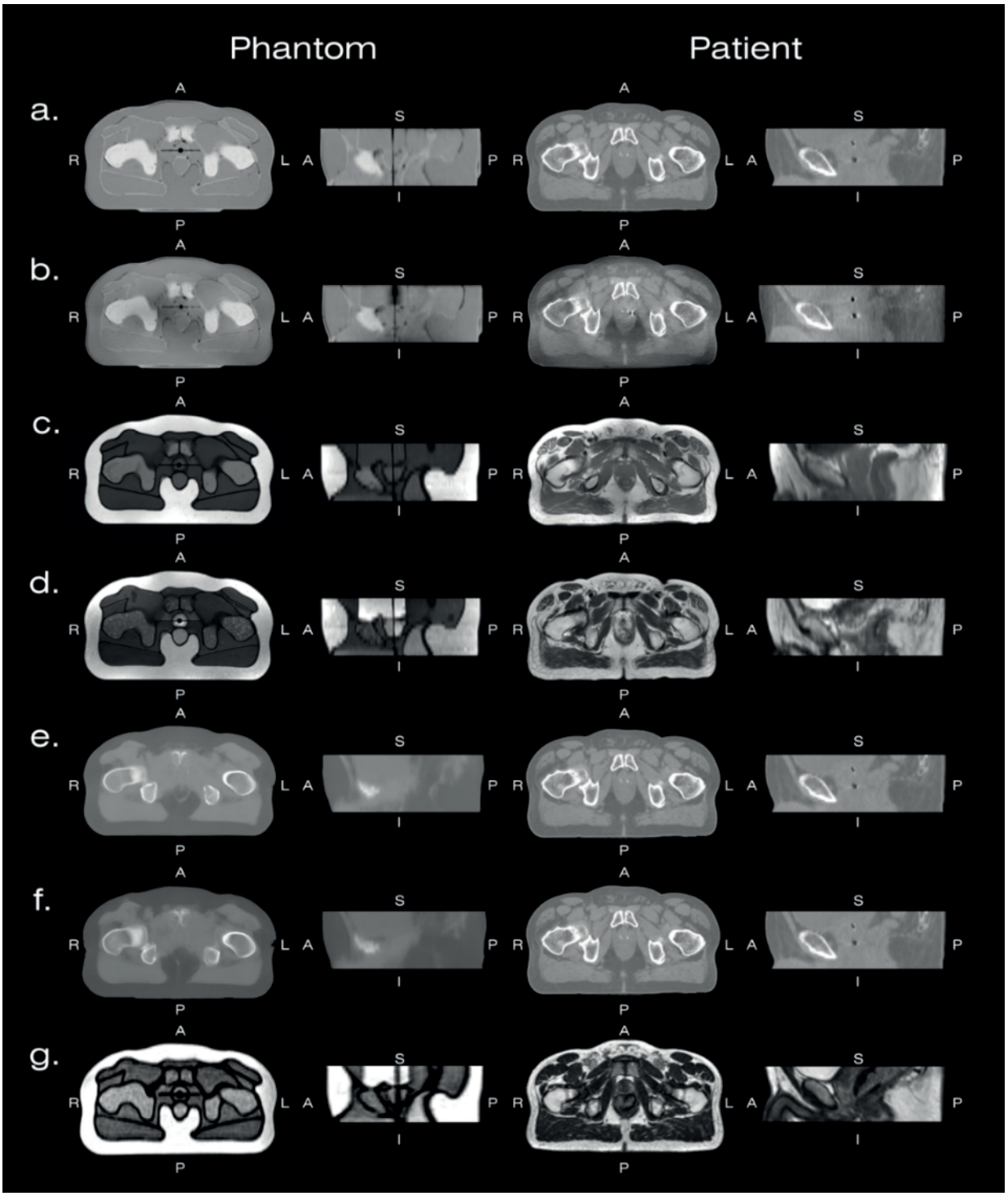


Figure 3.4.1. (a) CT, (b) CBCT, (c) 3.0T T1w TSE MRI, (d) 3.0T T2w TSE MRI, (g) 0.35T TRUFI MRI images of the phantom and an example patient in axial and sagittal planes. 1.5T and 3.0T-derived sCT images in the sagittal planes are presented in (e) and (f) with an example patient image for reference. The AP, LR and SI directions are labelled for reference.

	Tissue Type				
CT	$CTN_{phantom}$ (HU)	340 ± 29	-11 ± 8	39 ± 10	35 ± 11
	(HU)	376 ± 35		36 ± 3	36 ± 9
	(HU)	394 ± 35		37 ± 2	37 ± 10
	$CTN_{literature}$ (HU)	345 ± 21 [31]	-95 ± 9 [32]	38 [33]	36 [34]
1.5T		547 ± 17			
	$T1_{literature}$ (ms)	549 ± 42 [35]	96 ± 12.9 [36]	1042 ± 163 [39]	1317 ± 35 [35]
	$_{literature}$ (ms)	47 ± 2 [35]			
c.					
3.0T		597 ± 10			
	(ms)				
	$T2_{phantom}$ (ms)	50 ± 5	71 ± 3	40 ± 2	85 ± 3
	(ms)	49 ± 8 [35]			

Table 3.4.I. Table of CT numbers (a), 1.5T/3.0T T1 (b) and 1.5T/3.0T T2 (c) relaxation times ROI measurements from the phantom for trabecular bone, adipose tissues, skeletal muscle and prostate tissues. Sample in vivo literature values are presented for comparison. The CT numbers from the 1.5T- and 3.0T-derived sCT images are presented for reference.

3.4.B. Phantom Application: MRI-Only Simulation and Treatment Planning QA

The MAEs between the 1.5T/3.0T sCTs and corresponding CTs were 30 and 32 HU respectively. The bone DSC scores between the 1.5T/3.0T sCTs and corresponding CTs were 0.83 and 0.81 respectively. The sample OAR and PTV dose differences for the VMAT and IMRT doses calculated on the 1.5T/3.0T sCT, and CT images are reported in *Table 3.4.II*. The largest PTV D95% difference between plans calculated on the sCT and CT images was 2.9%. *Figure 3.4.2* shows dose distributions for the IMRT and VMAT plans overlaid on the reference CT image.

DRR images derived from CT and sCT and, projection X-ray images are shown in *Figure 3.4.3*. Alignment discrepancies between the 1.5T/3.0T sCT to CT and 1.5T/3.0T sCT DRRs to kV/kV orthogonal pair images are reported in *Table 3.4.III*. AP DRR and X-ray projections of the phantom are shown in *Figure 3.4.3*. The gamma pass rates and measured point dose differences were computed using the planned doses on the CT image as a reference. The measured percent point dose differences and 3%/3mm gamma pass rates for the delivered IMRT plan were 0.36% and 99.7% ± 0.5% respectively. The measured percent point dose differences and 3%/3mm gamma pass rates for the delivered VMAT plan were 1.67% and 99.7% ± 0.6% respectively.

	IMRT 1.5T sCT - CT	IMRT 3.0T sCT - CT	VMAT 1.5T sCT - CT	VMAT 3.0T sCT - CT
PTV D98% (%)	2.4	1.6	2.8	1.9
PTV D95% (%)	2.6	1.8	2.9	2
Rectum V50% (cc)	0.3	0.2	0.3	0.3
Rectum V70% (cc)	1.2	1.2	1.2	1.2
Bladder V65% (cc)	0.6	0.2	0.8	1.1
Bladder V80% (cc)	0.6	0.2	0.7	0.9
Left Femur D_{max} (%)	0.4	0.4	2.1	2
Right Femur D_{max} (%)	0.6	0.5	0.1	0.1

Table 3.4.II. Comparison of dose differences between doses calculated on the 1.5T/3.0T sCT and CT images. The OAR volume dose metrics are reported as the volume that receives equal to or more than the reported percentage of the prescription dose.

	DRR to X-ray pair			CT to CBCT		
	$\Delta Q_{LR}^{sCT,CT\ DRR}$ (mm)	$\Delta Q_{AP}^{sCT,CT\ DRR}$ (mm)	$\Delta Q_{SI}^{sCT,CT\ DRR}$ (mm)	$\Delta P_{LR}^{sCT,CT\ DRR}$ (mm)	$\Delta P_{AP}^{sCT,CT\ DRR}$ (mm)	$\Delta P_{SI}^{sCT,CT\ DRR}$ (mm)
1.5T sC	0.7 ± 0.5	0.8 ± 0.4	0.7 ± 0.9	0.1 ± 0.2	0.6 ± 0.4	0.7 ± 1.1
	0.6 ± 0.5	0.3 ± 0.3	0.8 ± 0.5	0.2 ± 0.2	0.8 ± 0.8	0.7 ± 0.5

Table 3.4.III. Summary result of the mean alignment discrepancies between orthogonal X-ray image pairs and, DRR images generated from the planning CT images. The alignment discrepancies between CBCT images and, planning CT and 1.5T/3.0T sCT images are reported. Alignment discrepancies are reported in the AP, LR and SI directions. These results are based on four separate phantom setups.

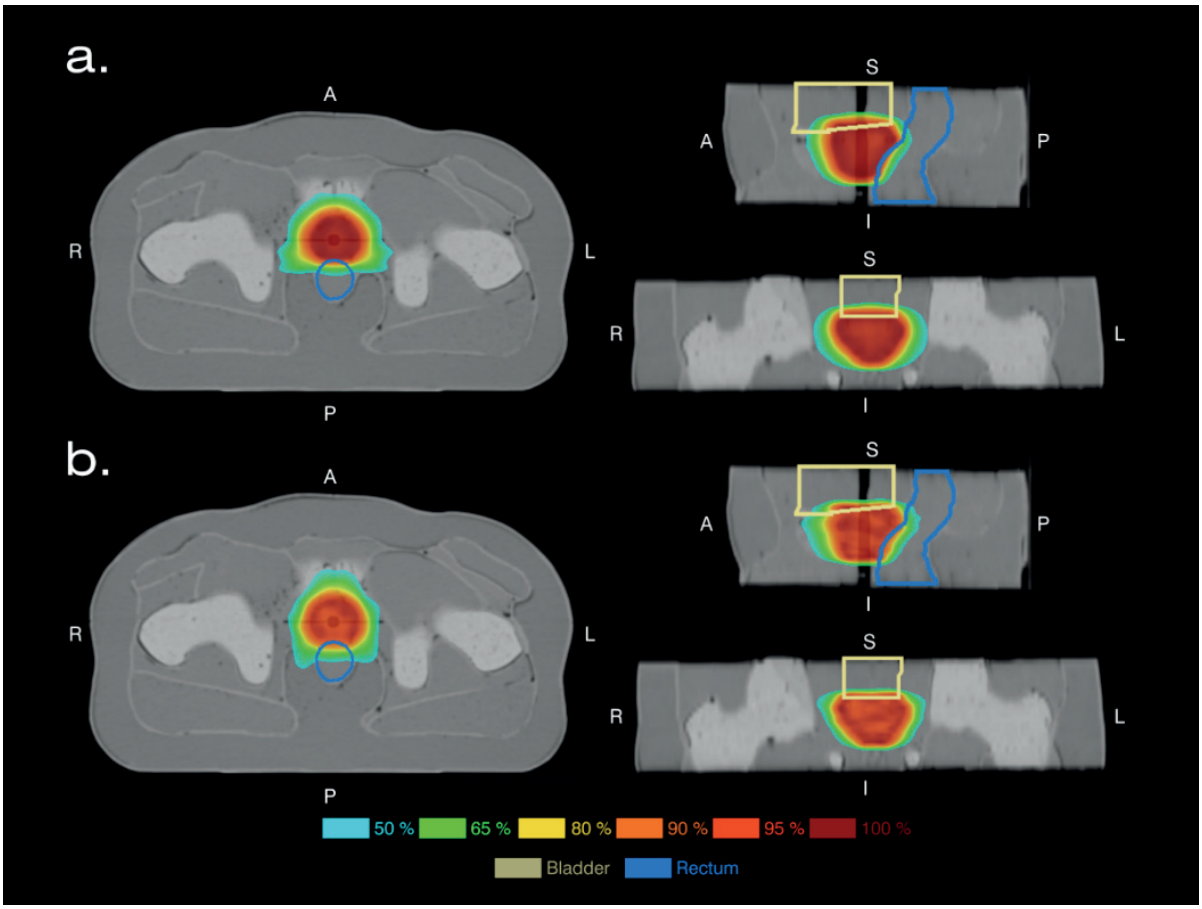


Figure 3.4.2. Dose color washes showing planned doses for the (a) IMRT and (b) VMAT plans on the reference CT image. Doses from 50% to 100% of the prescription dose are shown. The phantom bladder and rectum structures are shown. The AP, LR and SI directions are labelled for reference.

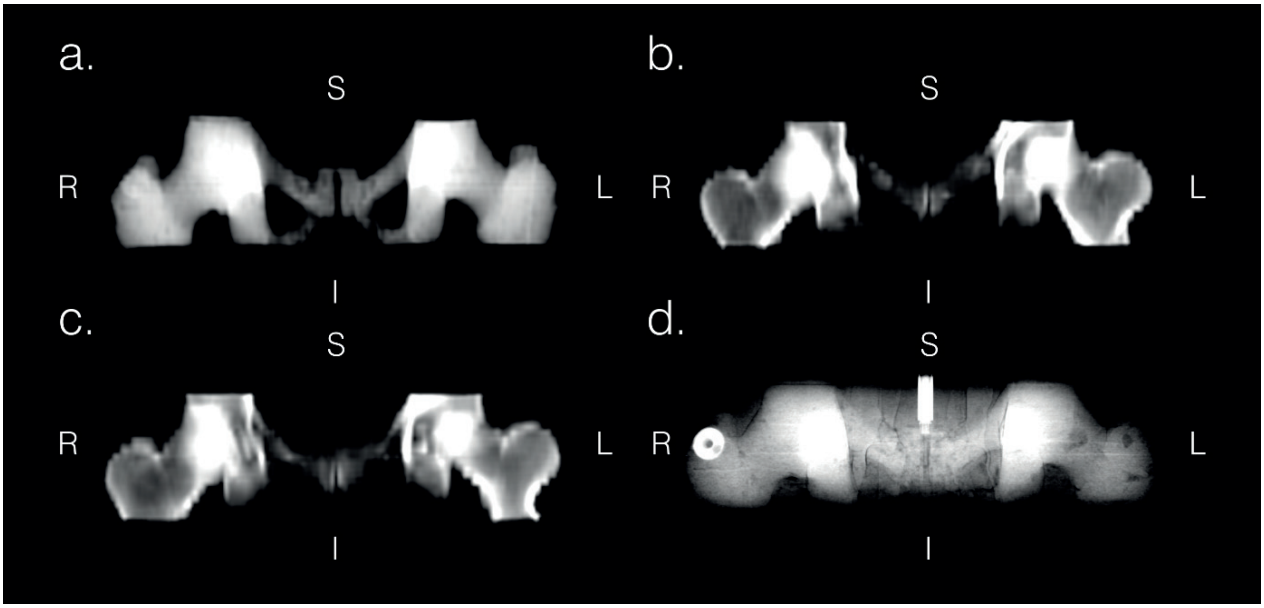


Figure 3.4.3. AP DRRs derived using phantom CT image (a), 1.5T sCT (b) and 3.0T sCT images (c). AP X-ray projection image of phantom taken on treatment unit (d). The LR and SI directions are labelled for reference.

3.5. Discussion

In this study a realistic anthropomorphic pelvis phantom was developed for QA testing MRI-based radiation therapy techniques. The tissue imaging characteristics of the phantom for 1.5T/3.0T MRI and CT were validated by physical measurements and tested by creating sCT images using the phantom. The CT numbers for bone, muscle and prostate tissues agreed well with *in vivo* literature measurements. The maximum absolute percent difference between the 1.5T and 3.0T T_1 values and T_2 values was 8.7% and 8.3% respectively. As an example application, the phantom was used for end-to-end testing of an MRI-only simulation and treatment planning workflow. The results demonstrated that the maximum CT-sCT MAE difference was 32HU and DSC score was 0.81. This is consistent with 36.5 ± 4.1 MAE values reported by the developers of the tested sCT method [30]. All observed planned dose metric variations between sCT and CT plans were under 3%. The alignment differences between MRI and CT-based images was less than 2 mm in each plane. The point dose differences were under 1.7% and the gamma-pass-rate was 99.7% at 3% / 3 mm for both delivered plans. These differences are small enough for clinical acceptability.

The main novelties of this phantom are that it is simpler to fabricate compared to previously published methods, produces realistic anthropomorphic CT and, T_1- and T_2-weighted MRI images, and can produce realistic sCT images using a commercially available sCT generation method. The phantom fabrication process of previously published phantoms have typically used deformable silicone balloons to encase the tissue-mimicking materials and professionally machined rigid structures [20–23]. Creation of the encasing structure of the phantom in this study only requires the

use of a 3D-printer. The tissue-mimicking materials used in previously published phantoms require a diverse range of materials to mimic different organs such as agarose gels doped with Gd-based contrast agents and NaF to simulate non-fat soft tissues, vegetable oils to mimic adipose tissue and petroleum jelly, K_2HPO_4, gypsum bandages, and a photopolymer-based 3D-printed structure to mimic pelvic bone [20,21]. The system of tissue mimicking materials presented in this work uses a single carrageenan-based material with additives to control CT number, T_1- and T_2-relaxation times. Additionally, the system of tissue-mimicking materials presented in this work does not exhibit shadowing artifacts allowing the generation of a relatively artifact free T_1- or T_2-weighted MRI images.

The stability of the phantoms imaging and physical characteristics over time were not directly investigated in this study. The phantom was tested for a period of 12 weeks without discernably exhibiting physical degradation. Previous studies have investigated the stability of carrageenan-based materials over time such as In et al who quantified the mechanical stability and T_1/T_2 relaxation time stability of carrageenan-based materials with agar and Gd contrast additives [43]. The density and compressive strength changes over 6 weeks were relatively stable for samples with 3% carrageenan as used in this study. The 3.0T T_1 relaxation times of 8 tested samples generally increased by 1.25 times compared to baseline after 5 weeks. The 3.0T T_2 relaxation times of eight tested samples generally decreased 0.75 to 0.95 times compared to baseline after 5 weeks [43].

There are limitations to the materials used in the phantom presented in this work. The adipose tissue surrogate CT number had the largest mean difference compared to in vivo measurements. Further work needs to be done to improve CT number matching of the adipose tissue surrogate material, possibly by using a mixture of oil and GMs with the carrageenan-based

materials. Another limitation is that the pelvic bone was treated as a single homogenous material, whereas human bone can be very heterogenous. The marrow cavity was excluded because it was challenging to 3D-print a small fillable cavity within the pelvic bone.

There are limitations to the phantom design. The phantom structure was designed to be 3D-printed as a single piece. This meant that the phantom structure needed to be rescaled to fit on the 3D-printer bed. For the purposes of this study, we do not think there are any major negative consequences to using a phantom that is smaller than typical adult male anatomy. If this phantom sees wider clinical use, industrial 3D-printers with larger print beds could be used to print phantoms true to the size of human anatomy without the need for design rescaling. Another drawback to the phantom design was that the phantom rectum and bladder structures do not allow for exact conformance to RTOG contouring guidelines due to the limited SI extent [42]. The SI extent was reduced because it reduced 3D-printing time and simplified the phantom design. This limitation can be remedied in future phantoms by using industrial 3D-printers which could allow rapid manufacturing of phantoms with larger craniocaudal extents or printing the phantom structure as interlocking modules.

The phantom could be used to objectively compare sCT image generation methods that require T_1- or T_2-weighted MRI image inputs. There are sCT algorithms that have been developed for the male pelvis site that use atlas-based, statistical decomposition algorithmic and, voxel-based deep- and machine-learning methods [16,30,44,46–48]. In this study we tested a single commercially available statistical decomposition algorithm-based sCT generator that takes a T_2-weighted MRI input image. Further testing is required to evaluate the phantoms performance with other sCT methods. The phantom was designed to match only the CT numbers, T_1- and T_2-relaxation times of human pelvic anatomy. The phantom materials were not designed to mimic the T_2* relaxation

times of pelvic bone or the chemical shift behaviors of adipose tissue and water. It is not expected that the phantom will be able to produce clinically realistic sCT images with methods that use multi-point Dixon (mDixon) or ultrashort echo-time (UTE) images as inputs [49,50].

