Title
Multimodal imaging and deep learning-based methods for improved dose calculation accuracy in photon and proton radiotherapy

Author
Scholey, Jessica Elizabeth

Publication Date
2021

Peer reviewed|Thesis/dissertation

Multimodal imaging and deep learning-based methods for improved dose calculation accuracy in photon and proton radiotherapy

by

Jessica Scholey

DISSERTATION

Submitted in partial satisfaction of the requirements for degree of

DOCTOR OF PHILOSOPHY

in

Bioengineering

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

AND

UNIVERSITY OF CALIFORNIA, BERKELEY

Approved:

DocuSigned by:

1644A2CD853841E... Peder Larson

Chair

DocuSigned by:

Steven Conolly

DocuSigned by:

Chris John Diederich

07F6E887F166488...

Committee Members

Acknowledgements

First and foremost, I would like to thank my research advisor, Peder Larson, who welcomed me into the world of Radiology and shaped my love of imaging. He allowed me a great deal of independence while also providing unwavering guidance and encouragement. Peder is the type of scientist, teacher, and leader I aspire to be. I would also like to thank Atchar Sudhyadhom, whose mentorship and continued collaboration has had profound impacts on my career as a junior physicist. This dissertation would not be possible without both of you.

I would also like to give special thanks to members of the bioengineering community: Dan Vigneron, who was my graduate advisor and has provided guidance on many topics ranging from course selection to career advice for junior faculty. Thank you to Chris Diederich, who was the Chair of my Qualifying Exam Committee, member of my Dissertation Committee, and has provided invaluable guidance on navigating being a BioE graduate student and physicist in Radiation Oncology. Thank you to Steven Conolly for serving on my Dissertation Committee and always sharing your innovative ideas for applying physics to the clinic. Thank you to Chunlei Liu for serving on my Qualifying Exam Committee and helping to expand my knowledge on quantitative MRI. Thank you to Joe Hsu for serving on my Qualifying Exam Committee and always contributing your clinical perspective to my research ideas. To my lab mates in Byers Hall: Abhe Rajagopal, Nick Dwork, Andrew Leynes, Kirti Magudia, Nikhil Deveshwar, Fei Tan, Sule Sahin, Dharshan Chandramohan, Philip Lee, Jasmine Hu, Luciano Vinas, and Elena Vasquez: the times spent brainstorming, working on class projects, and doing research together made this experience fun. Each of you are brilliant and kind, and I hope to continue working together in the future. I would also like to thank Cresini Tabaranza-David, Ana Maria Deluca, SarahJane Taylor, Kristin Olsen, Victoria Ross, and Elisa Yang for their administrative support over the years.

To my family in the Radiation Oncology Department at UCSF who have supported me since joining: Cathy Park, Emily Hirata, Tim Solberg, Olivier Morin, Martina Descovich, Alon Witztum, Gilmer Valdes, Sara St. James, Mary Feng, Felix Feng, Kavita Mishra, Sue Yom, Osama Mohamad, Tony Wong, Mekhail Anwar, Tomi Nano, Joey Cheung, Christina Calvin, Xin Miao, Nina Pitts, and the entire team. Balancing life as a medical physicist and PhD student was not easy and only possible through the collaborative environment fostered within our group.

Thank you to my former colleagues and friends at The University of Pennsylvania, where my career in medical physics began: Chris Ainsley, Kevin Teo, Michele Kim, Joe Dise, Dan Pinkham, Brendan Burgdorf, and Shannon O'Reilly.

Finally, I would like to thank my family for the unconditional love you have provided me over the years. Emma, for being my best friend and biggest cheerleader. Dad, for always supporting me and instilling in me the confidence to accomplish anything. Mom, for encouraging me to prioritize health and taking time to rest. Gul, for inspiring me to cherish each moment and live life to the fullest. Peter, for being both my biggest adventure and safest space. To my family who shaped my childhood: Marian, Trevor, Auntie Lindsay, Uncle Jamie, Katie, William, George, Uncle Andrew, Auntie Dawn, Lola, Nazli, Dervla, Nur, and Haldun. To my closest friends Anneka, Mari, Bobbi, Lauren, Stacey, Carly, Gina, Isha, Holly, Leanna, and Jess who always hold me up and give me joy.

I am so grateful for each of you!

Jess Scholey

December, 2021

Multimodal imaging and deep learning-based methods for improved dose calculation accuracy in photon and proton radiotherapy

Jessica Scholey

Abstract

Radiotherapy is one of the most common techniques used to treat cancer and is administered to over 60% of patients treated in the US. Computed tomography (CT) and magnetic resonance imaging (MRI) are powerful imaging modalities which have widespread applications in radiotherapy. Given the different underlying physics behind how image contrast is generated in kilovoltage CTs (kVCTs), megavoltage CTs (MVCTs) and MRIs, each has their own inherent strengths and limitations. The overarching goal of the research described in this dissertation is to use advanced image processing and deep learning methods to exploit the unique advantages of MRI, kVCT, and MVCT for improving dose calculation accuracy of photon and proton radiotherapy. Four primary projects are presented. The first includes the development of a novel multimodal imaging method used to estimate proton stopping power ratio using a combination of MRI and CT with results within 1% of physical measurements and were more accurate than the current clinical standard for the tissue types studied. The second includes the first work to synthesize MVCT images from MRI using a deep learning model in which the feasibility of utilizing MRI-derived synthetic CTs for radiotherapy treatment planning of head and neck cancer was demonstrated. The third is the first known application of deep learning to reduce uncertainty in the relationship between kVCT and electron density and stopping power ratio through learning of MVCT data, promising improvements for dose calculation accuracy. The synthetic images produced by the deep learning model were validated using results in tissue substitute phantoms. The fourth includes clinical considerations for implementing a commercial algorithm for producing MRI-derived synthetic CTs in a dataset of patients treated for prostate cancer. This study provides recommendations for optimizing the synthetic CT calibration curve for consistency with the users' clinically commissioned CT scanner.

Table of Contents

List of Figures

List of Tables

Abbreviations

A	Atomic mass
β	Particle velocity relative to speed of light
B_0	MRI main magnetic field
B_1	MRI radiofrequency field
c	Speed of light ($3.0\ x\ 10^8\ meters\ per\ second$)
CBCT	Conebeam computed tomography
CHNO	Carbon, hydrogen, nitrogen, oxygen
CNN	Convolutional neural network
CT	Computed tomography
DECT	Dual energy CT
DVH	Dose volume histogram
e	Electron charge ($1.602\ x\ 10^{-19}\ coulombs$)
ϵ_o	Vacuum permittivity ($8.854\ x\ 10^{-12}\ farads\ per\ meter$)
EBRT	External beam radiation therapy
GAN	Generative adversarial network
h	Hydrogen density
^1H	Proton density
H&N	Head and neck
HA	Hydroxyapatite ($Ca_5(PO_4)_3$)
HU	Hounsfield Unit
I_m	Mean ionization potential
ICRP	International Commission on Radiological Protection
ICRU	International Commission on Radiation Units and Measurements
IGRT	Image-guided radiation therapy
IMRT	Intensity-modulated radiation therapy
IRB	Institutional review board
kVCT	Kilovoltage computed tomography
Linac	Linear accelerator
MAE	Mean absolute error
m_e	Electron rest mass ($9.109\ x\ 10^{-31}\ kilograms$)
ME	Mean error
MRI	Magnetic resonance imaging

MVCT	Megavoltage computed tomography
OAR	Organ-at-risk
Org	Organic
ρ	Mass density
ρ_e	Relative electron density
PET	Positron emission tomography
PM	Physical measurement
PSNR	Peak signal-to-noise ratio
PTV	Planning target volume
QA	Quality assurance
ReLU	Rectified linear unit
RF	Radiofrequency
RMSE	Root mean square error
SBRT	Stereotactic body radiation therapy
sCT	Synthetic CT
sMVCT	Synthetic MVCT
SNR	Signal-to-noise ratio
SPR	Stopping power ratio
SSIM	Structural similarity index measure
T_1	Longitudinal relaxation time
T_2	Transverse relaxation time
TE	Echo time
TPS	Treatment planning system
TR	Repetition time
UC	Unified Composition
UTE	Ultrashort time-to-echo
μ	Linear attenuation coefficient
μ/ρ	Mass attenuation coefficient
VMAT	Volumetric modulated arc therapy
Z	Atomic number
ZTE	Zero-echo-time

Chapter 1 – Dissertation outline

In Chapter 2, I provide the background knowledge for this dissertation including the radiotherapy workflow and basic physics, dose calculation for photon and proton radiotherapy, and imaging with a focus on CT and MRI.

In Chapter 3, I provide details on data sources and tools used in subsequent chapters including an in-house created tissue substitute phantom, a commercial anthropomorphic head phantom, *in vivo* imaging datasets of patients treated for head and neck cancer and prostate cancer, CT calibration curves used to map Hounsfield Unit to electron density or stopping power ratio, and image evaluation and comparison metrics.

In Chapter 4, I describe the development of a novel multimodal imaging method used to estimate proton stopping power ratio using a combination of MRI and CT. Results of this method derived from imaging were compared to "ground truth" physical measurements calculated in tissue surrogate materials developed by our group. Results show that our multimodal imaging approach using MRI and MVCT provided results within 1% of physical measurements and were more accurate than the current clinical standard stoichiometric method for the tissue types studied.

In Chapter 5, I describe the first work to synthesize MVCT images from MRI using a deep learning model in which the feasibility of utilizing MRI-derived synthetic CTs for radiotherapy treatment planning of head and neck cancer was demonstrated. This work lays the foundation for a proposed paradigm shift whereby MRI is utilized for anatomical delineation while MRI-derived synthetic MVCT is used for radiotherapy dose calculation, which can all be performed on the treatment machine

and could provide substantial improvements to combined MRI-linear accelerator workflow and accessibility.

In Chapter 6, I describe work which was the first known application of deep learning to reduce uncertainty in the relationship between kVCT and electron density and stopping power ratio through learning of MVCT data, promising improvements for dose calculation accuracy. The synthetic images produced by the deep learning model were validated using results in tissue substitute phantoms.