3.6. Conclusions

The developed anthropomorphic pelvis phantom provides tissue-like contrast on T_1- and T_2-weighted MRI and CT and can be used to validate MRI-based radiation therapy processes through comparison with standard CT-based radiation therapy processes.

3.7. Bibliography

1. Rasch C, Barillot I, Remeijer P, Touw A, van Herk M, Lebesque JV. Definition of the prostate in CT and MRI: a multi-observer study. *Int J Radiat Oncol Biol Phys*. 1999;43(1):57-66.

2. Roach M, Faillace-Akazawa P, Malfatti C, Holland J, Hricak H. Prostate volumes defined by magnetic resonance imaging and computerized tomographic scans for three-dimensional conformal radiotherapy. *Int J Radiat Oncol*. 1996;35(5):1011-1018. doi:10.1016/0360-3016(96)00232-5

3. McLaughlin PW, Evans C, Feng M, Narayana V. Radiographic and anatomic basis for prostate contouring errors and methods to improve prostate contouring accuracy. *Int J Radiat Oncol Biol Phys*. 2010;76(2):369-378. doi:10.1016/j.ijrobp.2009.02.019

4. Schmidt MA, Payne GS. Radiotherapy planning using MRI. *Phys Med Biol*. 2015;60(22):R323. doi:10.1088/0031-9155/60/22/R323

5. Lee YK, Bollet M, Charles-Edwards G, et al. Radiotherapy treatment planning of prostate cancer using magnetic resonance imaging alone. *Radiother Oncol J Eur Soc Ther Radiol Oncol*. 2003;66(2):203-216.

6. Metcalfe P, Liney GP, Holloway L, et al. The Potential for an Enhanced Role for MRI in Radiation-therapy Treatment Planning. *Technol Cancer Res Treat*. 2013;12(5):429-446. doi:10.7785/tcrt.2012.500342

7. McLaughlin PW, Narayana V, Meirovitz A, et al. Vessel-sparing prostate radiotherapy: dose limitation to critical erectile vascular structures (internal pudendal artery and corpus cavernosum) defined by MRI. *Int J Radiat Oncol Biol Phys*. 2005;61(1):20-31. doi:10.1016/j.ijrobp.2004.04.070

8. Mutic S, Dempsey JF. The ViewRay System: Magnetic Resonance–Guided and Controlled Radiotherapy. *Semin Radiat Oncol*. 2014;24(3):196-199. doi:10.1016/j.semradonc.2014.02.008

9. Olberg S, Green O, Cai B, et al. Optimization of treatment planning workflow and tumor coverage during daily adaptive magnetic resonance image guided radiation therapy (MR-IGRT) of pancreatic cancer. *Radiat Oncol Lond Engl*. 2018;13. doi:10.1186/s13014-018-1000-7

10. Winkel D, Bol GH, Kroon PS, et al. Adaptive radiotherapy: The Elekta Unity MR-linac concept. *Clin Transl Radiat Oncol*. 2019;18:54-59. doi:10.1016/j.ctro.2019.04.001

11. Tijssen RHN, Philippens MEP, Paulson ES, et al. MRI commissioning of 1.5T MR-linac systems – a multi-institutional study. *Radiother Oncol*. 2019;132:114-120. doi:10.1016/j.radonc.2018.12.011

12. Independent Phantom Irradiation Is Both Necessary and Cost Effective for Clinical Trial Credentialing. *Int J Radiat Oncol • Biol • Phys*. 2015;92(3):501-503. doi:10.1016/j.ijrobp.2015.01.021

13. Tyagi N, Fontenla S, Zelefsky M, et al. Clinical workflow for MR-only simulation and planning in prostate. *Radiat Oncol*. 2017;12:119. doi:10.1186/s13014-017-0854-4

14. Fatemi A, Taghizadeh S, Yang CC, R. Kanakamedala M, Morris B, Vijayakumar S. Machine-Specific Magnetic Resonance Imaging Quality Control Procedures for Stereotactic Radiosurgery Treatment Planning. *Cureus*. 9(12). doi:10.7759/cureus.1957

15. Mutic S, Palta JR, Butker EK, et al. Quality assurance for computed-tomography simulators and the computed-tomography-simulation process: Report of the AAPM Radiation Therapy Committee Task Group No. 66. *Med Phys*. 2003;30(10):2762-2792. doi:10.1118/1.1609271

16. Maspero M, Savenije MHF, Dinkla AM, et al. Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy. *Phys Med Biol*. 2018;63(18):185001. doi:10.1088/1361-6560/aada6d

17. Arabi H, Dowling JA, Burgos N, et al. Comparative study of algorithms for synthetic CT generation from MRI: Consequences for MRI-guided radiation planning in the pelvic region. *Med Phys*. 2018;45(11):5218-5233. doi:10.1002/mp.13187

18. Liu L, Jolly S, Cao Y, Vineberg K, Fessler JA, Balter JM. Female pelvic synthetic CT generation based on joint intensity and shape analysis. *Phys Med Biol*. 2017;62(8):2935-2949. doi:10.1088/1361-6560/62/8/2935

19. Kraus KM, Jäkel O, Niebuhr NI, Pfaffenberger A. Generation of synthetic CT data using patient specific daily MR image data and image registration. *Phys Med Biol*. 2017;62(4):1358–1377. doi:10.1088/1361-6560/aa5200

20. Niebuhr NI, Johnen W, Güldaglar T, et al. Technical Note: Radiological properties of tissue surrogates used in a multimodality deformable pelvic phantom for MR-guided radiotherapy. *Med Phys*. 2016;43(2):908-916. doi:10.1118/1.4939874

21. Niebuhr NI, Johnen W, Echner G, et al. The ADAM-pelvis phantom-an anthropomorphic, deformable and multimodal phantom for MRgRT. *Phys Med Biol*. 2019;64(4):04NT05. doi:10.1088/1361-6560/aafd5f

22. Sun J, Dowling J, Pichler P, et al. MRI simulation: end-to-end testing for prostate radiation therapy using geometric pelvic MRI phantoms. *Phys Med Biol*. 2015;60(8):3097–3109. doi:10.1088/0031-9155/60/8/3097

23. Cunningham JM, Barberi EA, Miller J, Kim JP, Glide-Hurst CK. Development and evaluation of a novel MR-compatible pelvic end-to-end phantom. *J Appl Clin Med Phys*. 2019;20(1):265-275. doi:10.1002/acm2.12455

24. Liu T, Chen M, Song Y, Li H, Lu B. Quality improvement of surface triangular mesh using a modified Laplacian smoothing approach avoiding intersection. *PLOS ONE*. 2017;12(9):e0184206. doi:10.1371/journal.pone.0184206

25. Singhrao K, Fu J, Gao Y, Hu P, Yang Y, Lewis JH. OC-0186 A system of materials capable of mimicking soft tissues and bone with both CT and MR imaging. *Radiother Oncol*. 2019;133:S94-S95. doi:10.1016/S0167-8140(19)30606-1

26. Hattori K, Ikemoto Y, Takao W, et al. Development of MRI phantom equivalent to human tissues for 3.0-T MRI. *Med Phys*. 2013;40(3):032303. doi:10.1118/1.4790023

27. Yoshimura K, Kato H, Kuroda M, et al. Development of a tissue-equivalent MRI phantom using carrageenan gel. *Magn Reson Med*. 2003;50(5):1011-1017. doi:10.1002/mrm.10619

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

29. Bojorquez JZ, Bricq S, Acquitter C, Brunotte F, Walker PM, Lalande A. What are normal relaxation times of tissues at 3 T? *Magn Reson Imaging*. 2017;35(Supplement C):69-80. doi:10.1016/j.mri.2016.08.021

30. Siversson C, Nordström F, Nilsson T, et al. Technical Note: MRI only prostate radiotherapy planning using the statistical decomposition algorithm. *Med Phys*. 2015;42(10):6090-6097. doi:10.1118/1.4931417

31. Patrick S, Birur NP, Gurushanth K, Raghavan AS, Gurudath S. Comparison of gray values of cone-beam computed tomography with hounsfield units of multislice computed tomography: An in vitro study. *Indian J Dent Res*. 2017;28(1):66. doi:10.4103/ijdr.IJDR_415_16

32. Kim WH, Kim CG, Kim D-W. Optimal CT Number Range for Adipose Tissue When Determining Lean Body Mass in Whole-Body F-18 FDG PET/CT Studies. *Nucl Med Mol Imaging*. 2012;46(4):294-299. doi:10.1007/s13139-012-0175-3

33. van der Werf A, Dekker IM, Meijerink MR, Wierdsma NJ, de van der Schueren MAE, Langius JAE. Skeletal muscle analyses: agreement between non-contrast and contrast CT scan measurements of skeletal muscle area and mean muscle attenuation. *Clin Physiol Funct Imaging*. 2018;38(3):366-372. doi:10.1111/cpf.12422

34. Jadvar H, Ye W, Groshen S, Conti PS. [F-18]-fluorodeoxyglucose PET-CT of the normal prostate gland. *Ann Nucl Med*. 2008;22(9):787-793. doi:10.1007/s12149-008-0177-5

35. de Bazelaire CMJ, Duhamel GD, Rofsky NM, Alsop DC. MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: preliminary results. *Radiology*. 2004;230(3):652-659. doi:10.1148/radiol.2303021331

36. Rakow-Penner R, Daniel B, Yu H, Sawyer-Glover A, Glover GH. Relaxation times of breast tissue at 1.5T and 3T measured using IDEAL. *J Magn Reson Imaging JMRI*. 2006;23(1):87-91. doi:10.1002/jmri.20469

37. Chen Y, Jiang Y, Pahwa S, et al. MR Fingerprinting for Rapid Quantitative Abdominal Imaging. *Radiology*. 2016;279(1):278-286. doi:10.1148/radiol.2016152037

38. Fennessy FM, Fedorov A, Gupta SN, Schmidt EJ, Tempany CM, Mulkern RV. Practical considerations in T1 mapping of prostate for dynamic contrast enhancement pharmacokinetic analyses. *Magn Reson Imaging*. 2012;30(9):1224-1233. doi:10.1016/j.mri.2012.06.011

39. Barral JK, Gudmundson E, Stikov N, Etezadi-Amoli M, Stoica P, Nishimura DG. A robust methodology for in vivo T1 mapping. *Magn Reson Med*. 2010;64(4):1057-1067. doi:10.1002/mrm.22497

40. Gold GE, Han E, Stainsby J, Wright G, Brittain J, Beaulieu C. Musculoskeletal MRI at 3.0 T: Relaxation Times and Image Contrast. *Am J Roentgenol*. 2004;183(2):343-351. doi:10.2214/ajr.183.2.1830343

41. Yao L, Gai N. Fat-Corrected T2 Measurement as a Marker of Active Muscle Disease in Inflammatory Myopathy. *Am J Roentgenol*. 2012;198(5):W475-W481. doi:10.2214/AJR.11.7113

42. Gay HA, Barthold HJ, O'Meara E, et al. Pelvic normal tissue contouring guidelines for radiation therapy: a Radiation Therapy Oncology Group consensus panel atlas. *Int J Radiat Oncol Biol Phys*. 2012;83(3):e353-362. doi:10.1016/j.ijrobp.2012.01.023

43. In E, Naguib HE, Haider M. Mechanical stability analysis of carrageenan-based polymer gel for magnetic resonance imaging liver phantom with lesion particles. *J Med Imaging*. 2014;1(3):035502. doi:10.1117/1.JMI.1.3.035502

44. Kemppainen R, Suilamo S, Ranta I, et al. Assessment of dosimetric and positioning accuracy of a magnetic resonance imaging-only solution for external beam radiotherapy of pelvic anatomy. July 2019. doi:10.1016/j.phro.2019.06.001

45. Choi HJ, Kim YS, Lee SH, et al. Inter- and intra-observer variability in contouring of the prostate gland on planning computed tomography and cone beam computed tomography. *Acta Oncol Stockh Swed*. 2011;50(4):539-546. doi:10.3109/0284186X.2011.562916

46. Fu J, Yang Y, Singhrao K, et al. Deep learning approaches using 2D and 3D convolutional neural networks for generating male pelvic synthetic computed tomography from magnetic resonance imaging. *Med Phys*. 2019;46(9):3788-3798. doi:10.1002/mp.13672

47. Farjam R, Tyagi N, Deasy JO, Hunt MA. Dosimetric evaluation of an atlas-based synthetic CT generation approach for MR-only radiotherapy of pelvis anatomy. *J Appl Clin Med Phys*. 2019;20(1):101-109. doi:10.1002/acm2.12501

48. Arabi H, Dowling JA, Burgos N, et al. Comparison of synthetic CT generation algorithms for MRI-only radiation planning in the pelvic region. In: *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*. ; 2018:1-3. doi:10.1109/NSSMIC.2018.8824321

49. Hsu S-H, Cao Y, Huang K, Feng M, Balter JM. Investigation of a method for generating synthetic CT models from MRI scans of the head and neck for radiation therapy. *Phys Med Biol*. 2013;58(23). doi:10.1088/0031-9155/58/23/8419

50. Yang Y, Cao M, Kaprealian T, et al. Accuracy of UTE-MRI-based patient setup for brain cancer radiation therapy. *Med Phys*. 2016;43(1):262-267. doi:10.1118/1.4938266

CHAPTER 4 - Quantification of Fiducial Marker Visibility for MRI-Only Prostate Radiation Therapy Simulation

4.1. Abstract

Purpose: To objectively compare the suitability of MRI pulse sequences and commercially available fiducial markers (FMs) for MRI-only prostate radiation therapy simulation. Most FMs appear as small signal voids in MRI images making them difficult to differentiate from tissue heterogeneities such as calcifications. In this study we use quantitative metrics to objectively evaluate the visibility of FMs in 27 patients and an anthropomorphic phantom with a variety of standard clinical MRI pulse sequences and commercially available FMs.

Methods and Materials: FM visibility was quantified using the local contrast-to-noise-ratio lCNR, the difference between the 80^{th} and 20^{th} percentile iso-intensity FM volumes (V_{fall}) and the largest iso-intensity volume that can be distinguished from background: apparent-marker-volume (AMV). A larger lCNR and AMV, and smaller V_{fall} represents a more easily identifiable FM. The number of non-marker objects visualized by each pulse sequence was calculated using FM-derived template-matching. The FM-based Target Registration Error (TRE) between each MRI and the planning-CT image was calculated. Fiducial marker visibility was rated by two medical physicists with over three years of experience examining MRI-only prostate simulation images. The rater's classification accuracy was quantified using the F_1 score, which is the harmonic mean of the rater's precision and recall. These quantitative metrics and human observer ratings were used to evaluate FM identifiability in images from nine subtypes of T_1-weighted, T_2-weighted and GRE pulse sequences in a 27-patient study. A phantom study was conducted to quantify the visibility of 8

commercially available FMs.

Results: In the patient study, the largest mean lCNR and AMV and, smallest normalized V_{fall} were produced by the 3.0T multiple echo GRE (MEGRE) pulse sequence (T_1-VIBE, 2° flip angle, 1.23 ms and 2.45 ms TEs). This pulse sequence produced no false marker detections and TREs less than 2mm in the LR, AP and SI directions, respectively. Human observers rated the 1.23 ms TE GRE images with the best average marker visibility score of 100% and an F_1 score of 1. In the phantom study, the Gold-Anchor GA-200X-20-B (deployed in a folded configuration) produced the largest sequence averaged lCNR and AMV measurements at 16.1 and 16.7 mm^3 respectively.

Conclusions: Using quantitative visibility and distinguishability metrics and human observer ratings, the patient study demonstrated that MEGRE images produced the best gold FM visibility and distinguishability. The phantom study demonstrated that markers manufactured from platinum or iron-doped gold quantitatively produced superior visibility compared to their pure gold counterparts.

4.2. Introduction

Magnetic Resonance Imaging (MRI) has been adopted in many radiation therapy clinics because the modality provides excellent visualization of soft tissues compared to CT imaging [1–7]. Some radiation therapy centers have adopted MRI-only radiation therapy simulation imaging due to the benefits of removing CT from the radiation therapy planning workflow [8–13]. These benefits include eliminating uncertainties associated with multimodal image registration, reducing the number of imaging appointments a patient must attend and decreasing radiation exposure from imaging procedures.

Fiducial markers are used by some clinics performing MRI-only prostate simulation for patient positioning and alignment [12,14–17]. Fiducial marker-based prostate cancer patient alignment for MRI-only radiation therapy simulation imaging can be challenging because fiducial markers, calcifications and tissue heterogeneities all appear as signal voids on conventional MRI [18–21]. This can make it challenging to accurately distinguish fiducial markers from other signal voids. This is a significant problem since calcifications appear in over 40% of prostate cancer patients [22]. Incorrect fiducial marker identification could potentially result in patient misalignment. To mitigate this, it is important to identify fiducial markers and MRI pulse sequences that allow distinct visualization of fiducial markers for manual or automatic identification. Correct marker identification requires a clearly visible fiducial marker-induced signal void. Additionally, markers need to be distinguishable from other intraprostatic objects that produce a similar signal voids such as calcifications.

Previous work on the identification of optimal fiducial markers and MRI pulse sequences used human observers to rate the visibility of fiducial markers in MRI images. The visibility of gold fiducial markers for image-guided radiation therapy of rectal cancer was evaluated by Ende

et al [23]. Four different commercial fiducial markers were inserted into 5 patients. Four human observers scored marker visibility in images acquired using axial and sagittal T_2-weighted TSE and T_1-weighted 3D Gradient-Echo (GRE) pulse sequences. The 0.75 mm diameter Visicoil and Gold Anchor fiducial markers were determined to be the most visible on MRI. The use of T_1-weighted GRE MRI pulse sequences was recommended for fiducial marker identification and localization. No platinum fiducials were included in this study. Gurney-Champion et al tested the visibility of fiducials with a range of different diameters (0.28 – 0.6 mm) and iron doping percentages from 0 - 0.5% [24]. These tests were performed with using T_1-weighted Fast Free Echo with Dixon reconstruction and Steady State Free Precession (SSFP) pulse sequences. The visibility of fiducial markers was scored by human observers (radiologists). The highest ranked marker was a folded marker with a 0.5% iron content. While the iron-doped markers were more visible to observers, their appearance was not compared to intraprostatic calcifications or other tissue heterogeneities. Chan et al evaluated the visibility of two solid gold markers, four gold coils and one polymer marker from three vendors using 1.5T and 3.0T scanners. The markers were implanted in a bolus phantom and visibility was scored based on a qualitative assessment of the fiducial marker images. A limitation of this study was the use of a bolus that did not resemble human tissue on MRI. Depending on the specific material, bolus can appear hyperintense compared to human tissue because of the long T_1 and T_2 relaxation properties of these materials [21]. Studies of brachytherapy seed-induced signal void visibility in MRI have been done. Wachowicz et al characterized the susceptibility artifacts produced by a typical brachytherapy seed in a clinical 1.5T MRI scanner [25]. They found that the distortion patterns varied with the seed orientation relative to the main magnetic field and the read-encode direction. Some brachytherapy clinics use T_1-weighted fast-spoiled-gradient-echo (FSPGR) imaging for seed localization [26].

However, it is yet to be seen if these findings can directly translate to the prostate MRI-only simulation imaging workflow.

In this study we quantitatively assess the visibility and distinguishability of gold fiducial markers in prostate cancer patients with a range of common clinical MRI-only pulse sequences and, assess the visibility of eight different commercial fiducial markers in a phantom.

4.3. Methods and Materials

4.3.A. Patient Image Acquisition

Twenty-seven prostate cancer patients were retrospectively selected for this study. Patients underwent CT and MRI imaging for radiation therapy planning and simulation at the UCLA between June 2013 to July 2019. Patients received transperineal implantation of three (diameter = 1.2 mm, length = 3 mm) cylindrical gold fiducial markers (Civco Medical Solutions, Coralville, IA, USA) by a physician using template-guided 17-gauge needles. Only patients who received fiducial marker implantation prior to simulation imaging were included in this study. Patients were not included in the study if they had hip prosthesis (due to prostate-obscuring artifacts) or if less than three markers appeared in their MRI or CT images, as may be the case if a marker has migrated out of the prostate. To determine if a marker was missing, prostate contour-guided rigid registration of each MRI and CT image was performed, and the images were manually inspected to check if matching marker-like structures were present in both images. The study was conducted with Institutional Review Board (IRB) approval (IRB# 19-001731).

CT scans were acquired on a Sensation Open with an 82 cm bore (Siemens, Erlangen, Germany). CT scans were acquired with the following image parameters: 120kVp, exposure time = 500 mAs, tube current = 400mA, in-plane matrix = 512 x 512 pixels and a 1.5mm slice thickness.

The median in-plane resolution was 1.2 mm x 1.2 mm. Patients were in treatment position prior to imaging with a flat table, knee wedges and Vac-Lok immobilization cushioning.

MRI scans were acquired on the same day as the CT scans. Three types of pulse sequence were tested: T_1-weighted spin-echo, T_2-weighed spin-echo and T_1-weighted gradient-echo pulse sequences (*Figure 4.3.1*). The subtypes of T_1-weighted and T_2-weighted spin echo pulse sequences tested were T_1-weighted and T_2-weighted TSE, and T_1-weighted and T_2-weighted Fast Spin Echo (FSE). The subtypes of GRE pulse sequences included T_1-weighted GRE, T_1-weighted Volume Interpolated Breath-hold Exam (VIBE), T_1-weighted VIBE with Controlled Aliasing in Parallel Imaging Results in Higher Acceleration (CAIPIRINHA) [27]. The T_1-weighted LAVA-Flex pulse sequence was included as a FSPGR pulse sequence. A MEGRE pulse sequence was included with a pre contrast T_1-weighted VIBE image acquired at two separate TEs. A summary of the MRI pulse sequences, MRI scanner manufacturers, field strengths that were included in this study are shown in *Table 4.3.I*. The shorthand names for the pulse sequences used in this study are included in *Table 4.3.I.* MRI image acquisition parameters are included in *Appendix IV*. These pulse sequences were selected because they have either been included in previous MRI-only fiducial marker localization studies or are used in MRI-only prostate simulations [15,28–30]. Patient setup at the MRI was performed using a knee wedge on a curved couch and without laser-aided positioning. Patients who had images acquired on Siemens scanners were scanned using the 18-channel anterior phased array coils (18-Channel Body coil). Patients who had images acquired on GE scanners were scanned using an 8-channel anterior phased array coils (8-Channel body coil, GE Healthcare, Chicago, Il, USA).

Pulse Seque Type	MRI pulse sequence	Manufacturer	Field Streng	Acquisition Mode	Contrast administered?	Number o Patients
T_1-weighted Spin Echo		GE		2D	N	7
		Siemens		2D	N	1
	T_1-weighted TSE (T_1-TSE)	Siemens	3.0T	2D	N	11
T_2-weighted Spin Echo		GE		2D	N	7
		Siemens		2D	N	1
	T_2-weighted TSE (T_2-TSE)	Siemens	3.0T	2D	N	11
		GE		3D	Y	7
GRE	T_1-weighted GRE (T_1-GRE)	Siemens	1.5T	3D	Y	1
		Siemens		3D	Y	6
		Siemens		3D	Y	7
	T_1-weighted VIBE with CAIPIRINHA (T_1-VIBEcaipi)	Siemens	3.0T	3D	Y	4
Multiple-Echo GRE	T_1-weighted VIBE 1st echo (T_1-VIBEe1)	Siemens	3.0T	3D	N	8
	2nd echo (T_1-	Siemens	3.0T	3D	N	8

Table 4.3.I. Summary table of the MRI pulse sequences included in this study. The number of patients included in the st MRI scanner manufacturers, sequence type and contrast injection information are included in this table. The pulse sequences included in this study will be referred to using the shorthand names denoted in parenthesis in the MRI pulse sequence column.

Figure 4.3.1. Prostate MRI images of each pulse sequence with an implanted marker circled in red. Tested pulse sequences included (a) 1.5T T_1-FSE, (b) 1.5T T_2-FSE, (c) 1.5T T_1-LAVA, (d) 3.0T T_1-TSE, (e) 3.0T T_2-TSE, (f) 3.0T T_1-GRE, (g) 3.0T T_1-VIBE, (h) 3.0T T_1-VIBEcaipi, (i) 3.0T T_1-VIBEe1 and (j) 3.0T T_1-VIBEe2.

4.3.B. Marker Visibility and Distinguishability Metrics

Fiducial marker visibility was quantified using the lCNR, the signal void intensity falloff volume, V_{fall}, and AMV. The visibility metrics were selected because they have been shown to be related to human or machine-based object detectability [31–33]. Marker identification was performed in the Eclipse treatment planning system (Eclipse version 15.6.05, Varian Medical Systems inc, Palo Alto, CA). The visibility metrics are initialized with a 0.5 x 0.5 x 0.5 cm^3 ROI around the estimated marker position. This ROI is manually defined using the aid of the marker locations and relative positions in each patient's corresponding CT image. Inside this ROI the location of the voxel with minimum value was located. A bounding box was created consisting of all voxels wholly or partially within a 1 x 1 x 1 cm^3 region surrounding this location. The lCNR was computed by $\left|{(v-\mu_p)}/{\sigma_p}\right|$ where v is the minimum voxel value, μ_p and σ_p are the mean and standard deviation of the voxels inside the bounded region (*Figure 4.3.2a*). The MRI marker iso-contour was defined as the voxels contained within the 20% iso-intensity volume relative to the minimum marker signal. The iso-intensity volume is defined as the voxels the fulfill the threshold criteria: $\mu_p - t \times (\mu_p - v)$, where t is the threshold value. The AMV is defined as the volume of all the voxels enclosed by the MRI marker iso-contour.

The V_{fall} describes how sharply the marker signal changes from its peak minimum value to background. This allows differentiation between markers with clearly defined borders and markers with diffuse boundaries (*Figure 4.3.2b*). The V_{fall} is defined by $V_{fall} = V(t_i) - V(t_j)$, where $V(t_{i,j})$ is the marker volume at a certain threshold. In this study, the 80% and 20% iso-intensity volumes were used for defining V_{fall}.

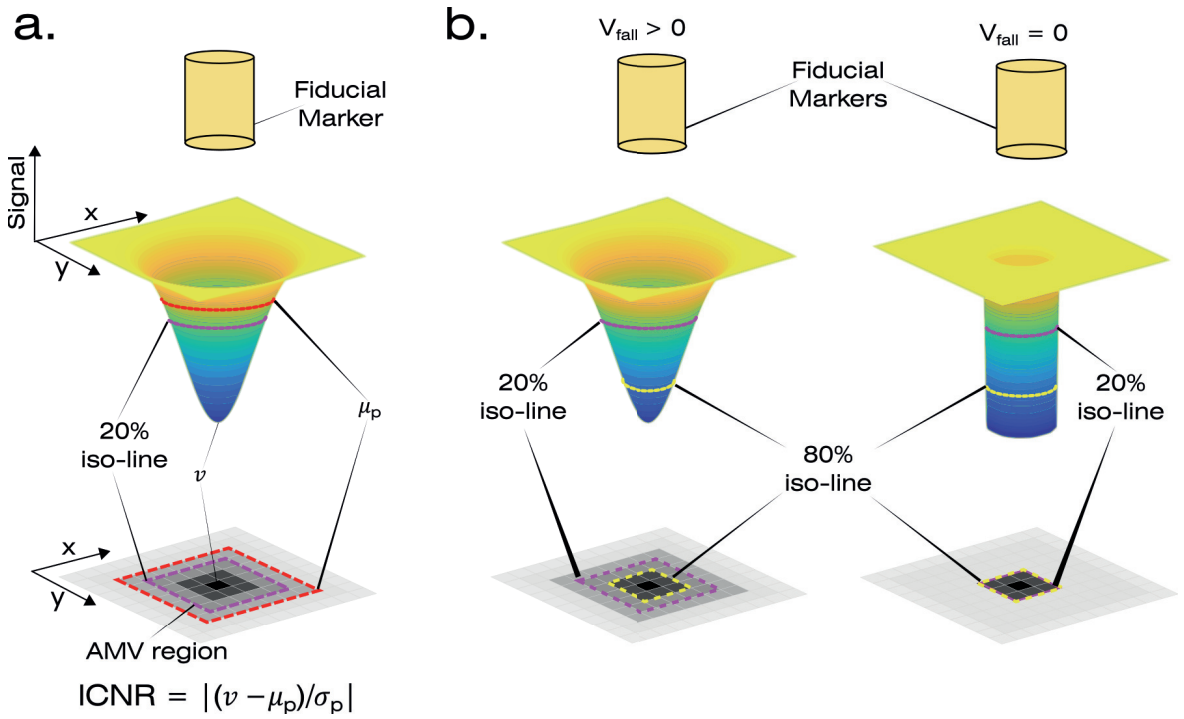


Figure 4.3.2. (a) A graphic showing the 2D and 3D representations of a fiducial marker-induced signal void. The lCNR is computed using the difference between the lowest and mean signal inside a region surrounding the marker. The AMV region encompasses the voxels that fall within the 20% iso-intensity region relative to the minimum signal voxel. In the above example the AMV would be computed using the voxels encompassed by the pink dashed line. (b) A graphic of two fiducial markers with different values of V_{fall}. V_{fall} is calculated by subtracting the volumes encompassed by the 20% and 80% iso-intensity lines.

Distinguishability metrics were calculated using marker-derived template-matching to evaluate the number of intraprostatic objects that appear similar to fiducial marker-induced signal voids. Patient- and pulse sequence-specific templates were created. Three 3D marker templates were created for each MRI image. Templates were created by using the same 1 x 1 x 1 cm^3 region bounding region for the visibility metric calculations. A patient- and image-specific normalized cross-correlation (NCC) map for each template. NCC method as used because it is a fairly standard means of quantifying similarity between images [34–36]. For each image a final NCC map was generated by taking the voxel-wise maximum of the NCC maps generated from the three marker templates. Suspect markers were identified by applying an 85% threshold to the final NCC map

and counting the number of distinct high NCC regions. A region was considered distinct if it shared no common edges with other high-NCC voxels. Voxels with a connectivity of 18 (voxels which have a shared face or edge) were clustered using the bwconncomp tool in Matlab. Suspect markers were classified as real markers, calcifications or tissue heterogeneities. Suspect markers were defined as fiducial markers if there was at least a single voxel overlap between the thresholded and clustered NCC map and the semi-automatically defined AMV. Calcifications were defined on MRI images using assistance from planning CT images. Calcifications were contoured on CT images in MIM. Calcifications were identified on MRI images by selecting signal voids which appeared in similar anatomical regions to calcifications in CT images. If one of the remaining high NCC regions had at least a single voxel overlap with a calcification contour, then it was counted as such. The remaining unaccounted high NCC regions were counted as tissue heterogeneities.

The TRE for each pulse sequence was calculated to establish the registration accuracy for fiducial marker contours generated using only MRI images. The TRE is a spatial accuracy test that depends on both the fiducial marker-induced signal void visibility and distinguishability. The TRE calculation framework used in this study was based on Fitzpatrick et al [37,38]. To calculate the TRE, an MRI fiducial marker was defined using the voxels enclosed within the semi-automatically defined AMV. The MRI images were rigidly registered to the patients planning CT image based on the centroid point locations of the MRI and CT fiducial markers in MIM (MIM Software Inc., Cleveland, OH). The CT marker centroids were defined based on a 1000 HU threshold applied to a 1 x 1 x 1 cm^3 bounding box around each marker. The images were then registered based on the marker centroids using point-based rigid registration. The target was based on the physician-drawn CT prostate contour. The TREs are reported in the LR, AP and SI directions.

Fiducial marker visibility was rated for each MRI pulse sequence by two medical physicists

with 3+ years of experience examining MRI-only prostate simulation images. Raters were given an instructional session about how fiducial markers may appear in different MRI pulse sequences. Raters were then shown masked prostate images (the same input as the visibility metrics calculations) in MIM and were instructed to mark signal voids that may resemble markers. Raters were allowed to view images in axial, sagittal or coronal planes, control the window/level and zoom to aid marker detection. Raters were requested to place a single point contour in the signal voids believed to be markers. No limit was set to the number of suspect markers that the raters could select. Raters were asked to rank their marker selection confidence on a scale of 1 (not very confident) to 3 (very confident) as outlined by Ende et al [38]. Raters were asked to rank marker visibility on a scale of 1 (markers not visible) to 5 (markers clearly visible). The classification accuracy was quantified using the F_1 score for each rater. The F_1 score is a measure of the accuracy of binary classification test [39]. The F_1 score defined by, $F_1 = 2 \cdot \frac{p \times r}{p \cdot r}$, where p is classification precision and r recall. The classification precision is defined by, $p = \frac{tp}{tp+fp}$, and recall is defined by, $r = \frac{tp}{tp+fn}$, where tp, fn and fp are the number of true positive (TP), false negative (FN) and false positive (FP) counts. An F_1 score of 1 denotes perfect precision and recall.

4.3.C. Commercial Marker Comparison

A comparison of commercially available markers was performed in a physical anthropomorphic phantom (*Figure 4.3.3*). The phantom was designed based on a patient's anatomy and 3D-printed using a MakerGear M2 3D printer (MakerGear LLC, Beachwood, OH). The phantom was printed with 3 cm thickness based on a 2D axial slice of a patient's anatomy. Muscle and prostate materials were a $GdCl_3$, agarose and carrageenan solution developed by Hattori et al [40]. Porcine fat was used to mimic fat and a gelatin-$CaCO_3$ mix was used to mimic bone. Eight fiducial markers from three

different manufacturers were compared. The marker name and corresponding manufacturers are shown in *Table 4.3.II*. A sample 0.91 mm diameter calcification, consisting of $CaCO_3$, was also implanted using a 19G needle. MRI images were acquired on a Siemens Skyra 3T and CT images were acquired on a Siemens Somatom. T_1-TSE, T_2-TSE, T_1-GRE and T_1-VIBE MRI images were acquired. The lCNR and AMV for each marker was calculated. Since V_{fall} is sensitive to marker orientation and size, it is not reported for commercial marker comparison.

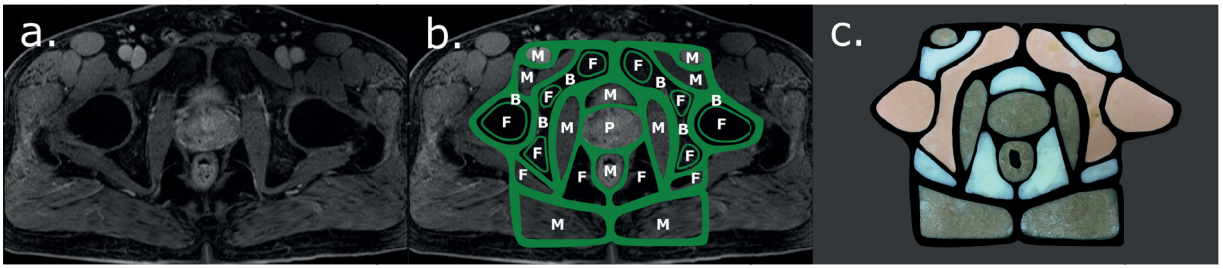


Figure 4.3.3. (a) T_1-weighted GRE MRI image that served as the template for the phantom design. (b) Fat (F), other soft tissues (M) and bone (B) were segmented. The surrounding structure was 3D-printed leaving open gaps for filler tissue material. (c) The final phantom was filled with materials for prostate, muscle, fat and bone.

Manufacturer	Marker Type	Needle Gauge (AWG)	Marker diameter (mm)	Marker length (mm)
Civco	Gold soft tissue	18	1.2	3
Civco	PointCoil	20	0.6	5
Civco	Polymark	18	1	3
Naslund	Gold Anchor GA-200X-20-B	20	0.40	20
Naslund	Gold Anchor GA-150-20	25	0.28	20
IBA	Visicoil Platinum 1	22	0.35	10
IBA	Visicoil Platinum 2	21	0.5	10
IBA	Visicoil Gold	21	0.5	10

Table 4.3.II. A list of tested markers, needle gauges and dimensions. Precision as reported by manufacturers.

4.4. Results

4.4.A. Retrospective Patient Study

Images were grouped according to MRI pulse sequence. Box plots and a summary table of the visibility metrics are shown in *Figure 4.4.1* and *Table 4.4.I.* The reported V_{fall} results were normalized by the AMV of each fiducial marker. The minimum mean and maximum number of intraprostatic marker-like objects are reported in *Table 4.4.II*. No incorrect markers were identified on the 1.5T T_1-TSE, 1.5T T_2-TSE and 1.5T T_1-GRE pulse sequences. No incorrect markers were identified on 3.0T T_1-VIBEcaipi and 3.0T T_1-GRE images. No FN counts were counted in the object distinguishability search. The TRE results for all tested pulse sequences are reported in *Table 4.4.III.* The human observer ratings are presented in *Table 4.4.IV*.

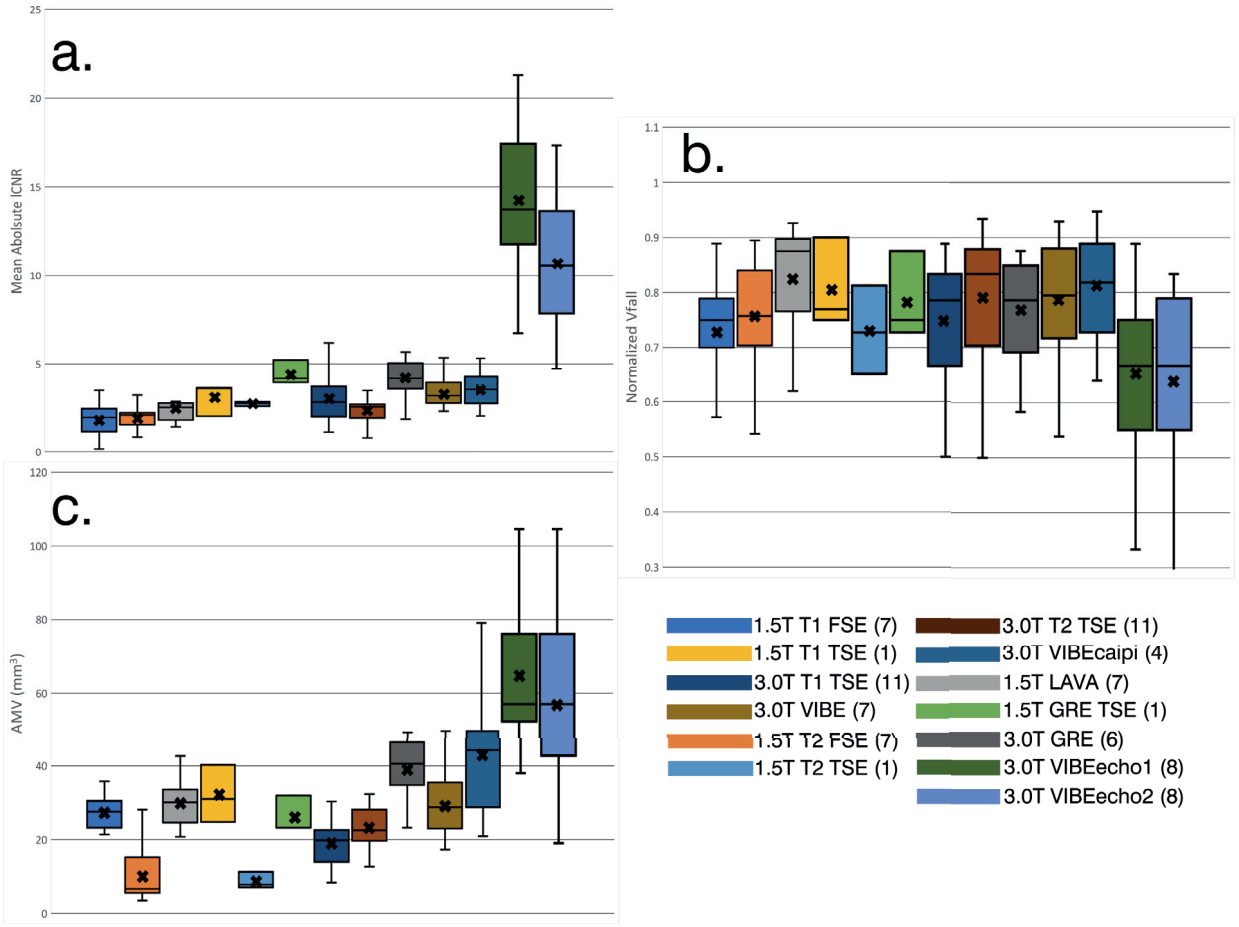


Figure 4.4.1. (a) Box plots of mean absolute marker lCNR for each pulse sequence in the patient study. Higher absolute lCNRs indicate markers with increased visibility. (b) Box plots of mean V_{fall} for each pulse sequence in the patient study. Lower V_{fall} values indicate markers with sharp edges. (c) Box plots of mean AMV for each pulse sequence in the patient study. A higher AMV indicates a larger marker size. A larger marker size indicates a marker with a larger apparent size. The x denotes the sample mean. The lines on each box represent (from top to bottom) the maximum, third quartile, median, first quartile and minimum values of each sample. The number of patients from each group are denoted in parenthesis.