In Chapter 7, I discuss clinical considerations for implementing synthetic images for treatment planning. In this work, I evaluated a commercial algorithm for producing MRI-derived synthetic CTs in a dataset of patients treated for prostate cancer using two common modalities including a conventional linear accelerator and the CyberKnife system. This study provides recommendations for optimizing the synthetic CT calibration curve for consistency with the users' clinically commissioned CT scanner and was the first to report on utilizing synthetic CT for treatment planning on CyberKnife.

In Chapter 8, I provide a summary of this dissertation and describe future research directions.

Chapter 2 – Background

2.1 Radiotherapy

Radiation therapy (or radiotherapy) is one of the most common therapies available for treating a wide range of cancers, with over 60% of cancer patients in the US treated curatively with radiation[1,2]. Since the discovery of X-rays by Wilhelm Conrad Röntgen in 1895, there have been significant technological advances allowing for increasingly complex and sophisticated delivery. In particular, the past 20 years have seen remarkable developments in computer technology, treatment planning software, dose calculation algorithms, and diagnostic and therapeutic delivery systems. These advancements have allowed for the emergence of techniques such as such as intensity-modulated radiation therapy (IMRT) and stereotactic body radiotherapy (SBRT), which allow clinicians to escalate dose delivered to tumor volumes while maintaining a sharp dose fall-off outside the tumor, minimizing toxicities to critical structures and normal tissues.

The journey of a typical patient treated with radiotherapy today begins with a cancer diagnosis, which is often made by a team of oncologists who determine radiation to be an appropriate course of treatment. After an initial consult in the radiation oncology department, the patient will receive an imaging scan (typically a computed tomography, or CT scan) which is used for three primary purposes: 1) to delineate (or "contour") the tumor and critical structures, 2) for radiotherapy treatment planning and dose calculation, and 3) as a reference image for patient alignment on the treatment machine prior to delivery. Contouring and treatment planning typically take about 1 week and radiotherapy delivery can take anywhere from ~1-7 weeks, depending on the treatment approach. This workflow is shown in Figure 2.1.

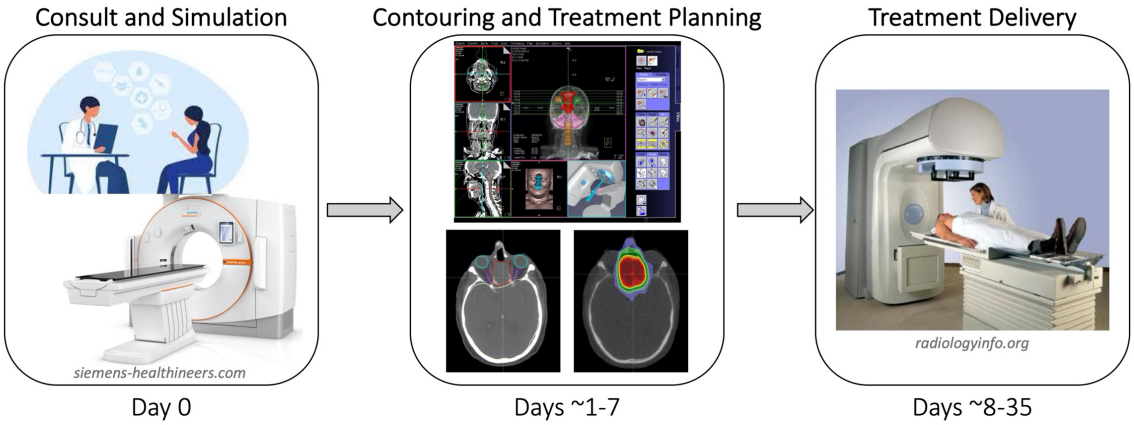

Figure 2.1: Radiotherapy workflow
Typical radiation oncology workflow consisting of consult and simulation appointments (left), contouring of anatomy and treatment planning (middle), and treatment delivery (right). During the simulation appointment, a CT scan is acquired and subsequently used for contouring anatomy and treatment planning. Prior to each treatment delivery, the patient is imaged for purposes of position set-up and verification using planar radiographs, CT, or MRI, depending on machine capabilities and treatment site.

2.2 Basic physics of radiotherapy

A basic understanding of photon interactions in matter is foundational to understanding how radiotherapy works. A comprehensive overview of these concepts can be found in several classic radiotherapy physics textbooks[3,4], with a brief description provided here. Whether considering photons used for imaging a patient (in the case of CT) or treatment (in the case of a linac), a photon beam is attenuated as it interacts with the material it is traversing via coherent scattering, the photoelectric effect, Compton scattering, or pair production. Coherent scattering (or classical scattering) occurs when a low-energy photon changes direction upon interacting with an atom with no energy transfer or deposition. The photoelectric effect occurs when a photon is absorbed by an atom, causing an atomic electron to be ejected. Compton scattering occurs when a photon interacts with an atomic electron, causing both the photon and electron to be scattered at angles that depend on the photon's incident

energy. Pair production occurs at energies higher than 1.02MeV when a photon interacts with the electromagnetic field of a nucleus, causing the production of an electron-positron pair. A photon is attenuated according to its linear attenuation coefficient (μ) which depends on the photon energy and the density of material being traversed. The mass attenuation coefficient (μ/ρ) is often used which depends on material atomic composition instead of density. The total mass attenuation coefficient ($(\mu/\rho)_{tot}$) is a sum of contributions from all photon interactions including coherent scattering ($(\sigma_{coh}/\rho)_{coh}$), photoelectric effect ($(\tau/\rho)_{P.E.}$), Compton scattering ($(\sigma_{C.S.}/\rho)_{C.S.}$), and pair production ($(\pi/\rho)_{P.P.}$) as shown in Equation 1.

$$(\mu/\rho)_{tot} = (\sigma_{coh}/\rho)_{coh} + (\tau/\rho)_{P.E.} + (\sigma/\rho)_{C.S.} + (\pi/\rho)_{P.P.} \quad (1)$$

Coherent scattering is only important at very low energies and high-Z materials and is usually omitted when considering therapeutic energies. Photoelectric, Compton scattering, and pair production phenomena are proportional to the photon energy (E), atomic number (Z), and electron density (ρ_e) through the following relationships:

$$(\tau/\rho)_{P.E.} \propto {Z^3}/{E^3} \quad (2)$$

$$(\sigma/\rho)_{C.S.} \propto \rho_e \quad (3)$$

$$(\pi/\rho)_{P.P.} \propto Z^2 \text{ (for energies up to ~20MeV)} \quad (4)$$

A plot of μ/ρ versus photon energy for water is shown in Figure 2.2 (values have been approximated for purposes of visualization). Of note, at lower photon energies typically used in diagnostic imaging (e.g. 70-140 kV), both photoelectric effect and Compton scattering contribute to photon attenuation, resulting in a dependence on material atomic number and energy. At higher energies typically used for therapy, (e.g. ~2-10MV), Compton scattering is the dominant photon interaction, resulting in a dependence only on material electron density. Mass attenuation coefficients for additional materials can be found in the literature[5].

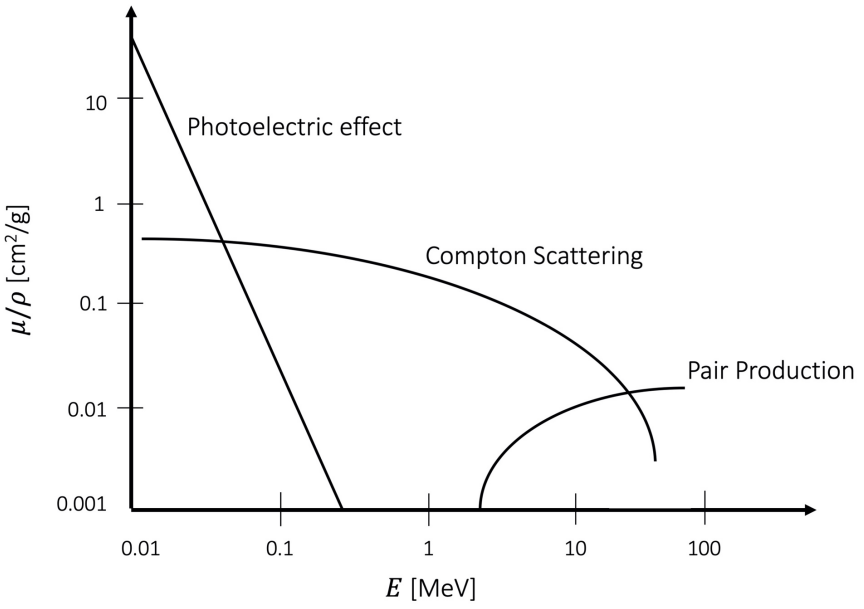

Figure 2.2: Mass attenuation coefficient in water
Mass attenuation coefficient versus photon energy for water (values have been approximated for purposes of visualization). At lower, diagnostic photon energies (e.g. 70-140 kV), both photoelectric effect and Compton scattering contribute to photon attenuation, resulting in a dependence on material atomic number and energy. At higher, therapeutic energies (e.g. ~2-10MV), Compton scattering is the dominant photon interaction, resulting in a dependence only on material electron density.

2.3 Dose calculation

Accurate dose calculation algorithms within a treatment planning system are essential for delivering high quality radiotherapy that agrees with the clinicians' intent. There are several in-house and commercial algorithms serving this purpose, each with varying levels of accuracy and computational requirements. A comprehensive review of these algorithms is outside the scope of this work and is available in the literature[6,7], however, a general overview of photon and proton external beam radiotherapy (EBRT) dose calculation is provided as foundation.