Visibility metrics: Mean lCNR, Normalized V_{fall} and AMV results from patient study.					
Field Strength	**Number of patients**	**Pulse Sequence**	**lCNR**	**Normalized V_{fall}**	**AMV (mm^3)**
1.5T	7	T_1 FSE	1.8 ± 0.9	0.7 ± 0.1	27.2 ± 4.4
1.5T	7	T_2 FSE	1.9 ± 0.6	0.7 ± 0.2	10.0 ± 7.2
1.5T	7	LAVA/FLEX	2.5 ± 0.8	0.8 ± 0.1	29.8 ± 6.0
1.5T	1	T_1 TSE	3.1 ± 0.9	0.8 ± 0.1	32.0 ± 7.8
1.5T	1	T_2 TSE	2.7 ± 0.1	0.7 ± 0.1	8.7 ± 2.3
1.5T	1	GRE	4.4 ± 0.7	0.8 ± 0.1	26.1 ± 5.0
3.0T	11	T_1 TSE	3.0 ± 1.5	0.8 ± 0.1	19.2 ± 6.2
3.0T	11	T_2 TSE	2.3 ± 0.6	0.8 ± 0.1	23.1 ± 5.3
3.0T	6	GRE	4.2 ± 1.4	0.8 ± 0.1	39.0 ± 7.9
3.0T	7	VIBE	3.2 ± 1.0	0.8 ± 0.1	29.1 ± 8.1
3.0T	4	VIBEcaipi	3.6 ± 1.0	0.8 ± 0.1	43.3 ± 19.0
3.0T	8	VIBEe1	10.7 ± 3.6	0.7 ± 0.1	65.2 ± 19.6
3.0T	8	VIBEe2	10.7 ± 3.6	0.6 ± 0.2	57.0 ± 20.8
Global			4.6 ± 4.3	0.8 ± 0.1	32.2 ± 19.5

Table 4.4.I. Mean and standard deviation results for the visibility metrics lCNR, V_{fall} and AMV for each tested pulse sequence separated by field strength.

Distinguishability metrics: Number of false positive marker-like objects detected in the patient study										
Field Strength	1.5 T			3.0T						
Pulse Sequence	T_1-FSE	T_2-FSE	T_1-LAVA	T_1-TSE	T_2-TSE	T_1-GRE	T_1-VIBE	T_1-VIBEcaipi	T_1-VIBEe1	T_1-VIBEe2
Number of Patients	7	7	7	11	11	6	7	4	8	8
Min	0	0	0	0	0	0	0	0	0	0
Mean	8.1	11.3	2.7	1.8	0.5	0	0.4	0	0	0
Max	23	36	8	11	6	0	3	0	0	0

Table 4.4.II. The minimum, mean and maximum number of FP marker-like objects detected for each MRI pulse sequence.

Field Strength	Number of patients	Pulse Sequence	LR direction error (mm)	AP direction error (mm)	SI direction error (mm)
1.5 T	7	T_1-TSE	0.3	2.8	2.1
1.5 T	7	T_2-TSE	1.5	1.0	0.8
1.5 T	7	T_1-GRE	0.9	0.5	1.2
1.5 T	1	T_1-FSE	1.8 ± 5.4	1.1 ± 0.7	0.9 ± 0.2
1.5 T	1	T_2-FSE	1.4 ± 0.7	1.8 ± 4.7	1.7 ± 1.2
1.5 T	1	T_1-LAVA	1.1 ± 0.7	1.5 ± 1.6	1.6 ± 2.1
3.0 T	11	T_1-TSE	1.0 ± 0.3	1.0 ± 0.7	1.4 ± 0.6
3.0 T	11	T_2-TSE	1.4 ± 2.4	0.9 ± 0.8	1.8 ± 3.3
3.0 T	6	T_1-GRE	0.8 ± 0.3	0.9 ± 0.2	1.3 ± 0.6
3.0 T	7	T_1-VIBE	1.4 ± 0.2	1.2 ± 0.3	1.0 ± 1.1
3.0 T	4	T_1-VIBEcaipi	0.5 ± 0.1	0.6 ± 0.3	0.9 ± 0.6
3.0 T	8	T_1-VIBEe1	1.6 ± 1.1	1.6 ± 0.9	1.1 ± 0.8
3.0 T	8	T_1-VIBEe2	1.7 ± 1.3	2.0 ± 1.1	1.5 ± 0.9
Global			1.28 ± 1.09	1.27 ± 1.05	1.35 ± 0.99

Table 4.4.III. TRE Mean and standard deviation results in the LR, AP and CC directions for each tested pulse sequence separated by field strength.

Field		of	F_1	detection	Mean marker visibility (%)	F_1	detection	Mean marker visibility (%)
1.5T	T_1-FSE	7	0.24	48	37	0.25	52	43
1.5T	T_1-GRE	1	1	100	100	1	67	60
3.0T	T_1-VIBE	7	0.86	76	71	0.86	57	49
				83	70			70

Table 4.4.IV. Summary table of human observer visibility ratings. The F_1 score is reported as a metric for fiducial marker classification accuracy in each of the tested MRI pulse sequences. An F_1 score of 1 denotes perfect precision and recall and a score of 0 denotes inaccurate precision and recall. The mean detection confidence and the marker visibility scores are presented for both raters.

4.4.B. MRI Visibility of Commercial Fiducial Markers

Figure 4.4.2 shows the markers implanted in the phantom for CT and each tested MRI pulse sequence. The Naslund Gold Anchor markers were able to be deployed as a line or a bundle. For this study we deployed the Naslund markers as a folded bundle. *Table 4.4.V* shows the lCNR and AMV for each fiducial marker and each MRI pulse sequence.

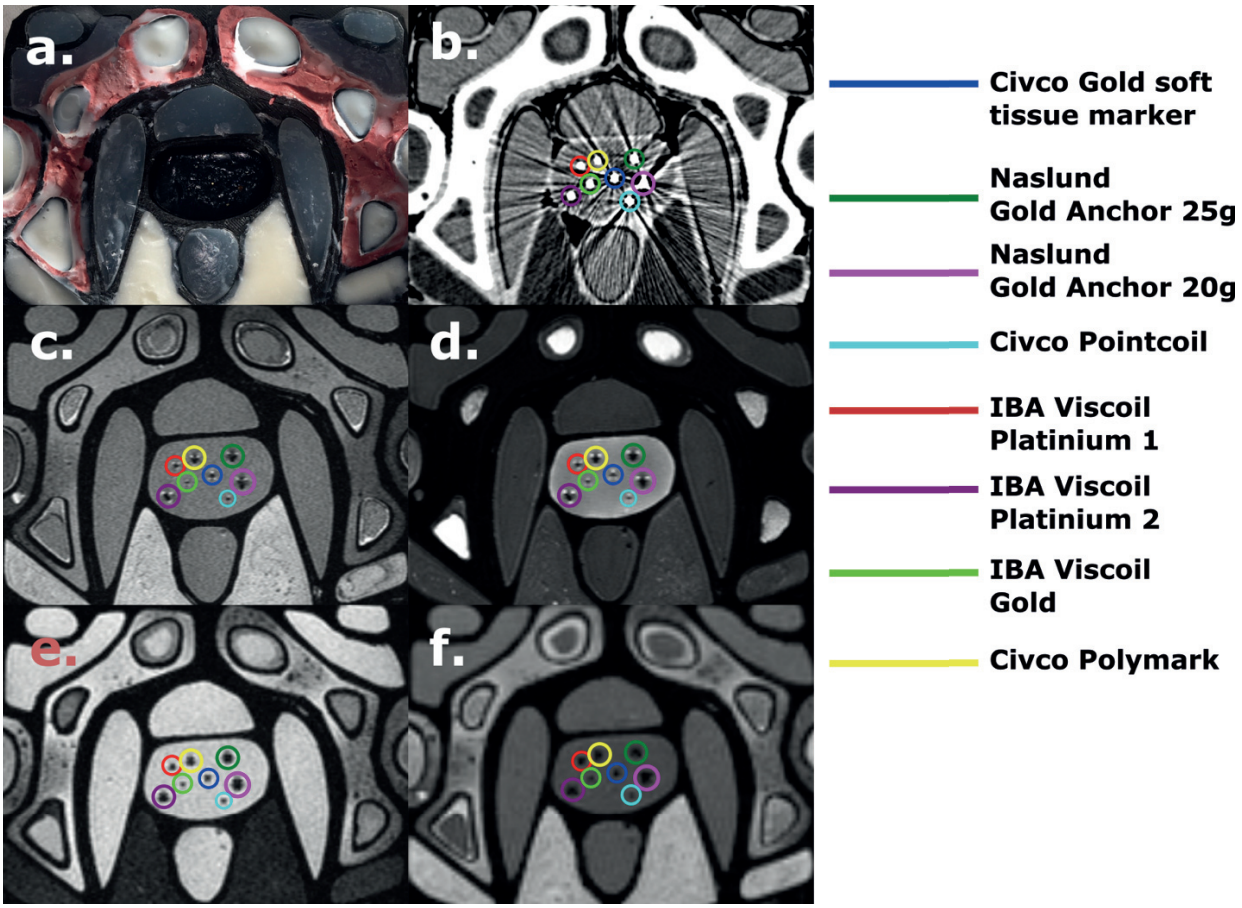


Figure 4.4.2. (a) a photo of the anthropomorphic male pelvis phantom, (b) CT, (c) T_1-TSE, (d) T_2-TSE, (e) T_1-GRE and (f) T_1-VIBE images of phantom.

	Pulse Sequence	IBA Visicoil Platinum 2	IBA ïsicoil Platinum 1	IBA Gold	Civco Polymark	Gold soft tissue marker	Gold Anchor GA-150-20	Civco Pointcoil	ıncho GA-200X-20-B	Calcification
lCNR										
	T_1-GRE	18.3	19.5		18.1	18.6	19.9	17.9	20.4	
AMV (mm^3)	T_1-TSE	1.8	3.0		4.8	3.0	4.2	3.6	4.8	
	T_1-VIBE	26.4	15.8		26.4	10.6	15.8	10.6	31.7	

Table 4.4.V. Comparison of AMV and lCNR values for the tested markers and simulated calcification in the phantom study.

4.5. Discussion

This study used quantitative metrics to evaluate the visibility of fiducial marker-induced signal voids in MRI images. The patient study involved testing the visibility and distinguishability of gold fiducial markers across a variety of commonly used sequences in MRI-only radiation therapy simulation. The quantitative metrics used to determine visibility and distinguishability were lCNR, V_{fall}, AMV and TRE. Human observers also rated the visibility of the markers in images from each pulse sequence. The phantom study sought to determine which commercially available fiducial marker-induced signal void displayed the best visibility across a variety of commonly used MRI pulse sequences in phantom.

The T_1-weighted GRE and VIBE pulse sequences (with the exception of the FSPGR T_1-LAVA sequence) produced the best visibility and distinguishability scores. Specifically, the 3.0T multiple-echo T_1-VIBEe1 and T_1-VIBEe2 sequences produced the highest mean lCNR and AMV scores and the smallest V_{fall} scores compared to the other tested MRI pulse sequences. No FPs were detected with this pulse sequence. The TREs for this pulse sequence were less than 2 mm in each plane. A 2 mm setup uncertainty is a commonly accepted error for fiducial marker-based prostate alignment [41]. The 3.0T multiple-echo T_1-VIBEe1 and T_1-VIBEe2 were the only pulse sequences which had a fiducial marker-induced signal void F_1-classification score greater than 0.96 and a mean marker visibility score greater than 95% for both human observers. These results are consistent with findings from other publications about the visibility of fiducial marker-induced signal voids. Ende et al qualitatively assessed the MRI visibility of gold fiducial markers for image-guided radiation therapy of rectal cancer [23]. They reported that a 3.0T T_1-weighted 3D GRE images produced superior marker visualization compared to sagittal and axially acquired T_2-weighted Spin-Echo images. The findings by Ende et al are consistent with this study where the human-

observer and quantitative visibility and distinguishability metrics of multiple-echo VIBE images produced superior marker visualization compared to GRE images acquired with longer TEs. This is also consistent with observations by Schieda et al who reported that the use of MEGRE images provides better detection of implanted seeds and fiducial markers when compared with both FSE and single-echo GRE [42]. MEGRE sequences are used for fiducial marker identification at some sites performing MRI-only radiation therapy simulation and the quantitative and qualitative results of this study support the use of this sequence for fiducial marker identification [15,17,28].

The phantom study demonstrated that the Gold Anchor GA-200X-20-B produced the largest lCNR and AMV with the GRE and T_1 VIBE sequences at 20 and 32 mm^3 respectively. The markers manufactured from platinum or iron-doped gold such as the platinum IBA Visicoil and Gold Anchors generally produced greater lCNR and AMV measurements than their pure gold counterparts. This is consistent with the findings by Gurney-Champion et al who reported that iron doped markers appeared more visible to human observers [24]. Ende et al tested the qualitative visibility an array of fiducial markers including a 0.5 mm diameter Visicoil, 0.75 mm Visicoil and Gold Anchor [23]. They reported that the Gold Anchor (deployed in its folded configuration) and 0.75 mm diameter Visicoil produced the best visibility. They attributed the visibility of the Gold Anchor to the iron-doping. They also attributed to the higher visibility scores of the 0.75 mm diameter over the 0.5 mm diameter Visicoil markers to the larger diameter of the marker producing a more pronounced signal void. Our study only tested the 0.25 mm and 0.5 mm diameter Visicoil fiducial markers. We also observed that for the 3D T_1-GRE and T_1-VIBE sequences that the larger diameter markers produced higher lCNR and AMV values. This study did not test the difference in visibility of the configuration of the various configurations of Gold Anchor fiducial markers and the Gold Anchors were deployed in a folded configuration.

There are pulse sequence and image acquisition factors that can affect the distinguishability of fiducial marker-induced signal voids. The distinguishability of signal voids and susceptibility artifacts will vary depending if GRE or spin-echo sequences are used. GRE imaging relies on the use of dephasing-rephasing gradient reversals resulting in image contrast being determined by T_2* relaxation [43]. This results in susceptibility artifacts being more pronounced in GRE images compared to spin-echo images [44]. Additionally, the size of fiducial marker-induced signal voids is dependent on factors including pulse sequence TE, the number of frequency encoding steps and slice thickness [42,45]. It is important to select the correct TE settings to allow differentiation between signal voids created by fiducial markers and background tissue (e.g. calcifications) [28,42].

Platinum has been proposed as a fiducial marker material. It has a higher magnetic susceptibility than gold and may create a larger artifact that is easier to identify. In this study two IBA Visicoil markers with identical dimensions, one manufactured from platinum and the other from gold, were compared. The platinum Visicoil marker had a larger AMV for T_2 TSE, GRE, and T_1 VIBE images. The gold marker had a larger AMV in the T_1 TSE images. The largest AMV increase between platinum and gold occurred in the T_1 VIBE images, where the AMV increased by 21.1 mm^3. However, the platinum marker exhibited a lower lCNR for all tested sequences.

Local distortions caused by implanted fiducials could potentially compromise localization accuracy. Fiducial markers do produce local distortions and the size of distortions depend on a number of factors including marker materials and sequence type. Fernandes et al observed that larger pulse sequence bandwidths decreased that magnitude of local distortions [30]. Markers manufactured from platinum or doped with iron also affect the extent of local distortions because of the effects that these elements have on changing the markers magnetic susceptibility. Nair et al observed that platinum markers appeared to produce more contrast than gold because of its high

magnetic susceptibility [46]. However, the effect on registration from susceptibility effects from fiducial markers has been shown to be small. Jonsson et al evaluated the susceptibility effects of gold fiducial markers in different geometries for prostate MRI-only simulations for 2D and 3D spin echo sequences. They found that the susceptibility effects resulted in sub mm fiducial marker positional differences [47].

There are limitations to this study. The patient study was retrospective, and the imaging parameters varied some between subjects (Supplemental Materials). The number of patients in each MRI sequence group varied from 1 (1.5T T_1-TSE, T_2-TSE and GRE) to 11 (3.0T T_1-TSE, T_2-TSE). While there is agreement between the human observer and quantitative visibility metrics, further studies with larger cohorts of patients are needed to validate the results in this study. Marker migration between MRI and CT simulation exams can potentially misidentification of the MRI signal voids. Studies have shown that the effect of marker migration is typically small such as Arpacı et al who evaluated marker migration between implantation and CT simulation for 39 patients [48]. The mean distance of migration was 1.0 ± 0.4 mm and the maximum distance was 1.4 ± 0.6 mm. The phantom study could be improved by testing these fiducial markers in patients. It is impractical to implant many markers inside patients; however, this study could help guide which markers warrant further patient testing. While only objects inside the prostate boundary were counted, the 1 cm^3 template size means that for locations near the prostate boundary, some tissue outside of the prostate was included in the NCC calculation.

4.6. Conclusions

This study used quantitative visibility metrics (lCNR, AMV, V_{fall}), distinguishability metrics (calculated using marker-derived template matching), TRE measurements and human observer

ratings to determine which fiducial marker and MRI sequence produced the best visibility of fiducial marker-included signal voids. The patient study demonstrated that multiple-echo T_1-VIBE sequences produced the best visibility of clinically used gold fiducial markers. The phantom study demonstrated that markers manufactured from platinum or iron-doped gold quantitatively produced superior visibility compared to their pure gold counterparts.

4.7. Appendix IV

of		manufacturer and model	field strength (T)		TR (ms)	Flip angle (°)	Axial resolution (mm)	ss	suppression on?	Contrast ?
7	T_1	GE HD Signa XT	1.5		483-717	90	0.74-0.82	5.0	N	N
7	T_2	GE HD Signa XT	1.5	91-	ɔvo 3-348 3	90	0.31-0.63	3.0	N	N
7	A/F	GE HD	1.5	-	3.5-	12	0.70-		Y	Y
1		Siemens Avanto	1.5		761	150	ɔ.80		N	N
2										
7		Siemens Skyra	3.0		600-800	130	ˈ.70-ɔ.80	5.0	N	N
2			3.0			130	0.70	5.0	N	N
1		Siemens Avanto	1.5		460 0	150	0.31	3.6	N	N
2			3.0			150	0.63	3.6	N	N

Appendix IV. Machine information and image acquisition parameters for all patients used in study.

umbe r of tients	nce	MR manufa er and mod	MRI field streng h (T)	TE (ms)		Flip angle (°)	Axial resoluti on (mm)	Slice thickne ss (mm)	Fat ppression on?	Contras t?
7		Siemens Skyra	3.0	101-105	-	160	0.63	3.6	N	N
2		Siemens	3.0		-	160	0.63	3.6	N	N
1		Sieme Avan	1.5	2.15		61	0.94	3.3	Y	Y
									Y	
4		Sieme Skyr	3.0	1.38-1.39		10	1.0	2.5	Y	Y
									Y	
2	VIBE CAIPI	Siemens Skyra	3.0	1.38-2.43	3.5-	9-10s	1.0-1.1	2.5	Y	Y
2	VIBE CAIPI	Siemens PrismaFit	3.0	1.38		10	1.0	2.5	Y	Y
8	VIBE	Siemens Skyr	3.0	1.23		2	1.6	3.6	Y	N
8	VIBE	Siemens Skyra	3.0			2	1.6	3.6	Y	N

Appendix IV (cont). Machine information and image acquisition parameters for all patients used in study.

4.8. Bibliography

1. Rasch C, Barillot I, Remeijer P, Touw A, van Herk M, Lebesque JV. Definition of the prostate in CT and MRI: a multi-observer study. *Int J Radiat Oncol Biol Phys*. 1999;43(1):57-66.