2.3.1 Photon radiotherapy

A dose calculation algorithm used for photon radiotherapy must model the energy deposition of a photon beam as it travels through a patient. In the case of EBRT, the therapeutic photon beam is typically produced by a linear accelerator (linac) that rotates around the patient to deliver radiation from different angles throughout treatment delivery. The majority of modern dose calculation algorithms approximate absorbed dose D at position $\vec{r}$ in a heterogeneous medium using the following integral:

$$D(\vec{r}) = \int \frac{\mu}{\rho} \Psi_p(\vec{r}') A(\vec{r} - \vec{r}') d^3\vec{r}' \quad (5)$$

where μ/ρ is the mass attenuation coefficient, $\Psi_p(\vec{r}')$ is the primary photon energy fluence produced by the linac, and $A(\vec{r} - \vec{r}')$ is the convolution kernel. Convolution kernels are matrices of dose distributions caused by scattered photons and electrons in water and are pre-calculated within the treatment planning system (TPS) using Monte Carlo. Material heterogeneities are accounted for by rescaling the dose spread kernel using the material density relative to water. In other words, the patient-specific parameter of interest for dose calculation in photon radiotherapy is the material mass density (ρ) or electron density relative to water (ρ_e), depending on the requirements of the TPS being used. In the clinic, density maps are estimated using a CT scan acquired for each patient and a CT calibration curve mapping Hounsfield Unit (HU) of each voxel to material density.

2.3.2 Proton radiotherapy

Using proton beams for therapeutic purposes was first proposed by Robert R. Wilson in 1946 and has since become an increasingly popular cancer treatment modality with the number of operational centers steadily increasing[8]. Unlike photons which interact through photoelectric, Compton, or pair production, protons interact through ionization and excitation caused by collisions between the protons and atomic electrons of the material being traversed and, to a lesser extent, radiative losses or nuclear

reactions. The rate of energy loss per unit length (known as the stopping power) is described by the Bethe-Bloch Equation:

$$-\frac{dE}{dx} = \frac{4\pi n z^2}{m_e c^2 \beta^2}\left(\frac{e^2}{4\pi\epsilon_o}\right)^2 \left(ln\left(\frac{2m_e c^2 \beta^2}{I(1-\beta^2)}\right) - \beta\right) \quad (6)$$

where n is the number of electrons in a given volume, z is particle charge (equal to 1 for protons), m_e is electron mass, c is the speed of light, β is the particle velocity relative to c, e is the electron charge, ϵ_o is the vacuum permittivity, and I is the mean ionization potential. Notably, a particle's stopping power is inversely proportional to the square of its velocity, giving rise to the distinct clinical advantage of depositing the majority of dose immediately before stopping at the distal end, as described by the characteristic Bragg Peak[9]. With appropriate choice of beam angles, full target coverage can be achieved while imparting negligible dose to healthy tissue beyond where the beam stops. Calculating the location of the Bragg Peak requires a patient-specific map of the ratio of stopping powers (or stopping power ratio, SPR) of the material being traversed relative to water. In the clinic, SPR maps are estimated using a CT scan acquired for each patient and a CT calibration curve mapping Hounsfield Unit (HU) of each voxel to material SPR.

2.4 Imaging in radiotherapy

Many different imaging modalities are used for diagnosis, treatment planning, and/or image guidance in radiotherapy including x-ray radiography, fluoroscopy, ultrasound, CT, magnetic resonance imaging (MRI), and positron emission tomography (PET), though the focus of this work is on CT and MRI. Prior to treatment planning and delivery, patients undergo a "simulation" appointment as shown in Figure 2.1. During simulation, a 3-dimensional (3D) image dataset is acquired with the patient in the same position used during treatment. This simulation scan is then used for two critical purposes: 1) all subsequent delineation of patient anatomy (normal and diseased tissues) and 2) calculation of radiation treatment plans and dose distributions. For the former, the optimal imaging modality depends on the

anatomy being visualized. For the later, the optimal imaging modality is CT because it provides a direct map of photon attenuation used for dose calculation and has therefore been the cornerstone of radiotherapy treatment planning.

2.4.1 CT

To acquire a CT image, a narrow beam of x-rays rotates around a patient while detectors mounted opposite the x-ray source measure photon attenuation. By taking a sufficient number of transmission measurements, a map of attenuation coefficients can be reconstructed[10,11]. During reconstruction, each axial plane is divided into small voxels where the attenuation at each voxel is known as a CT number. The CT number at each voxel is normalized to a corresponding HU value according to Equation 7, such that air is defined as -1000 HU and water is defined as 0 HU.

$$HU = \frac{\mu_{material} - \mu_{water}}{\mu_{water}} \, x \, 1000 \tag{7}$$

The most common photon spectra used for CT is in the 70-140kVp range (referred to as kVCT) where the photoelectric effect dominates and photon attenuation is proportional to Z^3, resulting in high tissue contrast favorable for visualizing anatomy. In radiotherapy, simulation scans are typically acquired at 120kVp. The relationships between HU and both ρ_e and SPR are often modeled as piecewise linear functions with two linear regions (for example, lower density as in lung, soft tissue, and higher density as in bone). A consequence of the explicit atomic number dependence in this relationship is that for a given HU there may be multiple possible solutions for electron density[12-14]. In other words, as different tissue types exhibit different mass densities, any solution to explicitly resolve electron density from kVCT HU necessitates an understanding of the atomic composition of the material. With only a standard single kVCT scan, there is not enough additional information to correct for this atomic number dependent effect. While a lesser concern for photon-based dose calculation where there is a smaller dosimetric impact of these uncertainties, even small inaccuracies in the translation of kVCT HU to

proton SPR can result in over- or under-dosing of the target volume and nearby organs at risk when using proton therapy[13,15]. When treating with proton therapy in clinical practice, these uncertainties are typically accounted for by adding an additional margin of 2.5-3.5% to each proton beam's range[16,17]. In efforts to address this issue, many groups have proposed solutions involving the incorporation of additional information provided by CT scans of differing energy, known as dual energy CT (DECT). This method considers that each CT scan taken at different energies will have differential responses/intensities due to different atomic numbers and density values, which can be used to solve for the effective atomic number and electron density. Results from prior studies[18-22] have shown great promise for high accuracy determination of both quantities. However, the necessity for two separate scans may limit the practically achievable accuracy[23] for DECT methods and recent studies[23-25] have shown that nominal noise and beam hardening artifacts may be of greater concern for DECT analysis methods compared with standard single energy kVCT techniques.

At higher x-ray energies (as in megavoltage or MV), the predominant mechanism of photon interaction is Compton scattering, which is directly dependent on the electron density of the medium rather than its atomic number. Therefore, HUs in MVCT images are directly dependent on electron density, with MVCT having demonstrated sufficient accuracy for dose calculation[26,27]. In bony tissue, for example, errors in estimating physical and relative electron density have been reduced from 3.6% and 2.6%, respectively, using kVCT to <1% using MVCT[28]. Furthermore, in contrast to DECT (and lower energy x-ray methods in general), higher energy x-rays are naturally less susceptible to beam hardening and photon starvation effects. With these advantages in terms of electron density linearity and resistance to beam hardening effects[29-33], MVCT imaging would seem to be more favorable than lower energy kVCT techniques for dose calculation. Yet key disadvantages such as lower image contrast, greater noise (due to increased scatter), and reduced availability (due to the potentially cost prohibitive higher energy x-ray accelerator requirements), have limited clinical implementation of

MVCT scans into the radiation oncology clinic to the special case of on-treatment image guidance and/or adaptation. Interest in overcoming these limitations has resulted in the development of technologies such as intermediate energy orthovoltage CT imaging[34] although, to date, no clinical system is available. In the case of proton therapy, proton CT has been investigated for providing more accurate maps between HU and SPR, and while proton CT would allow for a direct relationship between CT number and SPR[35,36] , its primary limitation is the high proton energy required to image a patient.[37].

2.4.2 MRI

Magnetic resonance imaging is a powerful imaging modality that utilizes the different magnetic properties of materials to generate image contrast. A clinical MRI system is composed of three main components: a large static magnetic field (typically 0.3-7 Tesla) used to create bulk magnetization of hydrogen nuclei, gradient coils used to induce small differences in phase, frequency, or applied magnetic field for distinguishing spatial and slice position of each nuclei, and a radiofrequency (RF) system used to transmit and/or receive signal-producing RF pulses. MRI pulse sequences are designed with optimal gradient and RF pulse settings (including flip angle, repetition time (TR), and echo time (TE)) to take advantage of differences in longitudinal (corresponding to T_1) and transverse (corresponding to T_2) relaxation times amongst different materials to produce image contrast. Detailed descriptions of the physics behind MRI can be found in the literature[11,38].

Over the past decade there has been a significant increase in the utilization of MRI in radiation oncology due to the improved soft tissue contrast provided relative to CT. MRI has been demonstrated to reduce inter- and intra-observer variability for contouring different disease sites[39-42] and used for both initial treatment planning and online adaptive radiotherapy through integration with linear accelerators[43]. In these contexts, the patient will typically receive both an MRI (for anatomical

visualization) and CT (for density information used in dose calculation). An MRI-only radiotherapy workflow would provide many advantages such as increased efficiency, reduction in the number of imaging scans required for each patient, and a decrease in image registration errors[44]. While MR images demonstrate favorable qualities for the delineation of many normal and diseased tissue types, they do not provide a direct map of photon attenuation, hindering its utility for dose calculation. This phenomenon can be easily demonstrated through examples of air and bone: air has very little density, minimally attenuates a photon beam, and appears dark on a CT image. Bone, on the other hand, is the densest of biological materials, significantly attenuates a photon beam, and appears bright on a CT image. However, despite their very different attenuation properties, both air and bone show up as dark signal voids in conventional MRI sequences due to low proton density (in the case of air and bone) and ultrashort transverse relaxation time (in bone). Techniques such as ultrashort time-to-echo (UTE) and zero-echo-time (ZTE) have been developed to image materials with very short transverse relaxation times, and while studies have indicated that these sequences have improved contrast in bone versus air, there are still instances of mislabeling[45] which would be unacceptable if used in radiotherapy dose calculation. These challenges in mapping MRI voxel values to photon attenuation have hindered the ability to use MRI for dose calculation.

The overarching goal of the research described in this dissertation is to use advanced imaging techniques to exploit the unique advantages of MRI, kVCT, and MVCT to improve the dose calculation accuracy of 3D image-based photon and proton EBRT. To reiterate: in Chapter 4, I describe the development of a novel method for more accurately estimating proton stopping power ratio using physical values estimated from a combination of MRI and CT. In Chapter 5, I describe the first work to synthesize MVCT images from MRI using a deep learning model, laying the foundation for a proposed paradigm shift whereby MRI is utilized for anatomical delineation while MRI-derived synthetic MVCT is used for radiotherapy dose calculation. In Chapter 6, I describe the first known

application of deep machine learning to reduce uncertainty in the relationship between kVCT and electron density and stopping power ratio through learning of MVCT data, promising improvements for dose calculation accuracy. Finally, in Chapter 7, I discuss clinical considerations for implementing synthetic images for treatment planning.

Chapter 3 – Data sources and tools

This chapter describes the datasets and tools referenced in subsequent chapters as outlined in Table 3.1.

Table 3.1: Datasets and tools
Datasets and tools used in the projects described in this dissertation.

	Chapters	Description
3.1 Tissue substitute phantom	4, 6	In-house phantom created to emulate the physical properties of skin, muscle, adipose, and spongiosa bone.
3.2 Anthropomorphic head phantom	6	Vendor-provided head phantom with tissue types meant to emulate soft tissue, inner bone, and outer bone.
3.3 *in vivo* head and neck dataset	5, 6	Paired T_1-weighted MRI, MVCT, and kVCT images for 120 patients treated for head and neck cancer.
3.4 *in vivo* prostate dataset	7	Paired MRI-derived synthetic CT and actual CT for 10 patients treated for prostate cancer.
3.5 CT calibration curves	4, 6	CT calibration curves generated to map CT Hounsfield Unit to relative electron density (3.5.1) or stopping power ratio (3.5.2).

3.1 Tissue substitute phantom

Phantoms are routinely used for quality assurance (QA) in radiation therapy and can be created in-house or acquired through a commercial vendor if available. Phantom characteristics depend on the specific application but should, in general, have homogeneous composition with physical properties that mimic biological tissues. Phantom materials can be considered mimicking of real tissues if they demonstrate reasonable equivalence with nominal physical parameters relevant to the modalities used in the study of interest. For example, in this work, tissue substitutes were considered reasonably equivalent to real biological tissues if their physical/compositional characteristics of interest for the respective modality were similar to each other including mass attenuation coefficients, water content, (^{1}H) proton density, and mass stopping power ratio.

3.1.1 Fabrication

Tissue substitute phantoms were created to mimic skin, muscle, adipose, and spongiosa bone based on each tissue's molecular composition provided in ICRU Report 44[46] and ICRP Report 23[47] by creating homogeneous mixtures of deionized water with appropriate ratios of gelatin (to mimic protein), coconut oil (to mimic fat), and/or pure hydroxyapatite (to mimic bone) as shown in Table 3.2.

Table 3.2: Tissue substitute composition
Tabulated compositions for tissue substitute phantoms created by measurements (by mass) of appropriate ratios of water, gelatin, coconut oil, and/or pure hydroxyapatite.

	Water	Protein	Fat	Hydroxyapatite	SDS
Skin	75.00	25.00	0	0	0
Muscle	74.78	19.97	5.0	0	0.25
Adipose	46.80	2.50	49.2	0	1.50
Spongiosa	26.61	11.83	47.43	12.81	1.32

Constituents of each mixture were carefully weighed using a high-precision scale (Practum313-1S, Sartorius Biotech, Germany). Skin phantoms were created by mixing fixed (by mass) amounts of gelatin with hot water until dissolved and allowing for solidification. Muscle and adipose phantoms were created similarly but including coconut oil (adipose solutions were created separately for each container to prevent congealing of the predominately oil mixture). Spongiosa bone phantoms were created similarly to skin phantoms but including hydroxyapatite powder. Phantoms containing coconut oil included a small amount of detergent/surfactant (SDS) to encourage mixing of oil with water which improved overall homogeneity. Measurements of physical density for each phantom were calculated from volume and mass determined using volumetric pipettes and the high-precision scale. As gelatin and coconut oil are not pure molecular compounds (but rather a mixture of multiple molecules), chemical composition of these substances was determined via carbon, hydrogen, nitrogen, and oxygen (CHNO) combustion analysis at a specialized microanalytic facility. Elemental composition and measured mass density were used to calculate ground truth ρ_e (Equation 8) and SPR (Equation 9: the

ratio of the Bethe-Bloch equation for a material relative to water and Equation 10: the Bragg Additivity Rule) for each tissue type.

$$\rho_e = \frac{\rho_m}{\rho_{H2O}} \frac{\left(\sum_i \frac{w_i Z_i}{A_i}\right)_m}{\left(\sum_i \frac{w_i Z_i}{A_i}\right)_{H2O}} \tag{8}$$

$$\mathrm{SPR} = \rho_e \frac{\ln[2m_e c^2 \beta^2 / I_m (1-\beta^2)] - \beta^2}{\ln[2m_e c^2 \beta^2 / I_{H2O} (1-\beta^2)] - \beta^2} \tag{9}$$

$$\ln I_m = \left(\sum_i \frac{w_i Z_i}{A_i} \ln I_i\right) / \left(\sum_i \frac{w_i Z_i}{A_i}\right) \tag{10}$$

In Equations 8-10, for each element i composing material m, w is the fraction by weight, Z is the atomic number, A is the atomic mass, I is the mean ionization potential, ρ_e is the electron density of the medium relative to water, ρ is the nominal mass density, m_e is electron mass, c is the speed of light, and β is the proton velocity relative to c. All calculations were performed using a fixed proton energy of 115MeV to minimize consequences of energy variability in practice[13].

The final phantom was created from nine identical containers of each tissue substitute material along with calibration materials used for the MRI (^{1}H) proton density calibration curve described in Chapter 4. Identical containers of each tissue type and calibration material were placed in various locations at each level (top, middle, and bottom) of the phantom, resulting in a phantom consisting of 57 total containers. An axial slice of the middle of the phantom is shown in Figure 3.1.

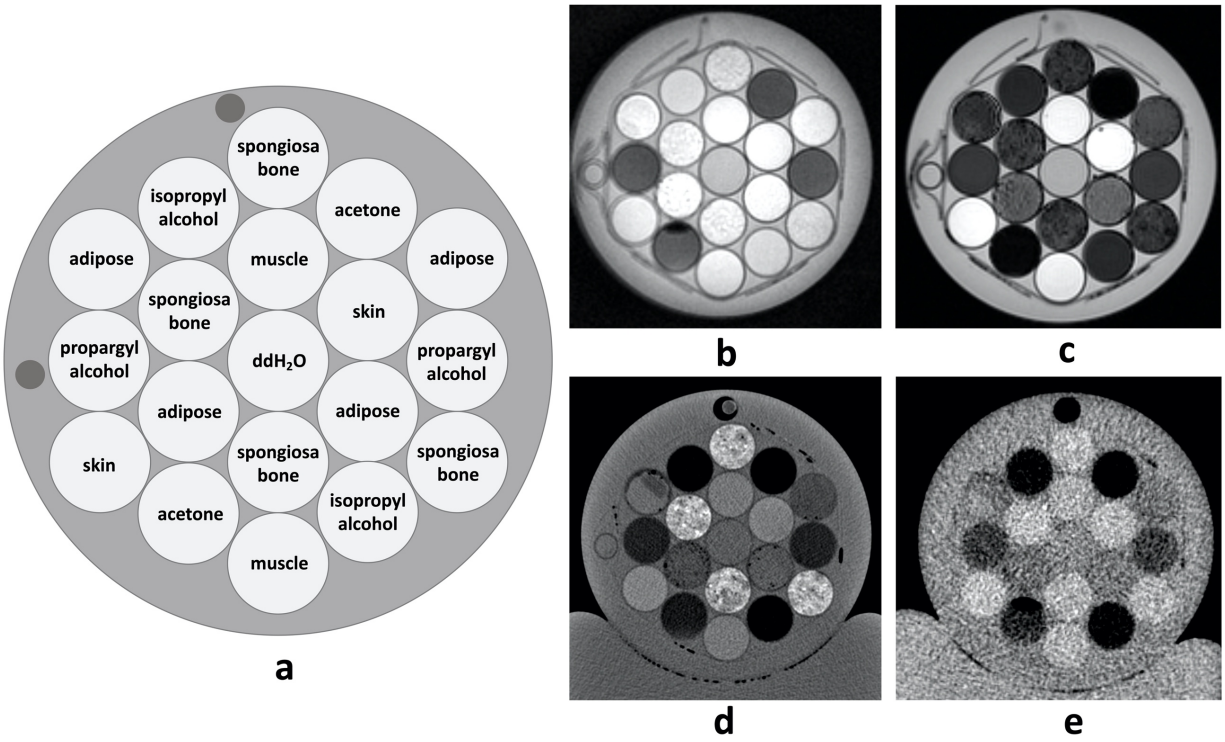

Figure 3.1: Tissue substitute phantom
Axial representation of a) example phantom configuration and corresponding images of b) ZTE (1H) proton density-weighted MRI, c) Dixon water-only MRI, d) kVCT, and e) MVCT. Each image depicts a single slice at the same level.

Image average intensity values were extracted using MIM Software (MIM Software Inc, v 6.8.3 Cleveland, OH) using regions of interest (ROIs) large enough to encapsulate all material within each container while minimizing volume averaging of the container itself.

3.1.2 Physical properties

To illustrate that tissue substitute materials were mimicking of real tissues, mass attenuation coefficients were computed for each phantom material from the atomic composition of each tissue substitute and elemental x-ray mass attenuation coefficients provided by NIST[48] assuming a single mono-energetic nominal mean energy of 80 keV for kVCT and 800 keV for MVCT. Proton mass stopping power ratios were computed for each phantom material from the atomic composition of each

tissue substitute assuming fixed proton energy of 115MeV. Literature values for tissue mass attenuation coefficients and mass stopping power ratios were derived from tissue compositions provided in ICRU 44[46]. These parameters are shown for phantom and literature tissues in Table 3.3.