2. Roach M, Faillace-Akazawa P, Malfatti C, Holland J, Hricak H. Prostate volumes defined by magnetic resonance imaging and computerized tomographic scans for three-dimensional conformal radiotherapy. *Int J Radiat Oncol*. 1996;35(5):1011-1018. doi:10.1016/0360-3016(96)00232-5

3. McLaughlin PW, Evans C, Feng M, Narayana V. Radiographic and anatomic basis for prostate contouring errors and methods to improve prostate contouring accuracy. *Int J Radiat Oncol Biol Phys*. 2010;76(2):369-378. doi:10.1016/j.ijrobp.2009.02.019

4. Schmidt MA, Payne GS. Radiotherapy planning using MRI. *Phys Med Biol*. 2015;60(22):R323. doi:10.1088/0031-9155/60/22/R323

5. Lee YK, Bollet M, Charles-Edwards G, et al. Radiotherapy treatment planning of prostate cancer using magnetic resonance imaging alone. *Radiother Oncol J Eur Soc Ther Radiol Oncol*. 2003;66(2):203-216.

6. Metcalfe P, Liney GP, Holloway L, et al. The Potential for an Enhanced Role for MRI in Radiation-therapy Treatment Planning. *Technol Cancer Res Treat*. 2013;12(5):429-446. doi:10.7785/tcrt.2012.500342

7. McLaughlin PW, Narayana V, Meirovitz A, et al. Vessel-sparing prostate radiotherapy: dose limitation to critical erectile vascular structures (internal pudendal artery and corpus cavernosum) defined by MRI. *Int J Radiat Oncol Biol Phys*. 2005;61(1):20-31. doi:10.1016/j.ijrobp.2004.04.070

8. Nakazawa T, Tateoka K, Saito Y, et al. Analysis of Prostate Deformation during a Course of Radiation Therapy for Prostate Cancer. *PLOS ONE*. 2015;10(6):e0131822. doi:10.1371/journal.pone.0131822

9. Kagawa K, Lee WR, Schultheiss TE, Hunt MA, Shaer AH, Hanks GE. Initial clinical assessment of CT-MRI image fusion software in localization of the prostate for 3D conformal radiation therapy. *Int J Radiat Oncol Biol Phys*. 1997;38(2):319-325.

10. Christiansen RL, Jensen HR, Brink C. Magnetic resonance only workflow and validation of dose calculations for radiotherapy of prostate cancer. *Acta Oncol Stockh Swed*. 2017;56(6):787-791. doi:10.1080/0284186X.2017.1290275

11. Balter JM, Lam KL, Sandler HM, Littles JF, Bree RL, Ten Haken RK. Automated localization of the prostate at the time of treatment using implanted radiopaque markers: Technical feasibility. *Int J Radiat Oncol*. 1995;33(5):1281-1286. doi:10.1016/0360-3016(95)02083-7

12. Jonsson JH, Karlsson MG, Karlsson M, Nyholm T. Treatment planning using MRI data: an analysis of the dose calculation accuracy for different treatment regions. *Radiat Oncol*. 2010;5:62. doi:10.1186/1748-717X-5-62

13. Hanvey S, Sadozye AH, McJury M, Glegg M, Foster J. The influence of MRI scan position on image registration accuracy, target delineation and calculated dose in prostatic radiotherapy. *Br J Radiol*. 2012;85(1020):e1256-e1262. doi:10.1259/bjr/26802977

14. Doemer A, Chetty IJ, Glide-Hurst C, et al. Evaluating organ delineation, dose calculation and daily localization in an open-MRI simulation workflow for prostate cancer patients. *Radiat Oncol*. 2015;10:37. doi:10.1186/s13014-014-0309-0

15. Persson E, Gustafsson C, Nordström F, et al. MR-OPERA: A Multicenter/Multivendor Validation of Magnetic Resonance Imaging-Only Prostate Treatment Planning Using Synthetic Computed Tomography Images. *Int J Radiat Oncol Biol Phys*. 2017;99(3):692-700. doi:10.1016/j.ijrobp.2017.06.006

16. Tyagi N, Fontenla S, Zelefsky M, et al. Clinical workflow for MR-only simulation and planning in prostate. *Radiat Oncol*. 2017;12:119. doi:10.1186/s13014-017-0854-4

17. Maspero M, Berg CAT van den, Zijlstra F, et al. Evaluation of an automatic MR-based gold fiducial marker localisation method for MR-only prostate radiotherapy. *Phys Med Biol*. 2017;62(20):7981. doi:10.1088/1361-6560/aa875f

18. Zeng GG, McGowan TS, Larsen TM, et al. Calcifications are potential surrogates for prostate localization in image-guided radiotherapy. *Int J Radiat Oncol Biol Phys*. 2008;72(4):963-966. doi:10.1016/j.ijrobp.2008.07.021

19. O'Neill AGM, Jain S, Hounsell AR, O'Sullivan JM. Fiducial marker guided prostate radiotherapy: a review. *Br J Radiol*. 89(1068). doi:10.1259/bjr.20160296

20. Stucchi P, Pozzi L, Bianchi C, et al. Intra- or periprostatic calcifications as natural fiducial markers for prostate localization in image-guided radiotherapy. *Phys Medica Eur J Med Phys*. 2016;32:65. doi:10.1016/j.ejmp.2016.01.223

21. Chan MF, Cohen GN, Deasy JO. Qualitative Evaluation of Fiducial Markers for Radiotherapy Imaging. *Technol Cancer Res Treat*. 2015;14(3):298-304. doi:10.1177/1533034614547447

22. Hong CG, Yoon BI, Choe H-S, Ha U-S, Sohn DW, Cho Y-H. The Prevalence and Characteristic Differences in Prostatic Calcification between Health Promotion Center and Urology Department Outpatients. *Korean J Urol*. 2012;53(5):330-334. doi:10.4111/kju.2012.53.5.330

23. Ende RPJ van den, Rigter LS, Kerkhof EM, et al. MRI visibility of gold fiducial markers for image-guided radiotherapy of rectal cancer. *Radiother Oncol*. 2019;132:93-99. doi:10.1016/j.radonc.2018.11.016

24. Gurney-Champion OJ, Lens E, van der Horst A, et al. Visibility and artifacts of gold fiducial markers used for image guided radiation therapy of pancreatic cancer on MRI. *Med Phys*. 2015;42(5):2638-2647. doi:10.1118/1.4918753

25. Wachowicz K, Thomas SD, Fallone BG. Characterization of the susceptibility artifact around a prostate brachytherapy seed in MRI. *Med Phys*. 2006;33(12):4459-4467. doi:10.1118/1.2364052

26. Blanchard P, Ménard C, Frank SJ. Clinical Use of Magnetic Resonance Imaging Across the Prostate Brachytherapy Workflow. *Brachytherapy*. 2017;16(4):734-742. doi:10.1016/j.brachy.2016.11.012

27. Breuer FA, Blaimer M, Heidemann RM, Mueller MF, Griswold MA, Jakob PM. Controlled aliasing in parallel imaging results in higher acceleration (CAIPIRINHA) for multi-slice imaging. *Magn Reson Med*. 2005;53(3):684-691. doi:10.1002/mrm.20401

28. Gustafsson Christian, Korhonen Juha, Persson Emilia, Gunnlaugsson Adalsteinn, Nyholm Tufve, Olsson Lars E. Registration free automatic identification of gold fiducial markers in MRI target delineation images for prostate radiotherapy. *Med Phys*. 2017;44(11):5563-5574. doi:10.1002/mp.12516

29. Ghose S, Mitra J, Rivest-Hénault D, et al. MRI-alone radiation therapy planning for prostate cancer: Automatic fiducial marker detection. *Med Phys*. 2016;43(5):2218. doi:10.1118/1.4944871

30. Dinis Fernandes C, Dinh CV, Steggerda MJ, et al. Prostate fiducial marker detection with the use of multi-parametric magnetic resonance imaging. *Phys Imaging Radiat Oncol*. 2017;1:14-20. doi:10.1016/j.phro.2017.02.001

31. Brunelli R, Poggio T. Face recognition: features versus templates. *IEEE Trans Pattern Anal Mach Intell*. 1993;15(10):1042-1052. doi:10.1109/34.254061

32. Tjan BS, Braje WL, Legge GE, Kersten D. Human efficiency for recognizing 3-D objects in luminance noise. *Vision Res*. 1995;35(21):3053-3069.

33. Dodge S, Karam L. Understanding How Image Quality Affects Deep Neural Networks. *ArXiv160404004 Cs*. April 2016. http://arxiv.org/abs/1604.04004. Accessed January 17, 2019.

34. Yoo J-C, Han TH. Fast Normalized Cross-Correlation. *Circuits Syst Signal Process*. 2009;28(6):819. doi:10.1007/s00034-009-9130-7

35. Stefano LD, Mattoccia S, Mola M. An efficient algorithm for exhaustive template matching based on normalized cross correlation. In: *12th International Conference on Image Analysis and Processing, 2003.Proceedings.* ; 2003:322-327. doi:10.1109/ICIAP.2003.1234070

36. Sarraf S, Saverino C, Golestani AM. A robust and adaptive decision-making algorithm for detecting brain networks using functional MRI within the spatial and frequency domain. In:

2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI). ; 2016:53-56. doi:10.1109/BHI.2016.7455833

37. Fitzpatrick JM, West JB, Maurer CR. Predicting error in rigid-body point-based registration. *IEEE Trans Med Imaging*. 1998;17(5):694-702. doi:10.1109/42.736021

38. Danilchenko A, Fitzpatrick JM. General approach to error prediction in point registration. In: San Diego, California, USA; 2010:76250F. doi:10.1117/12.843847

39. Saito T, Rehmsmeier M. The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLoS ONE*. 2015;10(3). doi:10.1371/journal.pone.0118432

40. Hattori K, Ikemoto Y, Takao W, et al. Development of MRI phantom equivalent to human tissues for 3.0-T MRI. *Med Phys*. 2013;40(3):032303. doi:10.1118/1.4790023

41. Enmark M, Korreman S, Nyström H. IGRT of prostate cancer; is the margin reduction gained from daily IG time-dependent? *Acta Oncol*. 2006;45(7):907-914. doi:10.1080/02841860600907337

42. Schieda N, Avruch L, Shabana WM, Malone SC. Multi-echo gradient recalled echo imaging of the pelvis for improved depiction of brachytherapy seeds and fiducial markers facilitating radiotherapy planning and treatment of prostatic carcinoma. *J Magn Reson Imaging*. 2015;41(3):715-720. doi:10.1002/jmri.24590

43. Chavhan GB, Babyn PS, Thomas B, Shroff MM, Haacke EM. Principles, Techniques, and Applications of T2*-based MR Imaging and Its Special Applications. *Radiographics*. 2009;29(5):1433-1449. doi:10.1148/rg.295095034

44. Zand KR, Reinhold C, Haider MA, Nakai A, Rohoman L, Maheshwari S. Artifacts and pitfalls in MR imaging of the pelvis. *J Magn Reson Imaging JMRI*. 2007;26(3):480-497. doi:10.1002/jmri.20996

45. Port JD, Pomper MG. Quantification and minimization of magnetic susceptibility artifacts on GRE images. *J Comput Assist Tomogr*. 2000;24(6):958-964. doi:10.1097/00004728-200011000-00024

46. Janardanan Nair V, Szanto J, Vandervoort E, et al. Feasibility, Detectability, and Experience With Platinum Seed Internal Fiducial Markers for CT-MRI Fusion and Real-time Tumor Tracking During Stereotactic Ablative Radiation Therapy. *Int J Radiat Oncol*. 2012;84(3, Supplement):S821. doi:10.1016/j.ijrobp.2012.07.2200

47. Jonsson JH, Garpebring A, Karlsson MG, Nyholm T. Internal Fiducial Markers and Susceptibility Effects in MRI—Simulation and Measurement of Spatial Accuracy. *Int J Radiat Oncol*. 2012;82(5):1612-1618. doi:10.1016/j.ijrobp.2011.01.046

48. Arpacı T, Uğurluer G, İspir EB, Eken A, Akbaş T, Serin M. Computed tomography based evaluation of prostatic fiducial marker migration between the periods of insertion and simulation. *Turk J Urol*. 2017;43(4):451-455. doi:10.5152/tud.2017.29726

CHAPTER 5 - A Generative Adversarial Network-Based (GAN-Based) Architecture for Automatic Fiducial Marker Detection in Prostate MRI-Only Radiation Therapy Images

5.1. Abstract

Purpose: Clinical sites utilizing MRI-only simulation imaging for prostate radiotherapy planning typically use fiducial markers for pretreatment patient positioning and alignment. Fiducial markers appear as small signal voids in MRI images and are often difficult to discern. Existing clinical methods for fiducial marker localization require multiple MRI sequences and/or manual interaction and specialized expertise. In this study, we develop a robust method for automatic fiducial marker detection in MRI simulation images of the prostate and quantify the prostate organ localization accuracy using automatically detected fiducial markers in MRI for pretreatment alignment using cone beam CT (CBCT) images.

Methods and Materials: In this study, a deep learning-based algorithm was used to convert MRI images into labelled fiducial marker volumes. 77 prostate cancer patients who received marker implantation prior to MRI and CT simulation imaging were selected for this study. Multiple-Echo T_1-VIBE MRI images were acquired, and images were stratified (at the patient level) based on the presence of intraprostatic calcifications. Ground truth (GT) contours were defined by an expert on MRI using CT images. Training was done using the pix2pix Generative Adversarial Network (GAN) image-to-image translation package and model testing was performed with five-fold cross validation. For performance comparison, an experienced medical dosimetrist and a medical physicist each manually contoured fiducial markers in MRI images. The percent of correct

detections and F_1 classification scores are reported for markers detected using the automatic detection algorithm and human observers. The patient positioning errors were quantified by calculating the target registration errors (TREs) from fiducial marker driven rigid registration between MRI and CBCT images. TREs were quantified for fiducial marker contours defined on MRI by the automatic detection algorithm and the two expert human observers.

Results: 96% of implanted fiducial markers were correctly identified using the automatic detection algorithm. Two expert raters correctly identified 97% and 96% of fiducial markers, respectively. The F_1 classification score was 0.68, 0.75 and 0.72 for the automatic detection algorithm and two human raters, respectively. The main source of false discoveries was intraprostatic calcifications. The mean TRE differences between alignments from automatic detection algorithm and human detected markers and GT were less than 1 mm.

Conclusions: We have developed a deep learning-based approach to automatically detect fiducial markers in MRI-only simulation images in a clinically representative patient cohort. The automatic detection algorithm-predicted markers can allow for patient setup with similar accuracy to independent human observers.

5.2. Introduction

Many institutions have begun implementing MRI-only simulation imaging for radiation therapy treatment planning of prostate cancer. MRI-only simulation imaging allows for excellent visualization of soft-tissue anatomy, the elimination of co-registration errors due multimodal image registration, reduction of radiation dose from simulation imaging and fewer imaging examinations for patients [1–6]. Clinics implementing MRI-only simulation imaging for prostate cancer typically use fiducial markers for pre-treatment organ localization either as a primary alignment technique or as a secondary check for soft-tissue-based alignment [7].

Using fiducial markers in prostate MRI-only simulation images can be challenging because fiducial markers, as well as calcifications and tissue heterogeneities, appear as signal voids in MRI. Calcifications occur in over 40% of prostate cancer patients and are a common source of fiducial marker misidentification [8]. Incorrect fiducial marker identification could potentially result in misalignment during treatment [7]. Multiple-echo gradient echo (MEGRE) MRI images have been shown to be useful for differentiating fiducial markers and intraprostatic objects such as calcifications and tissue heterogeneities in prostate MRI [9].

Institutions implementing MRI-only simulation imaging of the prostate have recommended for the development of quality assurance methods to validate fiducial marker positions and to aid time-consuming manual fiducial marker detection methods [7,10]. In the MR-PROTECT trial, two MEGRE sequence image sets were acquired to accurately localize marker positions during the relatively long MRI simulation sessions. Manual fiducial marker identification using MEGRE images can be time consuming because it requires identification of markers in multiple image sets [10]. Creating an automatic fiducial marker detection framework could decrease the time spent identifying markers, potentially increase marker detection accuracy, and aid in the

clinical transition of CT-based to MRI-only simulation workflows [8].

There have been implementations of semi-automatic and automatic fiducial marker detection approaches using MRI-only simulation images. Gustafsson et al presented a model-based semi-automatic fiducial marker identification workflow that selects suspect fiducial markers in MEGRE MRI sequences based on signal void volume and appearance changes [8]. Ghose et al used manifold learning and spectral clustering for fiducial marker localization in T_1 and T_2* gradient echo images [11]. Both methods resulted in lower fiducial marker sensitivity than manual detection by human observers. Ghose et al reported that expert radiation therapist observers were able to correctly identify all three implanted seeds in 11/15 patients included in that study. Maspero et al used template matching to detect and segment fiducial markers from a 3D analytical fiducial marker model with a uniform background [12]. This method was able to correctly detect implanted markers in 15/17 patients included in that study. Incorrect marker detections using this technique were due to intraprostatic calcifications. For these studies, the testing groups were not stratified according to the presence of intraprostatic calcifications.

Generative adversarial networks (GANs) have been used for several label-to-segmentation detection tasks of small objects in medical images [13,14]. GANs have been used to accurately detect objects that produce hypointense signals in MRI images. Zhang et al presented a conditional GAN-based semantic segmentation-based approach to detect multiple sclerosis lesions. These lesions appear hypointense in T_1-weighted brain MRI images [15]. The GAN-based architecture was compared to commonly used segmentation frameworks including UNet: a fully convolutional neural network (CNN), Global Convolutional Network (GCN): a deep encoder-decoder network with global convolution blocks and DeepMedic: a patch-based model with a customized 3D CNN architecture for brain lesion segmentation[16–18]. The Sørensen-Dice coefficients comparing

predicted and ground truth labels were 0.67, 0.41, 0.49 and 0.62 for the cGAN, UNet, GCN and DeepMedic models respectively. This GAN-based technique was used to automatically detect small objects that appear hypointense in MRI. Consequently, this method may be suitable to detect fiducial marker-induced signal voids in MRI.

In this work, we develop a conditional GAN-based approach to automatically detect fiducial markers in multiple-echo T_1-weighted prostate MRI images. Using automatically detected fiducial markers, we quantitatively compared the alignment errors from fiducial marker-guided rigid registration of fiducial markers detected on MRI to those in CBCT images. The alignment errors were quantified for fiducial marker segmentations defined by the automatic GAN-based detection algorithm and two independent human observers.

5.3. Methods and Materials

In this study, we develop a GAN-based algorithm for automatic detection of fiducial markers in MRI radiotherapy images and evaluate the patient setup accuracy using automatically detected fiducial markers. The development of an automatic GAN-based detection algorithm included training and testing data set creation, model hyperparameter tuning, model evaluation and alignment accuracy evaluation (*Figure 5.3.1*). Training and testing data were pre-processed by selecting suitable patients for study inclusion, structure contouring on CT, MRI and CBCT image sets and image augmentation for model training. Fiducial marker detection was done using 2D paired MRI/labeled images. Model hyperparameter tuning consisted of optimal model hyperparameter selection using 5-fold cross validation. Model evaluation included comparing markers detected using the automatic GAN-based detection algorithm with the optimal hyperparameters and human observers. The model predicted 2D labelled image slices were resized

back into their original 3D image volumes. The alignment accuracy for fiducial marker-based rigid registration of MRI simulation images and pre-treatment cone beam CT (CBCT) images was quantified. A comparison of alignment accuracies of fiducial markers defined by two independent human observers and the automatic GAN-based detection algorithm is presented.

5.3.A. Patient information and image acquisition settings

77 low-to-intermediate risk prostate cancer patients were retrospectively selected for this study. The median patient age was 71 years old. Patients underwent MRI and CT imaging for radiotherapy simulation and planning between Feb 2018 to Feb 2020. Patients received transperineal implantation of three cylindrical (diameter = 1.2 mm, length = 3 mm) gold fiducial markers (Civco Medical Solutions, Coralville, IA, USA) by a physician using template-guided 17-gauge needles with transrectal ultrasonography (TRUS) guidance. Patients received fiducial marker implantation prior to MRI and CT simulation imaging. 47 patients received MRI and CT imaging on the same day and 30 patients received MRI and CT imaging within a four-day window. 50 patients had a rectum-prostate separation gel implanted prior to CT simulation and MRI examinations (SpaceOAR® Hydrogel, Boston Scientific, MA, USA). Two patients had hip protheses. The study was conducted with Institutional Review Board (IRB) approval (IRB# 19-001731).