Table 3.3: Tissue substitute physical parameters
Tabulated physical parameters for mass attenuation coefficients at mean energies of kVCT (80 keV) and MVCT (800 keV), water content, hydrogen density, and proton mass stopping power ratios derived from tissue substitute phantom elemental composition versus real tissue composition from ICRU Report 44[46] and ICRP Report 23[47].

	μ/ρ (80 keV) $\times 10^{-1}$		μ/ρ (800 keV) $\times 10^{-2}$		Water content		Hydrogen density		SPR/ρ (115 MeV)	
	Phantom	Literature	Phantom	Literature	Phantom	Literature	Phantom	Literature	Phantom	Literature
Skin	1.81	1.81	7.78	7.80	0.75	0.59-0.72	0.100	0.100-0.101	0.99	0.99-1.00
Muscle	1.81	1.82	7.80	7.79	0.75	0.70-0.79	0.102	0.101-0.102	0.99	0.99-1.00
Adipose	1.82	1.80	7.88	7.87	0.47	0.11-0.31	0.115	0.112-0.116	1.00	1.00-1.03
Spongiosa	1.90	1.95	7.75	7.67	0.27	0.23	0.096	0.085	0.98	0.98

Additional metrics were assessed to demonstrate reasonable tissue substitute equivalence with biological tissue composition provided in the literature[46,47,49] including affective atomic number (Z_{eff}), physical density, and relative electron density which are shown in Table 3.4.

Table 3.4: Tissue substitute density values
Tabulated physical parameters for effective atomic number, physical density, and relative electron density derived from tissue substitute phantom elemental composition versus real tissue composition from ICRU Report 44[46] and ICRP Report 23[47].

	Z_{eff}		Physical density (g/cc)		Relative electron density	
	Phantom	Literature	Phantom	Literature	Phantom	Literature
Skin	7.46	7.51	1.06	1.09	1.05	1.07
Muscle	7.45	7.70	1.05	1.05	1.04	1.04
Adipose	7.03	6.56	0.95	0.95	0.96	0.95
Spongiosa	9.20	10.05	1.06	1.14	1.04	1.12

Physical measurement (PM) values for each parameter used to compute SPR were calculated directly using percentage (by mass) and elemental composition of each constituent. Uncertainties in PM values

were determined from the errors in mass measurements of the phantom components and propagated through each value used to compute SPR.

3.2. Anthropomorphic head phantom

A commercially-available anthropomorphic head phantom (RS-108, Radiologic Support Devices Inc., Long Beach, CA) was used to validate relative electron density (ρ_e) and SPR values of kVCT, MVCT, and synthetic MVCT images produced by the CycleGAN model described in Chapter 6. A 10mg sample of phantom bone material easily accessible inside the skull was extracted and sent for molecular composition determination using CHNO combustion analysis and energy dispersive x-ray spectroscopy. Because the deep learning model was trained on *in vivo* human data, phantom materials in this study were selected based on their similarity to values of *in vivo* tissues provided in the literature[46,47,50]. Elemental composition, physical density, relative electron density (calculated from Equation 8), and stopping power ratio (calculated from Equations 9 and 10) were compared between phantom tissues and literature *in vivo* human tissues to ensure the phantom was a reasonable surrogate. Of the phantom materials, bone closely matched composition, density, and SPR values from the literature (as shown in Table 3.5) and was selected for analysis. Other materials meant to generically mimic tissues in this head phantom did not mimic the typical mass density or elemental compositions found in human tissues[50] and were not further investigated.

Table 3.5: Anthropomorphic head phantom characteristics
Tabulated physical parameters for atomic composition, physical density, relative electron density, and stopping power ratio derived from physical measurements of the bone phantom material versus *in vivo* bone tissue composition from ICRU Report 44[46] and ICRP Report 23[47].

	Phantom Physical Measurements				Literature Values			
	% composition	Density (g/cc)	ρ_e	SPR	% composition	Density (g/cc)	ρ_e	SPR
Bone	H: 6 C: 46 N: 4 O: 31 Ca: 12 Trace < 1	1.18	1.12	1.13	H: 5-9 C: 40-60 N: 2-4 O: 30-45 Ca: 8-20	1.14 - 1.2	1.12 - 1.18	1.12 - 1.18

3.3 *in vivo* head and neck dataset

A dataset of kVCT, MVCT, and T_1-weighted (non-contrast) MR images for 120 patients treated for head and neck (H&N) cancer on an Accuray TomoHD system (TomoHD; Accuray, Sunnyvale, CA) were retrospectively recruited for studies described in Chapters 5 and 6. These studies and experimental protocols were approved under by the University of California, San Francisco Institutional Review Board (IRB) #14-15452 and carried out in accordance with relevant guidelines and regulations. Informed consent was waived by the IRB due to the retrospective, de-identified nature of the studies.

Each patient received a kVCT simulation scan in the head-first supine treatment position using a 5-point head and shoulder thermoplastic mask and tongue stent approximately two weeks prior to treatment using a Siemens Somatom CT scanner (Somatom; Siemens, Munich). kVCT scans were acquired at 120 kVp photon energy using a helical acquisition with reconstructed slices of voxel size $1x1x3mm^3$ (512 x 512 matrix). Immediately prior to treatment delivery, each patient received an MVCT scan in the head-first supine treatment position using a 5-point head and shoulder thermoplastic mask and tongue stent for purposes of image alignment on the TomoHD system. The MVCT scan closest in time to their kVCT simulation scan was included in this study to minimize differences in patient anatomy through potential changes in weight gain or loss and tumor growth or shrinkage. MVCT scans (3.5MV photon spectra) were acquired using the "coarse" pitch mode with reconstructed slices of voxel size $1x1x3mm^3$ (512 x 512 matrix). Diagnostic MRI scans were acquired in the head-first supine position approximately 0-6 weeks prior to treatment on 1.5 Tesla scanners using a T_1-weighted 3D sequence without contrast (TR 500-800ms, TE 7-15ms, flip angle 70-110^o, voxel size $1x1x3mm^3$).

For all patients, kVCT and MVCT scan volumes were co-registered to the space of their respective MRI scans using the automated 3D rigid registration algorithm in MIM to produce a match across the entire scan volume. CTs were cropped to match the MRI field-of-view and resampled to a

uniform 512x512 matrix. All MR images were histogram-matched to standardize intensity values of images acquired across different scanners[51].

3.4 *in vivo* prostate dataset

A dataset of CT and MRI-derived synthetic CT (sCT) images were acquired for 10 patients being treated with SBRT to the prostate who were prospectively recruited for the study referenced in Chapter 7. This study and its experimental protocols were approved under UCSF IRB #18-24578. Patients underwent informed consent prior to image acquisition. Same-day CT and MRI simulation in the head-first supine treatment position were performed on each patient. kVCT scans were acquired using a Siemens Somatom scanner at 120kVp photon energy using a helical acquisition with reconstructed slices of voxel size 1x1x3mm^3 (512 x 512 matrix). MRI simulation scans were acquired on a 3 Tesla scanner (MAGNETOM Vida, Siemens, Munich) using a radiotherapy-dedicated indexed couch overlay and large 18-channel UltraFlex coil (Siemens, Munich) suspended on a coil bridge to prevent deformation of the patients' surface caused by the coil. Dixon in-phase and out-of-phase images were acquired using a 3D T_1-weighted gradient echo sequence (TE_1/TE_2/TR=1.23/2.46/4ms, readout bandwidth=1090 Hz/pixel, 1x1x1.5mm^3, scan time 5 minutes). sCT datasets were generated using a vendor-provided bulk-assignment algorithm based on in-phase, out-of-phase, fat-only, and water-only MRIs. CT and sCT datasets were co-registered in MIM using an automated 3D rigid registration algorithm.

3.5 CT calibration curves

For the purposes of generating CT calibration curves, a CIRS (180 mm diameter) tissue characterization phantom (Model 062; CIRS, Norfolk, VA) containing nine tissue-surrogate plugs was scanned with kVCT and MVCT imaging modalities using the same imaging parameters as for *in vivo*

H&N data described in Chapter 3.3. This diameter phantom was chosen to better mimic scatter and beam hardening effects for the head and neck region studied.

3.5.1 Relative mass and electron density

Calibration curves relating CT HU with nominal mass density and CT HU with relative electron density were generated for the kVCT and MVCT scanners used. Batch number specific mass density (ρ) of each tissue surrogate plug was provided by the phantom vendor and used to generate the HU to ρ calibration curves. Relative electron density (ρ_e) was calculated for each type of tissue surrogate material using vendor provided batch number specific elemental compositions and mass density in Equation 8. HU to ρ and HU to ρ_e calibration curves are shown in Figure 3.2 and were used in subsequent HU to ρ or ρ_e conversions.

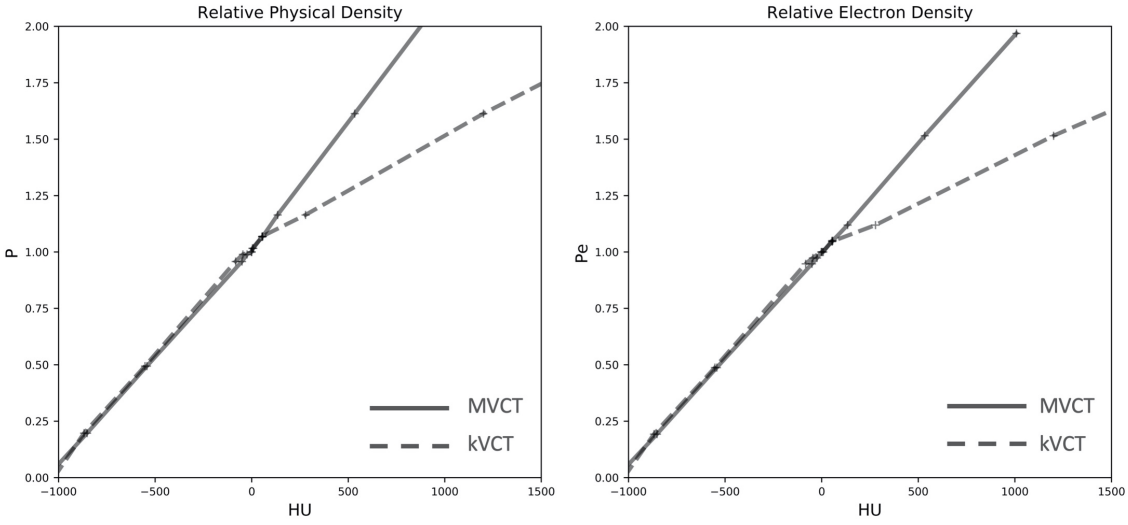

Figure 3.2. Hounsfield Unit to electron density calibration curves
Calibration curves of HU for kV (dashed line) and MV (solid line) CT plotted against relative (to water) physical (left) and electron (right) density. Curves were derived from a CIRS tissue surrogate phantom with various tissue mimicking materials and liquid water. Note the typical bilinear distribution for ρ and ρ_e mapped using lower (kV) energies, contrasted by the linear distribution at higher (MV) energies due to the negligible effects of the photoelectric effect at MV energies.