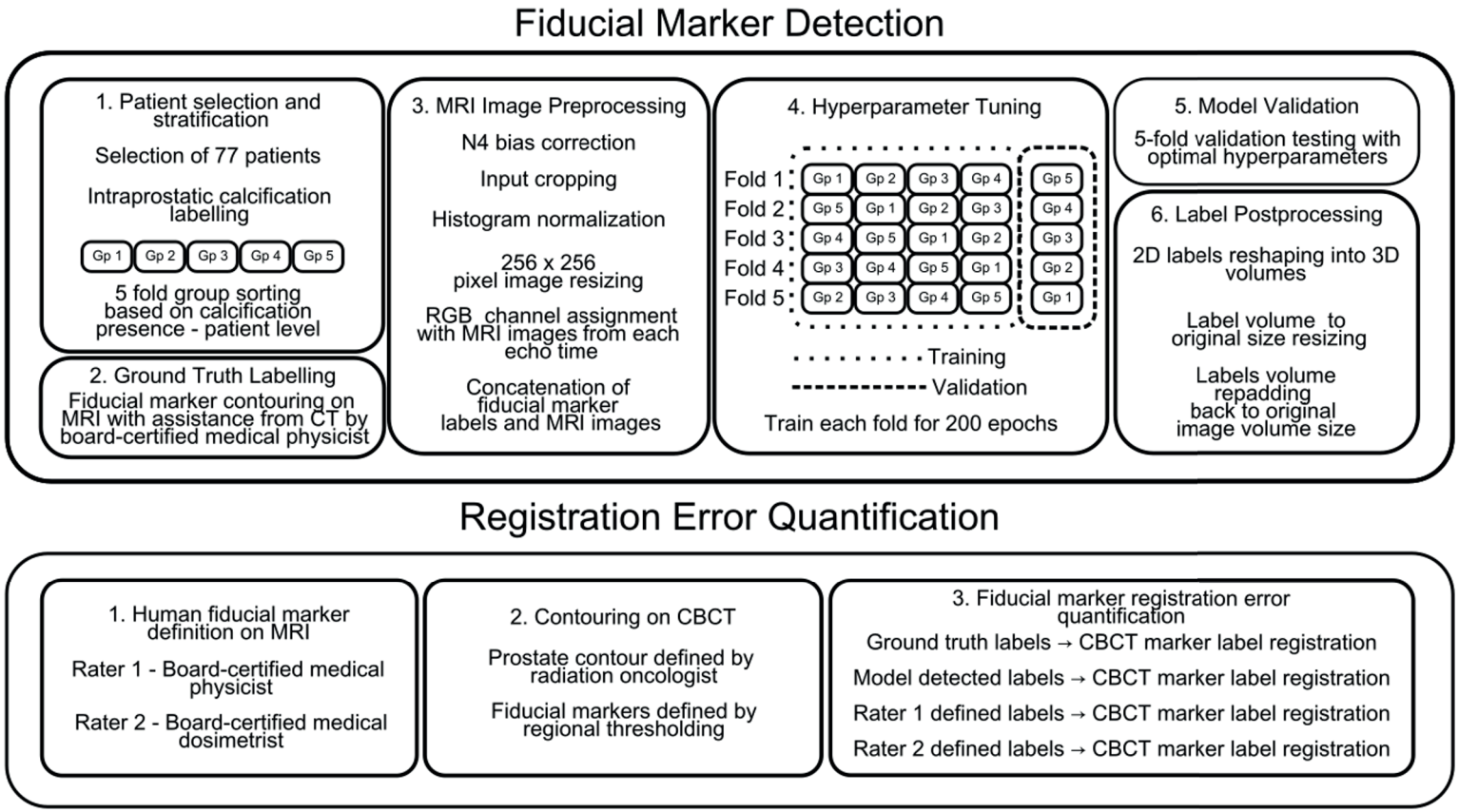


Figure 5.3.1. Workflow summarizing the fiducial marker detection and registration error quantification steps. The automatic fiducial marker detection workflow consists of patient selection and stratification, ground truth labelling, training image preprocessing, hyperparameter tuning, model validation and model-predicted label postprocessing. The registration error quantification steps included selecting independent human observers, CBCT image processing and registration error quantification by rigidly registering fiducial markers detected by the independent raters and GAN-based detection algorithm to fiducial markers in CBCT images. Gp denotes each group (or fold) in five-fold sorting and cross validation.

CT scans were acquired on a Sensation Open with an 82 cm bore (Siemens, Erlangen, Germany). CT scans were acquired with the following image parameters: 120kVp, exposure time = 500 mAs, tube current = 400mA, in-plane matrix = 512 x 512 pixels and a 1.5mm slice thickness. The median in-plane resolution was 1.2 mm x 1.2 mm. Patients were in treatment position prior to imaging with a flat table, knee wedges and Vac-Lok immobilization cushioning.

MRI images were acquired on a 3.0T Siemens (Siemens, Erlangen, Germany) Magnetom Vida (38 patients), 3.0T Siemens Skyra (3.0T) or a 1.5T Avanto Fit (2 patients). Multiple-echo 3D T_1-weighted Volumetric Breath-hold Exam (VIBE) MRI images were acquired for this study. Two echoes were acquired at 1.23 ms and 2.46 ms echo times and 4.17 ms repetition times. Images were acquired without contrast administration. This sequence was used in this study because it quantitatively produces good gold fiducial marker visibility and has been used in previous MRI-only detection workflows [12,19] . Patient setup at the MRI was performed using a knee wedge on a curved couch. Patients were scanned using 18-channel anterior phased array coils (18-channel body coil, Siemens, Erlangen, Germany).

56 patients were treated on a Varian Trilogy Novalis Tx linear accelerator (linac) and 21 patients were treated on a Varian TrueBeam linac (Varian Medical Systems, Palo Alto, California, USA). Pre-treatment CBCT images were acquired for all patients. The setup CBCT from each patient's first radiotherapy fraction was used in this study.

5.3.B. Data preprocessing

Fiducial markers and prostate organs were contoured on CT simulation, MRI and CBCT image sets. The prostate organ was delineated on CT simulation and MRI images by a board-certified radiation oncologist. Fiducial markers were delineated on CT simulation images by a board-

certified medical dosimetrist. Ground truth fiducial markers were defined on MRI images by a board-certified medical physicist. Ground truth fiducial markers were defined on MRI with assistance from prostate contour-guided rigid registration of each MRI and CT simulation image. Ground truth fiducial markers were identified by locating signal voids that appeared in the vicinity of the distinct positive contrast artifacts produced by fiducial markers in CT. Intraprostatic calcifications were identified and contoured on MRI images using the same method except the positive contrast artifacts produced by calcifications on CT were used for guidance. Intraprostatic calcifications that were greater than 2 mm in each axial plane were contoured. 48% of all patients had at least 1 intraprostatic calcification greater than 2 mm. An expert radiation oncologist contoured the prostate in CBCT images. Contouring was performed using MIM (version: 6.6.13, MIM Software Inc, Beachwood, OH, USA). Fiducial markers as visible on MRI, CT and CBCT images along with other delineated features such as calcifications are shown in *Figure 5.3.2.*

Segmentation labels, I_{label}, were created using the MRI fiducial marker and prostate contours. Multi-level masked volumes were created by setting the voxels inside the prostate contour to a value of 1, fiducial markers to a value of 2 and the background to 0. To be suitable for model training, images from each echo time were bias corrected, normalized based on intensity, resized and concatenated. N4 bias correction was applied to each MRI image in MIM [20]. Consensus histograms for images acquired at each echo time were created by counting the voxel intensities inside the prostate contours. A histogram-normalized image was created transforming each T_1-VIBE echo time image so that the histogram of the output images approximately matched the consensus histograms for each echo time. Each image set was resized to 256 x 256 x N_{slices} pixels using a box shaped kernel where N_{slices} is the number of image slices. The resized and normalized image sets were concatenated with the segmentation labels. Image preparation was completed by

concatenating each labelled image with each normalized echo time image.

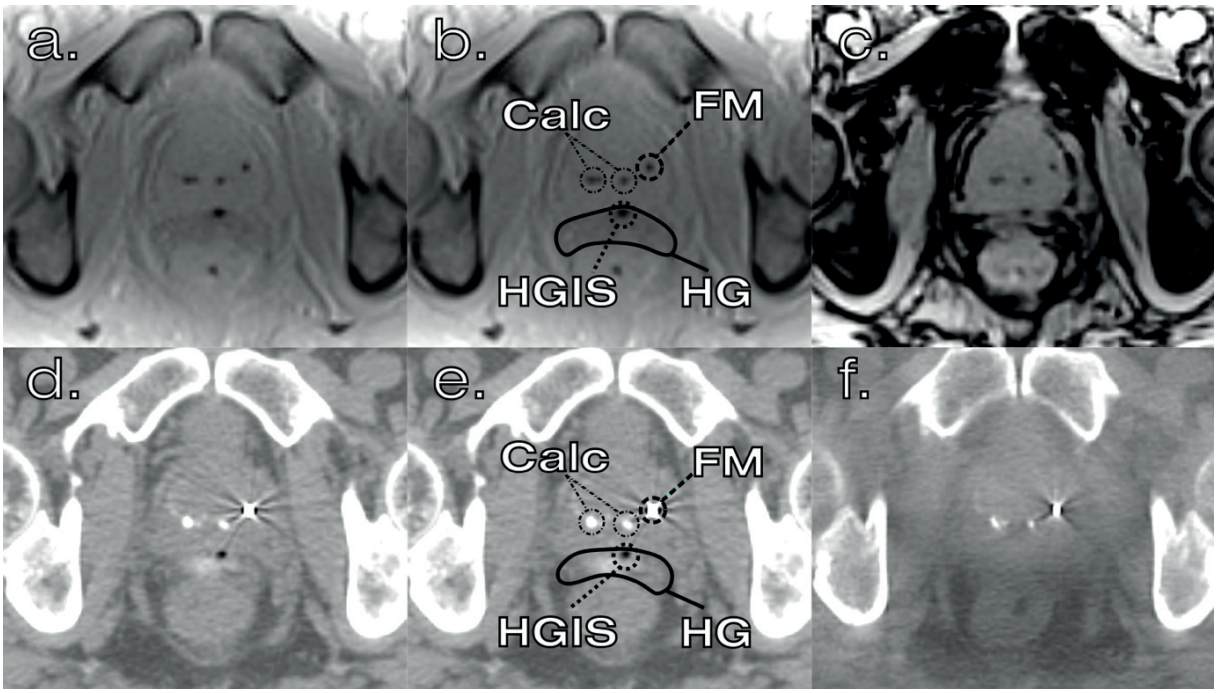


Figure 5.3.2. MRI (a,b,c) and CT (d,e,f) images of an axial section of prostate anatomy. MRI images acquired using the T_1-VIBE sequence at 1.23ms and 2.46ms echo times are shown in (a) and (c) respectively. Radiotherapy planning CT and pre-treatment CBCT images are shown in (d) and (f) respectively. Annotated 1.23ms echo time T1-VIBE and CT images are shown in (b) and (e) respectively. Fiducial markers (FM), calcifications (Calc), hydrogel spacer (HG) and hydrogel spacer injection sites (HGIS) structures are annotated.

5.3.C. Model training and evaluation

Training and testing data were created by stratifying patients according to calcification presence into 5 folds for hyperparameter tuning. Patients were initially stratified into two groups, those with detected intraprostatic calcifications, α, and those without, β. Both groups were split into 5 subgroups, $\alpha_1, \alpha_2,\ldots, \alpha_5$ and $\beta_1, \beta_2,\ldots, \beta_5$. Testing folds were created by combining $[\alpha_1\ \beta_1]$ and training folds were created by combining the exclusive disjunction of the testing group set i.e. $[\alpha_2,\ldots, \alpha_5\ \beta_1, \beta_2,\ldots, \beta_5]$. Image stratification was done at the patient level. Each fold was stratified so that about 40-45% of patients in the testing and training folds had intraprostatic calcifications,

which is a similar proportion to the general prostate cancer population [21,22].

In this study the pix2pix conditional GAN was used with a Pytorch backend on a computer containing a NVIDIA GeForce GTX 1070 GPU [23]. An overview of the pix2pix implementation used in this study is shown in *Figure 5.3.3.* Paired 2D MRI and labelled images were used for training and validation. The pix2pix input requires two co-registered paired images with voxel-wise correspondence. The GAN consists of a discriminator, D, and generator, G, that compete to minimize an objective loss function. The objective of G is to learn a distribution over data q from sampling input variables from a Gaussian distribution $p_{data}(q)$. D is a binary classifier that attempts to determine if a sample is from the training data or if it was generated by G. The generator network used in this study was a U-Net generator optimized for 256x256 pixel image inputs and the discriminator consisted of a three-layer convolutional network [16]. In addition to the adversarial generator and discriminator GAN losses, the model includes an additional least squares loss GAN (LSGAN). LSGAN is based on L_2 norm loss [24,25]. The LSGAN objective functions, V_{LSGAN}, for the discriminator and generator are described by:

$$\min_D V_{LSGAN}(D) = \frac{1}{2}\mathbb{E}_{x\sim p_{data}(x)}[(D(x)-1)^2] + \frac{1}{2}\mathbb{E}_{z\sim p_z(z)}[(D(G(z)))^2] \quad (1)$$

and

$$\min_G V_{LSGAN}(G) = \frac{1}{2}\mathbb{E}_{z\sim p_z(z)}\left[\left(D\left(G(z)\right)-1\right)^2\right] \quad (2)$$

where, $\mathbb{E}_{q\sim p_{data}(q)}$is the expectation value that a probability distribution q matches the real data distribution $p_{data}(q)$. In this study x is labelled image I_{label}, z is MRI image I_{MRI} and $G(z)$ is the generated image from z.

Model training data was loaded in serial batches and on a slice by slice basis. Each model was trained for 200 epochs. The learning rate, lr, and adaptive learning (Adam) optimizer momentum terms, β_1, were optimized as part of hyperparameter tuning [26]. The Adam optimizer

was used with a batch size of 1. Hyperparameter tuning was done by training each fold with different hyperparameter values and selecting hyperparameters that produced the smallest least squares loss across each group. The results from each testing fold were used to evaluate model performance using the optimal hyperparameters.

The 2D to 3D labelled image resorting and postprocessing steps are summarized in *Figure 5.3.4*. The detected fiducial marker outputs were segmented to evaluate model performance. The outputs from the automatic GAN-based detection algorithm were resized back to their original sizes using box-kernel interpolation. A threshold of 2 was applied to the model output to get a binary mask volume containing model-predicted fiducial marker segments (MPFM). The ground truth fiducial marker contours were converted to binary mask volumes containing ground truth fiducial markers (GTFM). Fiducial marker detection statistics were quantified by computing the F_1 scores using MPFM and GTFM volumes. The F_1 score is the harmonic mean of the precision and recall. The F_1 score reaches its best value at 1, representing perfect precision and recall, and worst at 0. Fiducial marker counting statistics were quantified by comparing the relative volume overlap between MPFM to ground truth fiducial markers. Using the ground truth calcification contours, a binary mask containing calcifications (GTCalc) was created to identify markers misidentified as calcifications. True positive (TP) counts were defined if the Sørensen–Dice coefficients between each MPFM and each GTFM was greater than 30%. False positive (FP) marker counts were quantified by counting the markers that had Sørensen–Dice coefficient >30% between MPFMs and objects in GTCalc or by 0% Sørensen–Dice coefficient between markers flagged in MPFM and GTFM. False negative (FN) marker counts were quantified by 0% Sørensen–Dice coefficient between markers flagged in GTFM and not in MPFM. The marker counting statistics were quantified by percent correct marker counts, marker count-based false

detection rate (FDR) and false negative rate (FNR).

Alignment accuracy was quantified by calculating the target registration error (TRE) from fiducial marker-based rigid registration using MRI and CBCT image sets. MRI fiducial marker contours were used from segmentations created using MPFM, GTFM and two independent human observers. For reference, an additional set of registrations was done using fiducial markers defined in the planning CT and CBCT images. To calculate the TRE, fiducial markers were defined in CBCT images by creating a spherical ROI with a 1cm radius encapsulating each fiducial marker and applying a 1000 HU threshold within the ROI. The largest continuous thresholded region was defined as a fiducial marker. The same method was employed to define fiducial markers in CT. The MRI defined fiducial marker contours were rigidly registered to the CBCT fiducial markers with the physician defined contours in CBCT images were used as the target volumes. Translational registrations without rotations were used to mimic the translations performed by a non-robotic linac treatment bed. TRE calculations were computed using the formulism presented by Fitzpatrick et al [27–29]. A board-certified medical physicist with 10 years of experience and a board-certified medical dosimetrist with over 6 years of experience were selected as independent raters. Raters were given an instructional session about how fiducial markers may appear in T_1-VIBE MRI images. Raters were blinded to the CT images. Raters were shown images from each acquired T_1-VIBE echo time and were allowed to view images in axial, sagittal or coronal planes, control the window/level and zoom to aid marker detection. Raters were requested to contour the signal voids considered to be fiducial markers. No limit was set to the number of suspect markers that the raters could select.

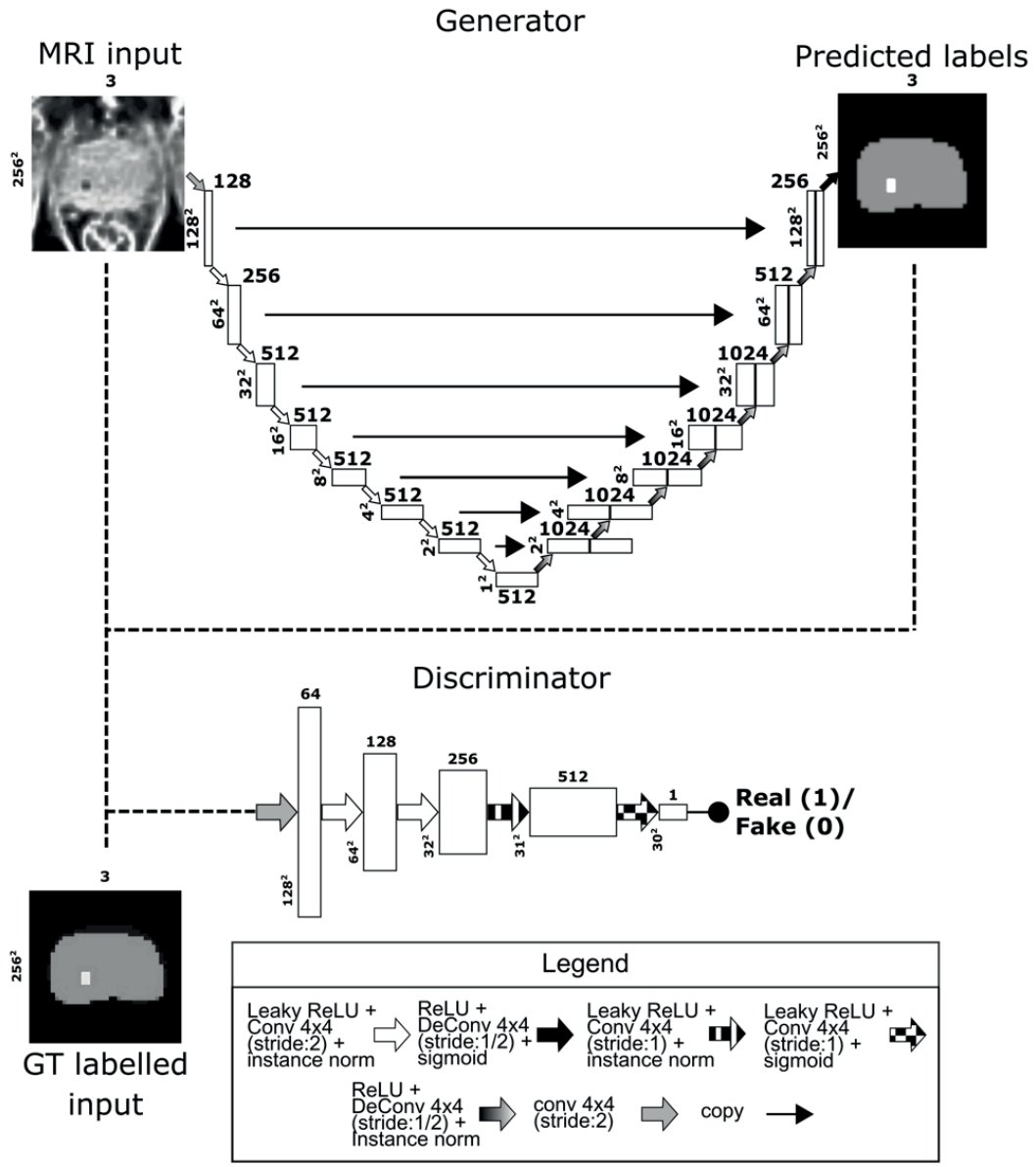


Figure 5.3.3. Summary workflow of the implementation of pix2pix for automatic fiducial marker detection. This study used co-registered MRI images and labelled fiducial marker images as inputs. The generator transforms the input MRI image into a labelled volume. The discriminator measures the similarity of the input image (which could be the target image from the dataset or the generator output) and tries to determine if it was produced by the generator. The weights of the discriminator are then adjusted based on the classification of the input/output image pairs and the input/target image pairs. The generator's weights are then adjusted based on the output of the discriminator and the output and target image differences. The weights are updated until the objective function is minimized.