3.5.2 Stopping power ratio (stoichiometric calibration)

Calibration curves relating CT HU and SPR were performed for the kVCT and MVCT scanners used in this work according to the stoichiometric method by Schneider *et al*[52]. Elemental composition for each plug provided by the vendor was used with corresponding atomic numbers, weights, and mean ionization potential*[53,54] to determine $\tilde{Z}_i$ and $\hat{Z}_i$ for element i as described[52]. Scanner-specific coefficients $K^{P.E.}$, K^{coh}, $K^{K.N.}$ characterizing the photoelectric effect, coherent scattering, and Klein-Nishina (Compton scattering) cross-sections, respectively, were computed from a linear regression fit. With these coefficients, HU values were calculated for 64 human tissues provided in the literature[47,55] and used to create a plot of tissue HU versus SPR as shown in Figure 3.3. Each reference tissue was classified as organ-like (shown in red), fat-like (shown in green), and bone-like (shown in blue), with three separate plots generate: one for each tissue type. The final stoichiometric curve was pieced together extending over a range of typical HU values for biological tissue from the separate fits with nodes at HU values of -200, -120, -20, +35, +100, and +140 as previously described[14].

*The work described in Chapter 4 used *Im* values published by Janni at al. in 1982[53]. Since the publication of that manuscript[56], updated *Im* values optimized for use with the Bragg additivity rule were published by Bär et al.[54] and used for all subsequent work (e.g. in Chapter 6). The kVCT and MVCT stoichiometric calibration curves shown in Figure 3.4 were created using values of Bär et al.

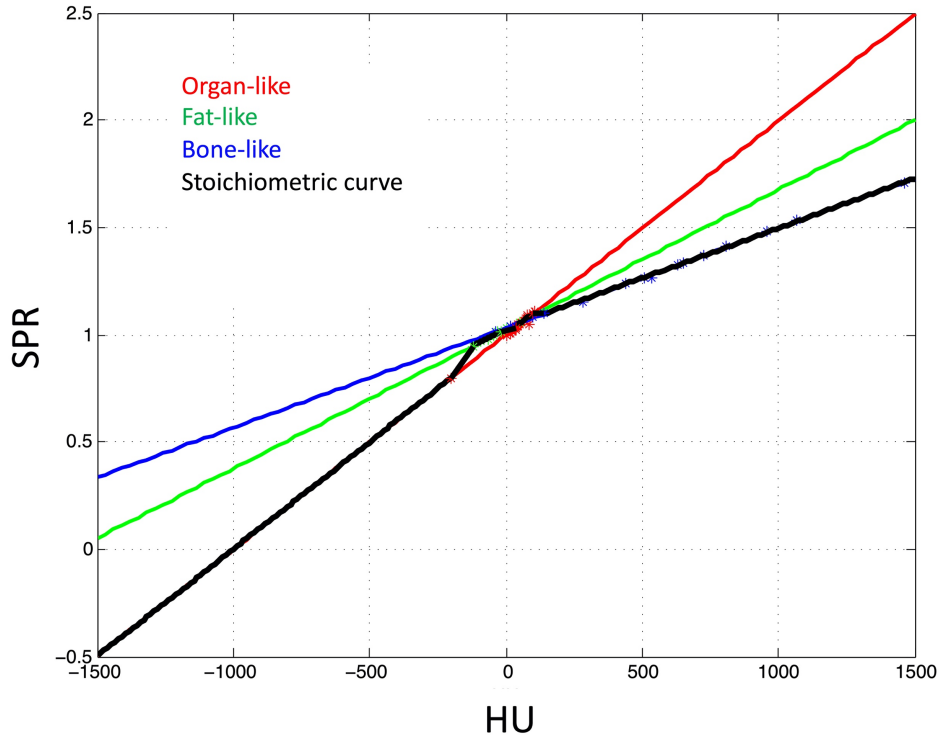

Figure 3.3: Stoichiometric calibration curve
A stoichiometric calibration curve (black line) which provides SPR from HU value was created for the kVCT scanner used in this study using the CIRS tissue surrogate phantom and parametric fits to literature values of organ-like (red line), fat-like (green line), and bone-like (blue line) human tissues.

In total, two stoichiometric CT calibrations were performed: one for the kVCT scanner used in this work (120kVp, 1mm slice thickness) and the other for the MVCT scanner used in this work (3.5MV, 1mm slice thickness). Details regarding stoichiometric calibrations for kV and MV energies are described in the literature[57]. All SPR calculations were performed using a fixed proton energy of 115MeV. HU to SPR calibration curves are shown in Figure 3.4 and were used in subsequent HU to SPR conversions.

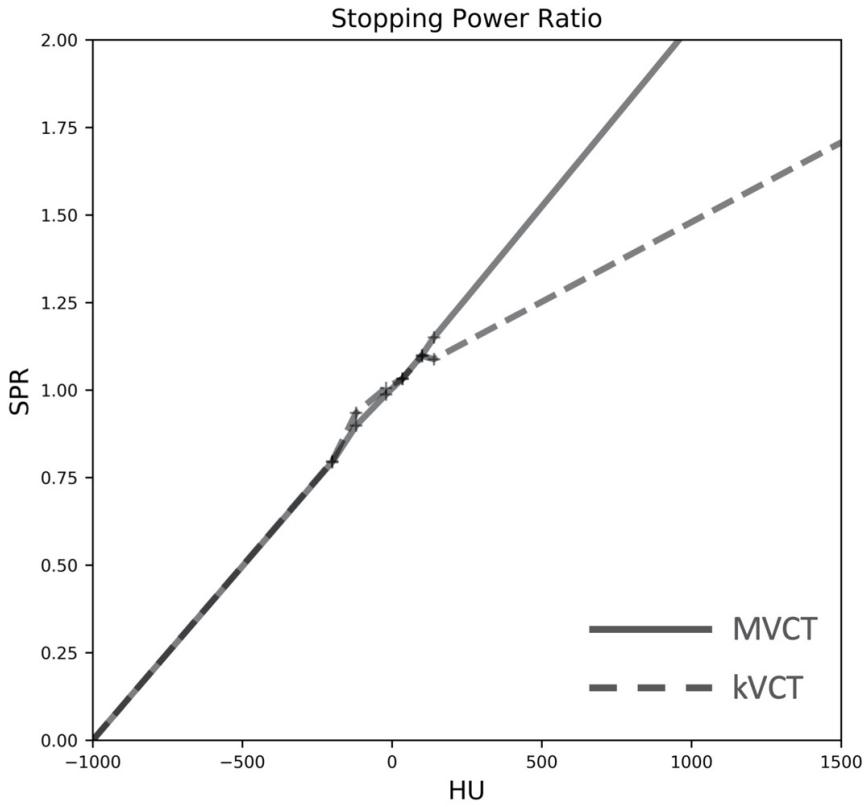

Figure 3.4. Hounsfield Unit to stopping power ratio calibration curve HU plotted against SPR derived from kV (dashed line) and MV (solid line) CT scans. All curves were derived from a CIRS tissue surrogate phantom with various tissue mimicking materials and liquid water. Note the typical bilinear distribution for SPR mapped using lower (kV) energies, contrasted by the linear distribution at higher (MV) energies.

For both ρ_e (Figure 3.2) and SPR (Figure 3.4) calibration curves, it is noted that there is a bilinear distribution for kVCT. On the other hand, MVCT calibration curves are linear with ρ_e and SPR. This energy dependent trend in CT imaging HU vs electron density is in line with the finding by Das *et al*[58] where they found that as photon energy increases, the relationship between HU and electron density becomes more linear. In addition, this relationship is fairly stable in the regime of lung to soft tissue ($\text{HU} \leq 100$) but becomes highly dependent on energy and material composition at higher electron density values as in bone ($\text{HU} > 100$).

3.6 Image evaluation metrics

There are several quantitative metrics commonly used to assess image quality. For example, the signal-to-noise ratio (SNR) of an image is a measure of desired signal relative to the level of background noise and peak-signal-to-noise ratio (PSNR) is a measure of the maximum signal relative to the level of background noise. There are also quantitative metrics used to compare one image to another. For example, mean error (ME) is the average error between a pixel (2D) or voxel (3D) in one image to its corresponding value in the comparison image. Similarly, mean absolute error (MAE) is the average absolute error value of pixel/voxel values between images. Root mean square error (RMSE) is the standard deviation of the residuals and is commonly used as a measure of accuracy. Finally, structural similarity index measure (SSIM) is a measure for quantifying similarity of structural information and spatial distribution between images.

Mathematically, the mean error metrics to assess differences between imaging datasets can be calculated at the pixel or voxel level. Letting N denote the total number of voxels shared in common between source X and target Y and X_i, Y_i denote the intensity values of the i^{th} voxel of the respective source and target then ME and MAE can be expressed as

$$\text{ME} = \frac{1}{N}\sum_{i=1}^{N} Y_i - X_i \tag{11}$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} |Y_i - X_i| \tag{12}$$

SNR, PSNR, RMSE, and SSIM are defined as

$$\text{SNR} = \frac{1}{K}\sum_{k=1}^{K} \frac{\mu_X^{(k)}}{\sigma^{(k)}} \tag{13}$$

$$\text{PSNR} = \frac{1}{K}\sum_{k=1}^{K} 20\log_{10}\left(\frac{I_{max}}{MSE^{(k)}}\right), \text{ where}$$

$$MSE = \frac{1}{N_k}\sum_{i=1}^{N_k}\left(Y_i^{(k)} - X_i^{(k)}\right)^2 \text{ and } I_{max} = 2^{12} - 1 \tag{14}$$

$$RMSE = \sqrt{\frac{1}{N_k}\sum_{i=1}^{N_k}\left(Y_i^{(k)} - X_i^{(k)}\right)^2} \tag{15}$$

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{16}$$

where $\mu_X^{(k)}$ is the mean of source X at slice k and $\mu_Y^{(k)}$ is the mirror case for another source Y such that the noise on slice k is given by the variable, $\sigma^{(k)}$, σ_{xy} is the covariance of x and y, $c_1 = (k_1 I_{max})^2$, $c_2 = (k_2 I_{max})^2$, $k_1 = 0.01$, and $k_2 = 0.03$.

Chapter 4 – MRI + CT for high accuracy SPR

Proton therapy is becoming an increasingly popular cancer treatment modality due to the proton's physical advantage in that it deposits the majority of its energy at the distal end of its track where the tumor is located. Range estimation errors caused by the difference in kVCT x-rays and proton interactions in matter can preclude the ability to utilize protons to their full potential. Applications of MRI in radiotherapy have increased over the past decade and using MRI to calculate SPR directly could provide numerous advantages. The purpose of this study was to develop a practical implementation of a novel multimodal imaging method for estimating SPR and compare the results of this method to physical measurements in which values were computed directly using tissue substitute materials fabricated to mimic skin, muscle, adipose, and spongiosa bone. For both the multimodal imaging method and physical measurements, SPR was calculated using the Bethe Bloch equation from values of relative electron density and mean ionization potential determined for each tissue. Parameters used to estimate SPR using the multimodal imaging method were extracted from Dixon water-only and (^{1}H) proton density-weighted zero echo time MRI sequences and CT, with both kVCT and MVCT used separately to evaluate the performance of each. For comparison, SPR was also computed from kVCT using the stoichiometric method, the current clinical standard. Results showed that our multimodal imaging approach using MRI with either kVCT or MVCT was in close agreement to SPR calculated from physical measurements for the four tissue substitutes evaluated. SPR values estimated using our method for MRI and MVCT were within 1% of physical measurements and of comparable or superior accuracy to the kVCT stoichiometric method for the four phantom tissue types studied.

4.1 Unified Compositions Model

There has been increasing interest in using MRI in radiotherapy due to its superior soft tissue contrast compared to CT, and could potentially offer distinct advantages for proton dose calculation if able to

directly provide SPR. A theoretical framework, named the Unified Compositions (UC) model, for achieving this was described by Sudhyadhom[59] in which a parametrized model was introduced to estimate the mean ionization potential (I_m) in human biological tissue. Using this framework, a tissue is categorized into three components, namely water, organic tissue, and mineralized bone with their relative percentages determined predominantly through MRI. The goal of this work was to outline and implement our multimodal imaging framework for the calculation of SPR at the voxel level through the estimation of relative electron density and mean ionization potential. In this study, I_m was calculated using the UC method, derived from the BAR as described below, and ρ_e was determined from CT imaging.

In the UC Model, human biological tissue is modeled using three general components: water, organic, and mineralized tissues. Under this assumption, the mean ionization potential can be computed on a per-voxel basis for the total material composing each voxel by translating Equation 10 to:

$$\ln(\mathrm{I_{voxel}}) = \frac{\left(\frac{\mathrm{w_{H_2O}Z_{H_2O}}}{\mathrm{A_{H_2O}}}\ln(\mathrm{I_{H_2O}})\right)+\left(\sum_{\mathrm{org}}\frac{\mathrm{w_{org}Z_{org}}}{\mathrm{A_{org}}}\ln(\mathrm{I_{org}})\right)+\left(\frac{\mathrm{w_{HA}Z_{HA}}}{\mathrm{A_{HA}}}\ln(\mathrm{I_{HA}})\right)}{\sum_{\mathrm{total}}\frac{\mathrm{w_{total}Z_{total}}}{\mathrm{A_{total}}}} \quad (17)$$

where w, Z, A, and I are parameters for water (H_2O), organic (org), and mineralized (hydroxyapatite (HA)) tissues.

In the original UC Model work, Equation 17 was simplified by assuming Z/A is identical for all biological molecules in the body, at the expense of a slight loss of accuracy, particularly in situations involving mixtures of water (Z_{H2O}/A_{H2O}=0.56) with bony tissue (Z_{HA}/A_{HA}=0.50). In this work, we were able to retain this accuracy by computing Z/A separately for the three constituents of biological tissue such that Z_{H2O}/A_{H2O} and Z_{HA}/A_{HA} are constants defined above while Z_{org}/A_{org} and Z_{total}/A_{total} were determined through a linear fit to their corresponding proton (hydrogen) densities. Using tabulated values of Z/A and the organic constituent hydrogen density by mass (h_{org}) for organic materials composing human biological tissues[49,60], Z_{org}/A_{org} was expressed as a function of h_{org} which can be

determined through imaging as described below. Similarly, values of Z_{total}/A_{total} for all molecules considered in human biological tissue was expressed as a function of total hydrogen density by mass (h_{total}), which can also be determined through imaging. Expressions for Z_{total}/A_{total} and Z_{org}/A_{org} as functions of h_{total} and h_{org}, respectively, are shown in Equation 18:

$$\frac{Z_{tot}}{A_{tot}} = 0.502h_{tot} + 0.499; \qquad \frac{Z_{org}}{A_{org}} = 0.490h_{org} + 0.500 \tag{18}$$

Mean ionization potential values for the three constituents of biological tissue were found using the same methodology as in the original work. I_{H2O} and I_{HA} are constants (75.3eV and 156.2eV, respectively[60]) while I_{org} was given through an exponential relationship between I_{org} and organic molecule hydrogen density by mass:

$$\mathrm{I_{org}} = \mathrm{A} \cdot \exp\left(\mathrm{Bh_{org}}\right) \tag{19}$$

The constants A and B are 93.23 and -3.47, respectively, while h_{org} is determined through imaging as described in section 4.2. Under these conditions, and assuming $\sum_i w_i = 1$, the mean ionization potential per voxel is reduced to the following expression:

$$\ln(\mathrm{I_{voxel}}) = \frac{\left(\mathrm{w_{H_2O}}\frac{\mathrm{Z_{H_2O}}}{\mathrm{A_{H_2O}}}\ln\left(\mathrm{I_{H_2O}}\right)\right)+\left(1-\left(\mathrm{w_{H_2O}}+\mathrm{w_{HA}}\right)\right)\left(\ln(\mathrm{A})-\mathrm{B}\cdot\mathrm{h_{org}}\right)\left(\frac{\mathrm{Z_{org}}}{\mathrm{A_{org}}}\right)+\left(\mathrm{w_{HA}}\frac{\mathrm{Z_{HA}}}{\mathrm{A_{HA}}}\ln\left(\mathrm{I_{HA}}\right)\right)}{\frac{\mathrm{Z_{total}}}{\mathrm{A_{total}}}} \tag{20}$$

All variables in Equation 20 for computing I_m per voxel can be determined a priori except for w_{H2O}, w_{HA}, h_{org} (used to compute Z_{org}/A_{org}) and h_{total} (used to compute Z_{total}/A_{total}), which can all be calculated using the proposed multimodal imaging framework.

4.2 Multimodal imaging framework

Percent water w_{H2O}

The Dixon technique used is a simple gradient echo MRI sequence that has been well described for separating water and fat ^{1}H (proton) signals[61] by exploiting the chemical shift between resonant frequencies of fat and water in a magnetic field, using in-phase and out-of-phase signals to produce

water-only and fat-only images. In theory, voxels in water-only images containing higher water content will produce higher signal. In this study, percent water was determined by taking the ratio of a voxel intensity value to the voxel of highest global intensity value.

Physical density and relative electron density

In the proposed framework, physical density is only used to determine percent hydroxyapatite (which is then used to calculate I_m) and relative electron density is only used to calculate SPR. A calibration curve was created relating HU values to physical density and relative electron density using the CIRS tissue surrogate phantom with known density values. Although measurements of ρ_e are typically derived from kVCT for treatment planning purposes, MVCT may be superior to kVCT for determination of electron density because MCVT HU values are largely independent of atomic number due to Compton scattering being the predominant mode of photon interaction. For this study, both kVCT and MVCT were used to assess and compare the accuracy of each. Physical and relative electron density calibration curves derived for both kVCT and MVCT are shown in Figure 3.2.

Percent hydroxyapatite w_{HA}

In this study, CT imaging was used to determine w_{HA} using an assumed relationship between physical density and calcium content derived from bone composition models as outlined by Zhou *et al.*[62]. Percent HA was determined using the fixed ratio of calcium mass to hydroxyapatite mass (39.9%). Percent calcium was determined empirically from physical density estimated from CT determined by a previous group, as shown in Equation 21[62].

$$Ca\,\% = -4.796\rho^3 + 7.761\rho^2 + 32.051\rho - 33.934 \quad (21)$$

Proton (hydrogen) density h_{org} and h_{total}

A ZTE gradient-echo MRI sequence[63] with $TE \ll T_2$, small flip angle, and short TR is (^{1}H) proton density weighted and can provide a reasonable approximation of the voxel total hydrogen content, with higher image intensity values corresponding to higher total hydrogen content. Values of h_{total} were

determined by dividing the total hydrogen content per voxel (in units of volume) derived from ZTE MRI by the physical density derived from CT, converting h_{total} to per unit mass. Materials of known hydrogen density including water, acetone, 92% (by volume) isopropyl alcohol, and propargyl alcohol, were used to create a calibration curve relating image intensity to voxel total (^{1}H) proton content as shown in Figure 4.1[59].