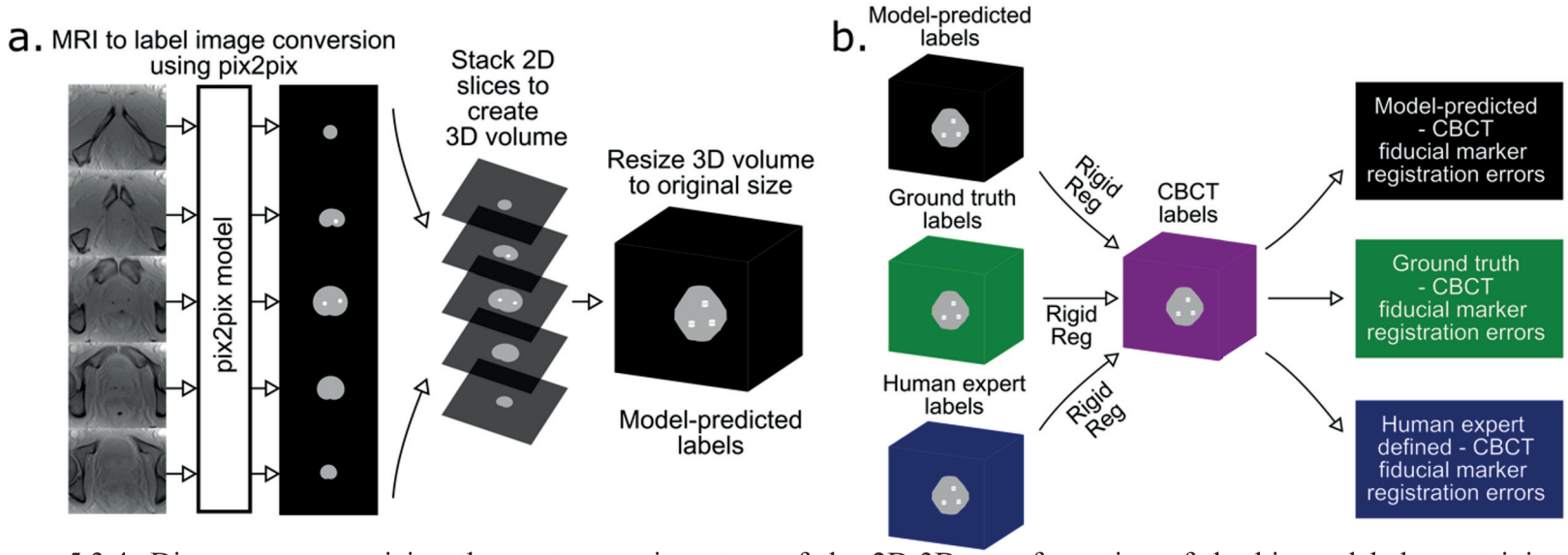


Figure 5.3.4. Diagrams summarizing the postprocessing steps of the 2D-3D transformation of the binary labels containing automatically detected markers (a) and the fiducial marker-based registration steps (b). (a) The 2D-3D image volume transformation included stacking each labelled 2D pix2pix image output in the original image orientation followed by resizing/repadding to the original image volume size. The registration error quantification (b) was done using three sets of labelled volumes: those automatically detected by the model, the ground truth label volumes, and those that were labelled by human experts. Each image set was rigidly registered to fiducial markers in CBCT images and the corresponding TRE was quantified for each labelled volume-CBCT registration.

5.4. Results

The percent marker count accuracy, F_1 scores, FDR and FNR results for the GAN-based detection algorithm and human raters are presented in *Table 5.4.I.* A lr of 0.0003 and β_1 of 0.05 were identified as optimal hyperparameters for the final training and testing steps. Each fold was trained for 200 epochs.

A total of 228 fiducial markers were included in this study. The GAN-based detection algorithm, human rater-1 and human rater-2 correctly identified 220, 221 and 219 fiducial markers, respectively. The GAN-based detection algorithm, human rater-1 and human rater-2 produced 25, 10 and 6 false positive counts, respectively. The GAN-based workflow, human rater-1 and human rater-2 produced 1, 1 and 3 false negative counts, respectively. Calcifications accounted for 84%, 60% and 67% of the false discoveries for the GAN-based detection algorithm, human rater-1 and human rater-2 respectively. There was one instance of a signal void in the vicinity of the hydrogel injection site resulting in a false marker count for human rater 1.

Box plots of the TREs using ground truth-, human observer-, GAN-based detection algorithm and CT-defined fiducial marker contours are shown in *Figure 5.4.1.* In eight patients, a single fiducial marker that was visible on the MRI/CT simulation images was not visible in CBCT images. In one patient, two fiducial markers that were visible on MRI/CT simulation images was not visible in CBCT images. We suspect that the markers may have dislodged between simulation and treatment imaging.

alidation Gro	Detection Rate (%)			F_1 score								
	GAN-based algorith	Rater 1	Rater 2	GAN-based algorith	Rater 1	Rater 2	GAN-based algorith	Rater 1	Rater 2	GAN-based algorith	Rater 1	Rater 2
1	96	98		0.69	0.69		13	8	2	0	0	0
2	91			0.69			19	7	4	0	0	0
3	100	96		0.68	0.77		4	2	0	0	2	2
4	96	100		0.67	0.75		4	4	6	2	0	4
5	100			0.72			10	0	0	0	0	0
Global	96	97		0.68	0.75		11	5	3	0	0	1

Table 5.4.I. Table summarizing the marker detection rates, F_1 scores, false discovery rates and false negative rates for testing fold in addition to the global means. The F_1 score is calculated using the harmonic mean of the classifica precision and recall, where an F_1 score of 1 represents perfect precision and recall. The main source of false discov for both human observers and the GAN-based detection algorithm were intraprostatic calcifications.

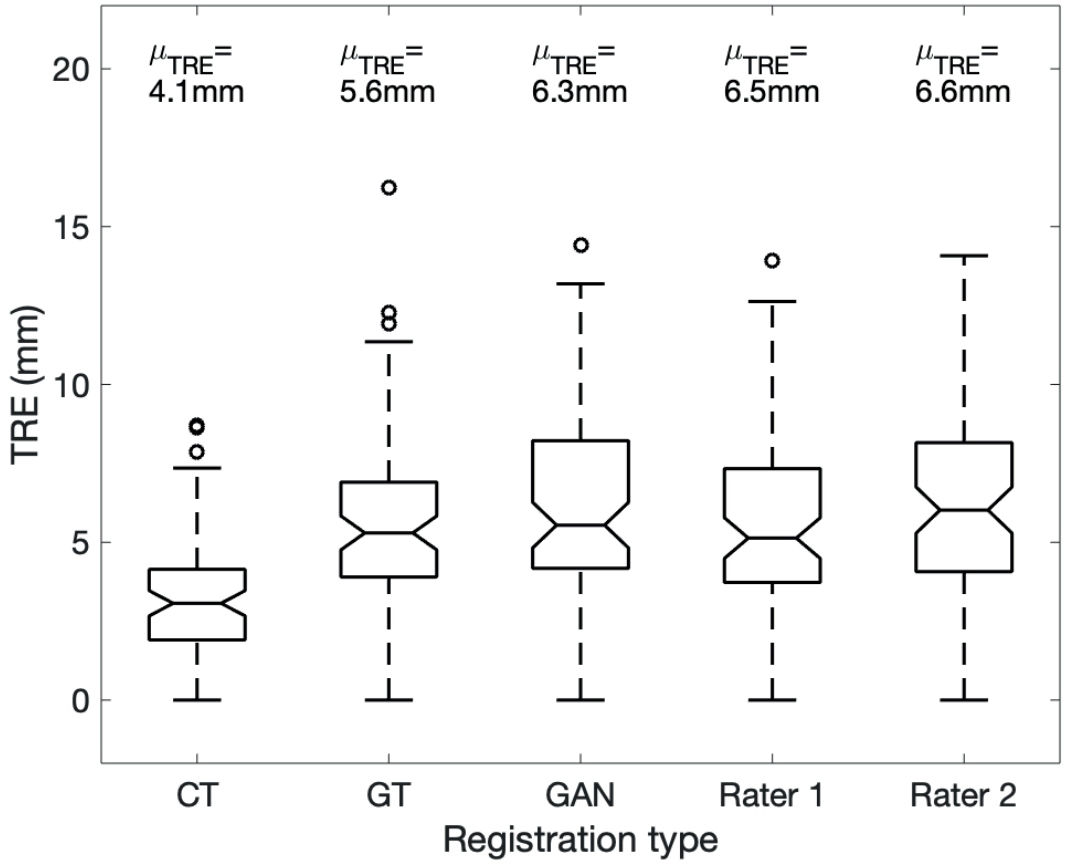


Figure 5.4.1. Box plots showing the TRE distributions from marker contours generated from CT, MRI Ground Truth (GT), GAN and human raters to markers in CBCT images. TRE calculations were based on fiducial marker-based rigid registrations. CBCT prostate contours were drawn by an experienced radiation oncologist. The mean TRE values are indicated by μ_{TRE}.

5.5. Discussion

The purpose of this study was to develop a deep learning-based algorithm for automatic detection of fiducial markers for MRI-only simulation imaging, and to compare the detection accuracy and fiducial marker-based patient setup accuracy of automatically detected fiducial markers to trained human observers. A GAN-based detection algorithm was used to detect fiducial markers in MEGRE images acquired for 77 prostate cancer patients. The GAN-based detection algorithm correctly identified 96% of 220 implanted fiducial markers. Two human raters correctly identified

97% and 95% of implanted fiducial markers respectively. TRE calculations were based on fiducial marker based rigid registrations using fiducial markers defined in CBCT images and segmented fiducial markers on MRI images. The mean TRE differences between ground truth, GAN-based detection algorithm, and human rater-defined fiducial markers on MRI and CT were less than 2.5 mm.

Intraprostatic calcifications were a major source of incorrect marker counts for both the GAN-based detection algorithm and human observers. The use of MEGRE MRI sequences was intended to minimize the number of false marker detections due to calcifications. However, previously published studies have demonstrated that intraprostatic calcifications can still be a source of false fiducial marker detections for automatic fiducial marker detection frameworks [8,11,12]. As part of an extension of this study we tested if calcifications detected using the GAN-based detection algorithm had a detrimental effect on registration. The detected suspect fiducial markers in MRI images were registered to fiducial markers in CBCT images. For example, if a suspect marker was seen to be a calcification in the CBCT image it was treated as an additional fiducial marker. T-tests with a 95% confidence interval demonstrated that there was no statistical difference between TREs (with targets defined in CBCT images) calculated using the automatically detected markers with calcifications and the MRI ground truth defined markers. Calcifications typically do not dislodge in the time between simulation and treatment imaging and could possibly be used as natural fiducial markers [30,31].

This study provides insight into optimal fiducial marker implantation procedures for prostate MRI-only simulation. Current fiducial marker implantation guidelines recommend that three intraprostatic markers are implanted, with one marker positioned in the right base, one in the left mid-gland, and one in the right apex of the gland to ensure that the implant configuration is a

triangle with at least 1-cm spacing [32,33]. In this study, most of the markers that displaced between simulation and treatment day imaging were implanted at the prostate apex. Marker injection location was also a factor in marker-induced signal void discernibility by both the GAN-based detection algorithm and human observers. Sources of false negative fiducial marker counts included markers implanted in or near signal voids created by hydrogel injections sites and markers on the prostate/rectum interface. Therefore, markers implanted in the contralateral midzone and base of the prostate organ may be unreliable for automatic marker detection because of their potential to dislodge. Additionally, markers implanted near the prostate/rectum interface and inside hydrogels may have obscured visibility in MRI images.

There are improvements that could be made to the automatic GAN-based detection algorithm. The current version of the GAN-based detection algorithm uses a greater than 0% DICE overlap condition for marker counting. The rationale behind setting a low DICE limit was to be able to detect smaller markers that might have small overlap with ground truth. A potential source of error from using this condition could be if an incorrect fiducial marker is identified close to the boundary of a real fiducial marker. This could lead to an incorrect marker being identified as a real fiducial marker. Further studies could be conducted to evaluate and establish the optimal criterions for fiducial marker definition.

There are limitations to using a retrospective dataset for model training and evaluation. The clinical use of the automatic GAN-based detection algorithm would require compatibility with prospectively acquired data. Prospectively acquired data may contain feature variations not present in a retrospectively trained model. Acquisition of prospective data could potentially have data variations such as acquisition setting errors, institution specific image acquisition variations or patient specific anatomical variations. The purpose of this study was to provide a proof of concept

of the automatic GAN-based detection algorithm. Studies with prospectively acquired data accounting for simulated acquisition setting differences and a larger cohort of patients could help make the automatic GAN-based algorithm more robust.

A deep learning-based marker detection algorithm could potentially be incorporated in a prostate MRI-only simulation workflow to assist marker localization and definition. Clinics that have commissioned MRI-only prostate radiotherapy treatments have recommended acquiring CT images for fiducial marker position verification [7]. However, the acquisition of CT scans is contrary to the purpose of MRI-only simulation imaging. This GAN-based marker detection algorithm can potentially be used in a marker detection and automatic contouring framework with fiducial markers verified and checked by a dosimetrist or medical physicist.

Additional studies can be conducted with the automatic GAN-based detection algorithm. The ability of the model to improve discrimination between fiducial markers and calcifications can be tested by creating models with images acquired at finer slice thicknesses and with additional echoes. The clinical utility of the automatic workflow can be further tested with a larger cohort of prostate cancer patients receiving MRI-only simulation examinations. The generalization of this algorithm to other anatomical sites such as the rectum and pancreas may also be worth investigating [34,35].

5.6. Conclusions

We have developed a deep learning-based marker detection algorithm to automatically detect gold fiducial markers in MRI images with a clinically representative cohort of prostate cancer patients. Fiducial marker-based pretreatment setup using automatically detected fiducial markers can be done with similar accuracy to fiducial markers identified on MRI by independent human observers.

5.7. Citations

1. Rasch C, Barillot I, Remeijer P, Touw A, van Herk M, Lebesque JV. Definition of the prostate in CT and MRI: a multi-observer study. *Int J Radiat Oncol Biol Phys*. 1999;43(1):57-66.

2. Roach M, Faillace-Akazawa P, Malfatti C, Holland J, Hricak H. Prostate volumes defined by magnetic resonance imaging and computerized tomographic scans for three-dimensional conformal radiotherapy. *Int J Radiat Oncol*. 1996;35(5):1011-1018. doi:10.1016/0360-3016(96)00232-5

3. McLaughlin PW, Evans C, Feng M, Narayana V. Radiographic and anatomic basis for prostate contouring errors and methods to improve prostate contouring accuracy. *Int J Radiat Oncol Biol Phys*. 2010;76(2):369-378. doi:10.1016/j.ijrobp.2009.02.019

4. Schmidt MA, Payne GS. Radiotherapy planning using MRI. *Phys Med Biol*. 2015;60(22):R323. doi:10.1088/0031-9155/60/22/R323

5. Lee YK, Bollet M, Charles-Edwards G, et al. Radiotherapy treatment planning of prostate cancer using magnetic resonance imaging alone. *Radiother Oncol J Eur Soc Ther Radiol Oncol*. 2003;66(2):203-216.

6. Metcalfe P, Liney GP, Holloway L, et al. The Potential for an Enhanced Role for MRI in Radiation-therapy Treatment Planning. *Technol Cancer Res Treat*. 2013;12(5):429-446. doi:10.7785/tcrt.2012.500342

7. Greer P, Martin J, Sidhom M, et al. A Multi-center Prospective Study for Implementation of an MRI-Only Prostate Treatment Planning Workflow. *Front Oncol*. 2019;9. doi:10.3389/fonc.2019.00826

8. Gustafsson Christian, Korhonen Juha, Persson Emilia, Gunnlaugsson Adalsteinn, Nyholm Tufve, Olsson Lars E. Registration free automatic identification of gold fiducial markers in MRI target delineation images for prostate radiotherapy. *Med Phys*. 2017;44(11):5563-5574. doi:10.1002/mp.12516

9. Schieda N, Avruch L, Shabana WM, Malone SC. Multi-echo gradient recalled echo imaging of the pelvis for improved depiction of brachytherapy seeds and fiducial markers facilitating radiotherapy planning and treatment of prostatic carcinoma. *J Magn Reson Imaging*. 2015;41(3):715-720. doi:10.1002/jmri.24590

10. MR-PROTECT: Clinical feasibility of a prostate MRI-only radiotherapy treatment workflow and investigation of acceptance criteria | Radiation Oncology | Full Text. Accessed July 26, 2020. https://ro-journal.biomedcentral.com/articles/10.1186/s13014-020-01513-7

11. Ghose S, Mitra J, Rivest-Hénault D, et al. MRI-alone radiation therapy planning for prostate cancer: Automatic fiducial marker detection. *Med Phys*. 2016;43(5):2218. doi:10.1118/1.4944871

12. Maspero M, Berg CAT van den, Zijlstra F, et al. Evaluation of an automatic MR-based gold fiducial marker localisation method for MR-only prostate radiotherapy. *Phys Med Biol*. 2017;62(20):7981. doi:10.1088/1361-6560/aa875f

13. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*. 2018;321:321-331. doi:10.1016/j.neucom.2018.09.013

14. Xue Y, Xu T, Zhang H, Long LR, Huang X. SegAN: Adversarial Network with Multi-scale L1 Loss for Medical Image Segmentation. *Neuroinformatics*. 2018;16(3):383-392. doi:10.1007/s12021-018-9377-x

15. Zhang C, Song Y, Liu S, et al. MS-GAN: GAN-Based Semantic Segmentation of Multiple Sclerosis Lesions in Brain Magnetic Resonance Imaging. In: *2018 Digital Image Computing: Techniques and Applications (DICTA)*. ; 2018:1-8. doi:10.1109/DICTA.2018.8615771

16. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, eds. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Lecture Notes in Computer Science. Springer International Publishing; 2015:234-241. doi:10.1007/978-3-319-24574-4_28

17. Peng C, Zhang X, Yu G, Luo G, Sun J. Large Kernel Matters -- Improve Semantic Segmentation by Global Convolutional Network. In: ; 2017:4353-4361. Accessed March 13, 2020. http://openaccess.thecvf.com/content_cvpr_2017/html/Peng_Large_Kernel_Matters_CVPR_2017_paper.html

18. Kamnitsas K, Ledig C, Newcombe VFJ, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med Image Anal*. 2017;36:61-78. doi:10.1016/j.media.2016.10.004

19. Singhrao K, Ruan D, Fu J, et al. Quantification of fiducial marker visibility for MRI-only prostate radiotherapy simulation. *Phys Med Biol*. 2020;65(3):035015. doi:10.1088/1361-6560/ab65db

20. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: Improved N3 Bias Correction. *IEEE Trans Med Imaging*. 2010;29(6):1310-1320. doi:10.1109/TMI.2010.2046908

21. Shoskes DA, Lee C-T, Murphy D, Kefer J, Wood HM. Incidence and significance of prostatic stones in men with chronic prostatitis/chronic pelvic pain syndrome. *Urology*. 2007;70(2):235-238. doi:10.1016/j.urology.2007.04.008

22. Smolski M, Turo R, Whiteside S, Bromage S, Collins GN. Prevalence of prostatic calcification subtypes and association with prostate cancer. *Urology*. 2015;85(1):178-181. doi:10.1016/j.urology.2014.09.026

23. Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-Image Translation with Conditional Adversarial Networks. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. ; 2017:5967-5976. doi:10.1109/CVPR.2017.632

24. Mao X, Li Q, Xie H, Lau RYK, Wang Z, Smolley SP. Least Squares Generative Adversarial Networks. In: *2017 IEEE International Conference on Computer Vision (ICCV)*. ; 2017:2813-2821. doi:10.1109/ICCV.2017.304

25. Sanchez I, Vilaplana V. Brain MRI super-resolution using 3D generative adversarial networks. Published online April 11, 2018. Accessed March 20, 2020. https://openreview.net/forum?id=rJevSbniM

26. Bock S, Weiß M. A Proof of Local Convergence for the Adam Optimizer. In: *2019 International Joint Conference on Neural Networks (IJCNN)*. ; 2019:1-8. doi:10.1109/IJCNN.2019.8852239

27. Fitzpatrick JM, West JB, Maurer CR. Predicting error in rigid-body point-based registration. *IEEE Trans Med Imaging*. 1998;17(5):694-702. doi:10.1109/42.736021

28. Danilchenko A, Fitzpatrick JM. General approach to error prediction in point registration. In: ; 2010:76250F. doi:10.1117/12.843847

29. Liu L, Tian Z, Zhang Z, Fei B. Computer-aided Detection of Prostate Cancer with MRI: Technology and Applications. *Acad Radiol*. 2016;23(8):1024-1046. doi:10.1016/j.acra.2016.03.010

30. Sbai A, Thariat J, Tachfouti N, Pan Q, Lagrange J-L. Calcifications intraprostatiques : marqueurs fiduciels naturels dans la radiothérapie guidée par l'image du cancer de la prostate. *Cancer/Radiothérapie*. 2014;18(8):740-744. doi:10.1016/j.canrad.2014.07.161

31. Hanna SA, Neves-Junior WFP, Marta GN, Haddad CMK, da Silva JLF. Role of Intra- or Periprostatic Calcifications in Image-Guided Radiotherapy for Prostate Cancer. *Int J Radiat Oncol*. 2012;82(3):1208-1216. doi:10.1016/j.ijrobp.2011.03.059

32. Shinohara K, Roach M. Technique for implantation of fiducial markers in the prostate. *Urology*. 2008;71(2):196-200. doi:10.1016/j.urology.2007.10.011

33. Linden RA, Weiner PR, Gomella LG, et al. Technique of outpatient placement of intraprostatic fiducial markers before external beam radiotherapy. *Urology*. 2009;73(4):881-886. doi:10.1016/j.urology.2008.10.071

34. Gurney-Champion OJ, Lens E, van der Horst A, et al. Visibility and artifacts of gold fiducial markers used for image guided radiation therapy of pancreatic cancer on MRI. *Med Phys*. 2015;42(5):2638-2647. doi:10.1118/1.4918753

35. van den Ende RPJ, Rigter LS, Kerkhof EM, et al. MRI visibility of gold fiducial markers for image-guided radiotherapy of rectal cancer. *Radiother Oncol*. 2019;132:93-99. doi:10.1016/j.radonc.2018.11.016

CHAPTER 6 - Conclusions and Future Work

6.1. Conclusions

The goals of this thesis were to develop clinical solutions to improve QA processes for prostate MRI-only radiation therapy simulation imaging and optimize fiducial marker detection methods to aid pretreatment patient positioning using MRI-only simulation images. Chapters 2 and 3 of this thesis involved the development of technologies to improve QA testing processes of EBRT treatments utilizing MRI-only simulation imaging. Chapter 4 was focused on using quantitative criteria to identify the optimal clinical MRI sequences for use with clinically used gold fiducial markers and identify an optimal commercially available fiducial marker for MRI-only simulation imaging. Chapter 5 summarized work on the development of an automatic fiducial marker detection algorithm which could potentially aid in the clinical transition of CT-based to MRI-only simulation imaging workflows. This concluding chapter will summarize the major conclusions from each chapter and discuss future studies

The goal of Chapter 2 was to create a generalizable model using carrageenan-based tissue-surrogate materials for MRI and CT imaging. Using carrageenan, water, Gd-contrast, agarose, GMs and $CaCO_3$, a generalizable fit model was developed that covered a range of T_1 /T_2 relaxation times and CT numbers from 82 to 2180 ms, 12 to 475 ms and -117 to 914 HU, respectively. Using this generalizable fit model, nine diverse tissue surrogates were created with good T_1/T_2 relaxation time and CT number agreement compared with published in vivo measurements. Therefore, this system of materials can mimic MRI and CT characteristics for bone and a wide range of soft tissues such as adipose tissue and skeletal muscle.

There are additional experiments that can be performed using the system of materials

developed in Chapter 2 including testing additional additives, testing the T_2* relaxation times of this system of materials and quantifying the longevity of the system of materials. Testing additional additives such as superparamagnetic iron oxide-based contrast agents may provide additional modes of T_1/T_2 relaxation. Quantifying the T_2* relaxation properties this system of materials may also provide insights about how to improve the bone-tissue surrogate materials. Finally, further experiments can be done to characterize the stability of the imaging characteristics of this system of materials over time.

In Chapter 3, a realistic anthropomorphic pelvis phantom was developed for QA testing MRI-based radiation therapy techniques. The phantoms imaging utility at 1.5T and 3.0T was quantified in addition to its ability to produce realistic sCT images using commercially available software. Dosimetry was also characterized by delivery of two types of commonly used prostate radiation therapy plans. This study demonstrated that the developed anthropomorphic pelvis phantom provided tissue-like contrast on T_1 and T_2 weighted MRI and CT images and can be used to validate MRI-based radiation therapy processes through comparison with standard CT-based radiation therapy processes.

Further studies using the work demonstrated in Chapter 3 can be done. The phantom can be adapted for several therapeutic and imaging QA applications. The generalizability of the phantom fabrication method could aid in the development of multimodal phantoms for benchmarking multisite MRI-based radiation therapy clinical trials. Additionally, the phantom could be printed with deformable interfaces, which could increase its utility for MRI-guided radiation therapy applications. This could be done using dual extruder 3D-printers containing a deformable filament (such as thermoplastic polyurethane) and a solid-at-room temperature filament (such as PLA). Using a dual extruder 3D-printer a tissue-mimicking MRI/CT phantom

with a mixture of solid and deformable walls could be fabricated and filled with the deformable carrageenan-based materials presented in this work. Additionally, the framework outlined in this work could allow for modular MRI/CT phantoms to be developed potentially incorporating many deformable organs.

In Chapter 4, the visibility of fiducial marker-induced signal voids in MRI images was quantitatively characterized in a patient and phantom study. The patient study involved testing the visibility and distinguishability of gold fiducial markers across a variety of commonly used sequences for MRI-only simulation imaging for radiation therapy. The quantitative metrics used to determine visibility and distinguishability were lCNR, Vfall, AMV and TRE. Human observers also rated the visibility of the markers in images from each pulse sequence. The phantom study sought to determine which commercially available fiducial marker-induced signal void displayed the best visibility across a variety of commonly used MRI pulse sequences in phantom. The major conclusions from the patient study were that non-contrast-enhanced T_1-weighted MEGRE MRI sequences produced the best visibility of clinically used gold fiducial markers. The phantom study demonstrated that markers manufactured from gold doped with iron quantitatively produced superior visibility compared to their pure gold counterparts.

The work presented in Chapter 4 can be expanded by investigating the rate of marker displacements in prospective patient studies. In this study gold doped with iron markers were deployed in phantom as an aggregate cluster (as opposed to a line). Most commercially used fiducial markers are deployed as a single solid cylindrical entity and can easily be dislodged. Marker dislocation is a common failure mode of fiducial marker-based positioning and alignment. Prospective studies may help elucidate if fiducial markers deployed as aggregate clusters dislodge at a higher rate than conventional cylindrical fiducial markers.

In Chapter 5, a deep learning-based framework for automatic detection of fiducial markers for MRI-only simulation imaging was developed. Using the automatically detected markers, the framework's detection accuracy and alignment quality was compared to human observers. A GAN-based deep learning architecture was used to classify fiducial markers in MEGRE images acquired for 77 prostate cancer patients. In conclusion, a deep learning-based marker detection framework to automatically detect and classify gold fiducial markers in a clinically representative cohort of prostate cancer patients was developed. Using fiducial markers detected using the automatic GAN-algorithm, pre-treatment prostate organ alignment can be done with similar accuracy to independent human observers.

Additional experiments to expand the work presented in Chapter 5 can be conducted including comparing the detection accuracy of 2D vs 3D deep-learning algorithms and prospective clinical implementation experiences of using automatic fiducial marker detection workflows can be evaluated. In the study presented in Chapter 4, a 2D GAN was utilized because insufficient data was available to train a 3D GAN. However, with sufficient data both 2D and 3D models could be developed, and their classification accuracy could be compared.

6.2. Closing Remarks

The clinical implementation of MRI-only simulation imaging for radiation therapy has scientific, technological, clinical and economic challenges to its widespread use. However, the technology can bring immediate benefits to radiation therapy clinics such as improvements in imaging for treatment response. This thesis was focused on developing scientific and technological approaches for prostate MRI-only simulation imaging. There still remain scientific, technological and clinical challenges to MRI-only simulation imaging such as adaptation of the technology for other clinical

sites. Evaluations of the cost/benefit analyses have yet to be done on the economic impact of the adoption of MRI-only simulation and cogitation of this topic is warranted. This closing section will include a brief discussion of these topics because they have important ramifications for radiation therapy clinics planning to invest in an expensive MRI scanner.

There are scientific and technological challenges to the clinical implementation of MRI-only simulation imaging. Radiation therapy is used to treat diseases located in anatomical regions such as the head and neck, thorax, abdomen, pelvis and extremities. This thesis was focused on a subtype of pelvic cancers that affects men. More work needs to be done to allow for robust MRI-only simulation imaging for radiation therapy of other treatment sites. Some of the tools and methods developed in this thesis are applicable to other cancer sites. The methods used in the development of generalizable tissue mimicking materials and the development of anthropomorphic phantoms can potentially be adapted for other radiation therapy treatment sites [1]. Additionally, methods used in identification and automatic detection of fiducial markers can be utilized for treatment sites such as the abdomen where fiducial markers are used for pre-treatment patient positioning and alignment [2,3]. However, more methods need to be developed to allow MRI-only simulation imaging of multiple treatment sites. Motion management strategies such as 4DMRI have yet to see widespread application for MRI-only simulation of thoracic and abdominal cancers which may be subject to respiratory and cardiac motion [4,5].

The clinical justification of MRI-only simulation does require additional investigation. A major argument for the implementation of MRI-only simulation is improved tumor definition resulting from the elimination of multimodal image registration errors from traditional treatment planning workflows. However, errors can be introduced in parts of the treatment workflow that may overwhelm the reduction in systematic errors from treatment planning such as incorrect pre-

treatment patient setup [6]. Multisite clinical trials and normal tissue toxicity studies are warranted to show the therapeutic benefits of using MRI-only simulation imaging over the current treatment standard [6,7].

There are economic considerations that may dissuade clinics adopting dedicated MRI scanners for radiation therapy simulation. A typical MRI simulator can cost in excess of $1 million compared to a CT simulator which can cost between $500,000 to $750,000 [8–10]. MRI scanners also have additional costs associated with hiring specialized staff such as operators, engineers and safety personnel as well as costs associated with MRI specific shielding requirements [11]. The increased operation cost may be passed on to patients. This may incentivize clinics to only invest in a CT scanner for radiation therapy simulation purposes. However, there are several ways that the patient and institution costs can be reduced from the use of MRI-only simulation imaging for radiation therapy. Insurance companies typically reimburse for a single simulation scan per radiation therapy plan[11]. Patients receiving MRI-only simulation could thereby have lower out-of-pocket costs because they would not be receiving an additional CT simulation scan. Additionally, the amount billed for MRI-only simulation scans may decrease over time because insurance companies may adapt to the presence of new technologies for radiation therapy treatments. This is especially demonstrated by the changes in radiation therapy technologies and treatment techniques over the past 20 years [12]. Additionally, clinics can save on the cost of adopting MRI-only simulation and gauge patient throughput by retrofitting diagnostic MRI scanners with flat couches and in room lasers for patient positioning. This type of resource sharing may allow clinics to gain the benefits of MRI-only radiation therapy simulation without directly investing in a new MRI scanner and may help provide economic justification for buying a new scanner if the patient throughput becomes sufficiently high.

Finally, the clinical use of MRI-only simulation imaging has the potential to improve treatment response assessment. An advantage of MRI-only simulation imaging is that patient images can be acquired in treatment position which can allow improvements to the quality of treatment response. Studies have shown that using MRI images acquired in treatment position decreases bladder and rectal toxicity for patients who had CTV contours derived from MRI simulation images [13]. Acquisition of multimodal images in treatment position may allow for more consistent organ positioning because of the reduction in anatomical deformations between simulation and pre-treatment imaging. For example, acquisition of immobilized MRI images for treatment planning of head and neck cancers may reduce GTV definition uncertainties. For some head and neck cancer patients treated at UCLA, GTVs are defined using both planning CT and a diagnostic PET/CT images. Diagnostic PET/CT images are acquired on a curved couch with the patient in supine position and an anterior head tilt. Immobilized head and neck images for radiation therapy simulation are acquired on a flat couch with minimal head tilt. The biomechanically different anatomical positions of head anatomy make GTV definition challenging without the use of deformable image registration. Therefore, acquisition of multimodality images in treatment position for tumor definition for treatment planning may result in reduced radiation therapy margins (and possibly normal tissue toxicity) [14–16]. Additionally, the acquisition of immobilized MRI images may improve the ease of acquiring longitudinal imaging for treatment response. Longitudinal imaging can involve acquisition of pre-, intra- and post-treatment images to assess therapeutic tumor response. A number of novel functional MRI-based tools have been utilized in radiology departments including pH imaging for brain tumor and diffusion weighted MRI for prostate cancer post-treatment response [17–20]. Radiology departments that acquire images for treatment response typically do so without immobilization. This makes it difficult to assess if

changes in post-treatment tumor volume or appearance is due to differences in patient position and/or differences in interdepartmental scanning protocols. Acquisition of treatment response and longitudinal images in treatment position could allow for radiation therapy clinics to more accurately assess if locoregional tumor failures are due to errors in treatment delivery or to evaluate tumor response to radiation during fractionation.

There are numerous challenges to the clinical implementation of MRI-only simulation imaging. Some of these challenges include the need for scientific and technological developments for multisite MRI-only simulation imaging, the need for clinical trials to evaluate the long-term benefits of using this technology and cost/benefit analyses of this technology in radiation therapy clinics. This work was designed to address some scientific and technological challenges to clinical development of MRI-only simulation imaging for radiation therapy of the male pelvis. The development of solutions to the multisite clinical implementation challenges will allow for the full potential of MRI-only simulation imaging for radiation therapy to be achieved such as a potential reduction in treatment margins, reduction in treatment induced toxicities and improvements in locoregional tumor control. The completion of this work represents a stride toward the achievement of that goal.

6.3. Bibliography

1. Singhrao K, Fu J, Wu HH, et al. A novel anthropomorphic multimodality phantom for MRI-based radiotherapy quality assurance testing. *Medical Physics*. n/a(n/a). doi:10.1002/mp.14027

2. Park CK, Pritz J, Zhang GG, Forster KM, Harris EER. Validating fiducial markers for image-guided radiation therapy for accelerated partial breast irradiation in early-stage breast cancer. *Int J Radiat Oncol Biol Phys*. 2012;82(3):e425-431. doi:10.1016/j.ijrobp.2011.07.027

3. Slagowski JM, Colbert LE, Cazacu IM, et al. Evaluation of the Visibility and Artifacts of 11 Common Fiducial Markers for Image Guided Stereotactic Body Radiation Therapy in the Abdomen. *Practical Radiation Oncology*. Published online January 24, 2020. doi:10.1016/j.prro.2020.01.007

4. Du D, Caruthers SD, Glide-Hurst C, et al. High quality T2-weighted 4D magnetic resonance imaging for radiation therapy applications. *Int J Radiat Oncol Biol Phys*. 2015;92(2):430-437. doi:10.1016/j.ijrobp.2015.01.035

5. Owrangi AM, Greer PB, Glide-Hurst CK. MRI-only treatment planning: benefits and challenges. *Phys Med Biol*. 2018;63(5):05TR01. doi:10.1088/1361-6560/aaaca4

6. Tenhunen M, Korhonen J, Kapanen M, et al. MRI-only based radiation therapy of prostate cancer: workflow and early clinical experience. *Acta Oncologica*. 2018;57(7):902-907. doi:10.1080/0284186X.2018.1445284

7. Metcalfe P, Liney GP, Holloway L, et al. The Potential for an Enhanced Role for MRI in Radiation-therapy Treatment Planning. *Technol Cancer Res Treat*. 2013;12(5):429-446. doi:10.7785/tcrt.2012.500342

8. Sarracanie M, LaPierre CD, Salameh N, Waddington DEJ, Witzel T, Rosen MS. Low-Cost High-Performance MRI. *Sci Rep*. 2015;5. doi:10.1038/srep15177

9. Sistrom CL, McKay NL. Costs, charges, and revenues for hospital diagnostic imaging procedures: differences by modality and hospital characteristics. *J Am Coll Radiol*. 2005;2(6):511-519. doi:10.1016/j.jacr.2004.09.013

10. Mayo-Smith WW, Rhea JT, Smith WJ, Cobb CM, Gareen IF, Dorfman GS. Transportable versus Fixed Platform CT Scanners: Comparison of Costs. *Radiology*. 2003;226(1):63-68. doi:10.1148/radiol.2261012047

11. Chakravarty A, Debnath J. Life cycle costing as a decision making tool for technology acquisition in radio-diagnosis. *Med J Armed Forces India*. 2015;71(1):38-42. doi:10.1016/j.mjafi.2014.10.004

12. Shen X, Showalter TN, Mishra MV, et al. Radiation Oncology Services in the Modern Era: Evolving Patterns of Usage and Payments in the Office Setting for Medicare Patients From 2000 to 2010. *JOP*. 2014;10(4):e201-e207. doi:10.1200/JOP.2013.001270

13. Sander L, Langkilde NC, Holmberg M, Carl J. MRI target delineation may reduce long-term toxicity after prostate radiotherapy. *Acta Oncologica*. 2014;53(6):809-814. doi:10.3109/0284186X.2013.865077

14. Nuyttens J, Quint S, Es H de R, Baartman EA, Dirkx ML. Impact of Different Immobilization Devices and Setup Correction Protocols on Required CTV-PTV Margins in Patients With a Sarcoma of the Lower Extremities. *International Journal of Radiation Oncology, Biology, Physics*. 2012;84(3):S170. doi:10.1016/j.ijrobp.2012.07.441

15. Immobilization versus no immobilization for pelvic external beam radiotherapy. Accessed June 19, 2020. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6035901/

16. White P, Yee CK, Shan LC, Chung LW, Man NH, Cheung YS. A comparison of two systems of patient immobilization for prostate radiotherapy. *Radiat Oncol*. 2014;9:29. doi:10.1186/1748-717X-9-29

17. Hauser T, Essig M, Jensen A, et al. Prediction of treatment response in head and neck carcinomas using IVIM-DWI: Evaluation of lymph node metastasis. *Eur J Radiol*. 2014;83(5):783-787. doi:10.1016/j.ejrad.2014.02.013

18. Schreuder SM, Lensing R, Stoker J, Bipat S. Monitoring treatment response in patients undergoing chemoradiotherapy for locally advanced uterine cervical cancer by additional diffusion-weighted imaging: A systematic review. *J Magn Reson Imaging*. 2015;42(3):572-594. doi:10.1002/jmri.24784

19. Nunes TF, Szejnfeld D, Szejnfeld J, et al. Assessment of Early Treatment Response With DWI After CT-Guided Radiofrequency Ablation of Functioning Adrenal Adenomas. *American Journal of Roentgenology*. 2016;207(4):804-810. doi:10.2214/AJR.16.16207

20. Consolino L, Anemone A, Capozza M, et al. Non-invasive Investigation of Tumor Metabolism and Acidosis by MRI-CEST Imaging. *Front Oncol*. 2020;10. doi:10.3389/fonc.2020.00161