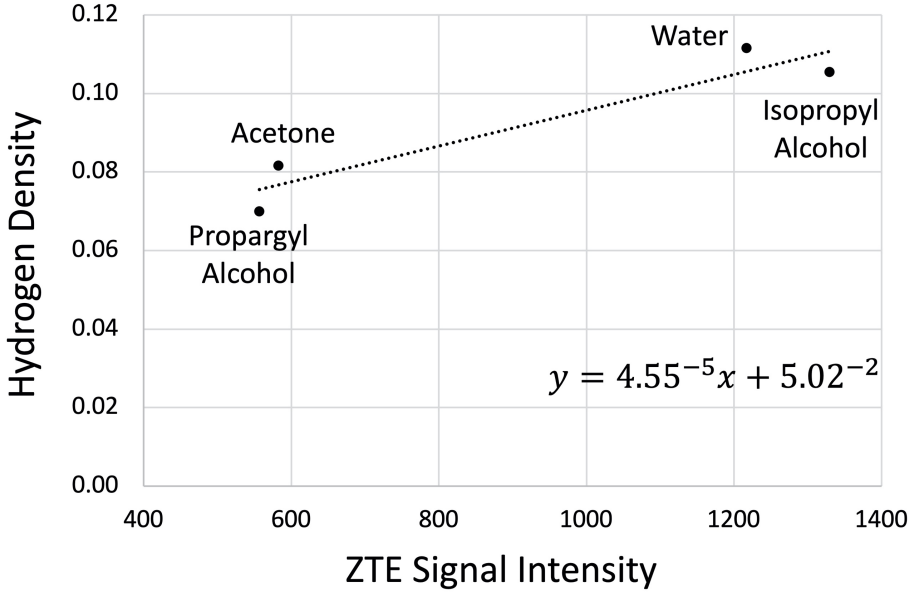

Figure 4.1: Zero-echo time MRI hydrogen density calibration curve
Calibration curve relating ZTE image signal intensity and hydrogen density. The organic molecule hydrogen density (h_{org}) was determined by subtracting the hydrogen content contributions from water (h_{H_2O}) and mineral (h_{HA}), as shown in Equation 22.

These materials were selected because they have a known (^{1}H) proton density (ranging from 0.07 to 0.11) and span the hydrogen density range of tissue-mimicking materials evaluated in this study (ranging from 0.09 to 0.11). Of note, this calibration curve is scanner and sequence parameter dependent (including resolution) and should be performed separately for all scanners.

$$h_{org} = \frac{h_{tot} - (h_{\mathrm{HA}} \cdot w_{HA} + h_{H_2O} \cdot w_{H_2O})}{1 - (w_{H_2O} + w_{HA})} \quad (22)$$

Multimodal Imaging Pipeline

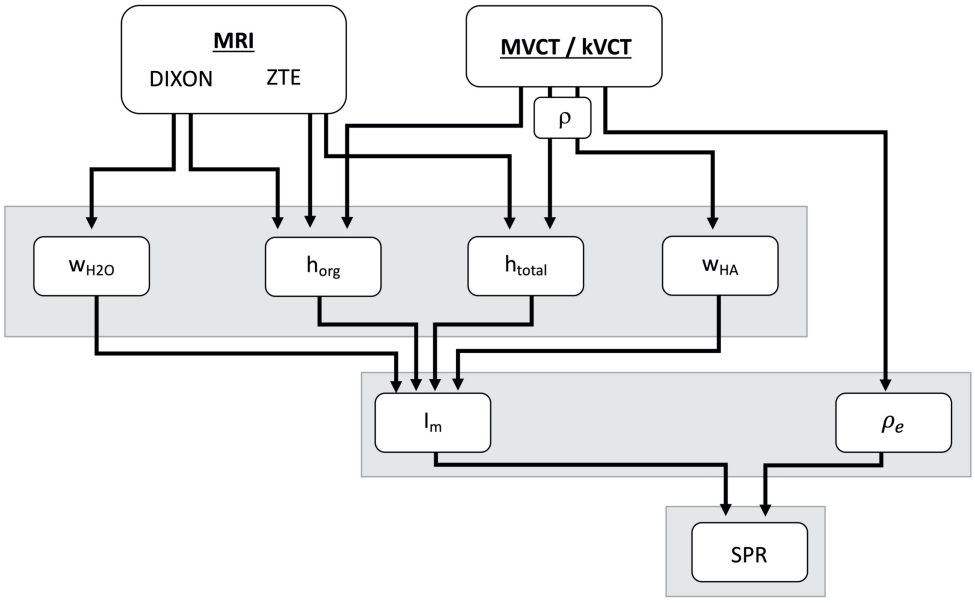

Figure 4.2: Multimodal imaging pipeline
Schematic illustrating the multimodal imaging pipeline used for determination of SPR using the UC method.

The full multimodal imaging pipeline for determining SPR is depicted in Figure 4.2. MRI datasets were acquired on a 3 Tesla scanner (MR750; GE Healthcare, Milwaukee, WI) using 3D acquisition and 2mm slice thickness, with all images acquired using a transmit/receive quadrature birdcage-type head coil. This type of head coil allowed for a highly homogenous signal acquisition[64] not requiring specific compensation for B_1 inhomogeneities. The Dixon water-only image was acquired using flip angle of 1°, TR of 3.77ms, and TE of 2.32ms and 1.12ms for the in-phase and out-of-phase images, respectively. The ZTE scan was acquired using flip angle of 0.6°, TR of 1.34ms, and TE of 0.016ms. kVCT images were acquired on a Siemens Somatom scanner using 120kVp energy and 1mm slice thickness. MVCT images were acquired on a TomoHD system using the 3.5MV energy MVCT beam with fine pitch and 1mm slice thickness reconstruction.

4.3 Sensitivity analysis

The sensitivity of our SPR determination methodology was evaluated by calculating changes in I_m and SPR due to variations in parameters measured using each imaging technique, specifically, HU (from CT) and water fraction and hydrogen density (from MRI). Variations in physical density and relative electron density were also separately used to assess the sensitivity of I_m and SPR on material density. CT number variations of ± 5 HU were used for the model sensitivity analysis, which is the recommended tolerance value for CT number accuracy of water and field uniformity provided by AAPM Task Group Report 66 for simulator and treatment planning applications in radiotherapy[65]. Although HU tolerances typically vary with imaging energy and material, ± 5 HU was considered reasonable for the purposes of assessing model sensitivity for tissue substitutes used in this study.

Quantitative MRI is a developing field that has primarily focused on methods for more specific tumor delineation, radiotherapy dose painting, interpretation of treatment response, and identification of normal tissue toxicity[66,67]. Unlike in CT, MR image intensity values do not correspond to physical or electron density, i.e. a pixel value in MRI is relative to other pixel values and depends on several factors related to material composition and scan parameters. A comprehensive investigation of quantitative MRI is beyond the scope of this study; however, our model sensitivity was evaluated based on perturbations of $\pm 10\%$ hydrogen density and water content derived from ZTE and DIXON scans, respectively. This accuracy range is consistent with results seen in this and prior work[59], and considered reasonable for the purposes of assessing model stability.

4.4 Results

Table 4.1 summarizes the results of each parameter used to calculate I_m using the multimodal imaging framework for the four tissue substitutes evaluated.

Table 4.1: Im calculation parameters
Tabulated values of parameters required for Im calculation estimated from multimodal imaging (MRI, UC-kVCT and UC-MVCT) versus physical measurements (PM).

	w_{H_2O} (%)	w_{HA} (%)	h_{tot} (%)	h_{org} (%)	ρ (g/cc)
Skin	91.5 (MRI)	6.5 (UC – kV)	10.9 (UC – kV)	5.9 (UC – kV)	1.05 (UC – kV)
		11.4 (UC – MV)	10.3 (UC – MV)	5.4 (UC – MV)	1.11 (UC – MV)
	75.0 ± 0.15 (PM)	0 ± 0.0 (PM)	9.4 ± 0.02 (PM)	4.0 ± 0.01 (PM)	1.06 ± 0.002 (PM)
Muscle	83.9 (MRI)	4.8 (UC – kV)	11.1 (UC – kV)	24.7 (UC – kV)	1.03 (UC – kV)
		7.6 (UC – MV)	10.7 (UC – MV)	15.2 (UC – MV)	1.06 (UC – MV)
	74.8 ± 0.15 (PM)	0 ± 0.0 (PM)	9.8 ± 0.02 (PM)	5.7 ± 0.01 (PM)	1.04 ± 0.002 (PM)
Adipose	38.3 (MRI)	0.5 (UC – kV)	11.9 (UC – kV)	12.3 (UC – kV)	0.97 (UC – kV)
		1.5 (UC – MV)	11.7 (UC – MV)	12.3 (UC – MV)	0.98 (UC – MV)
	46.8 ± 0.09 (PM)	0 ± 0.0 (PM)	12.0 ± 0.02 (PM)	12.8 ± 0.03 (PM)	0.95 ± 0.02 (PM)
Spongiosa	17.5 (MRI)	15.8 (UC – kV)	9.6 (UC – kV)	11.4 (UC – kV)	1.16 (UC – kV)
		7.9 (UC – MV)	10.5 (UC – MV)	11.4 (UC – MV)	1.06 (UC – MV)
	26.6 ± 0.05 (PM)	12.8 ± 0.03 (PM)	9.0 ± 0.02 (PM)	10.0 ± 0.02 (PM)	1.06 ± 0.002 (PM)

W_{H2O} estimated from the water-only Dixon MRI deviated from PM values by approximately 8.5%, 9.1%, 9.1%, and 16.5% for adipose, muscle, spongiosa, and skin, respectively. W_{HA} was calculated from physical density determined from CT, therefore all tissues evaluated using the multimodal imaging method showed some amount of hydroxyapatite due to non-zero physical density. The average W_{HA} values estimated from kVCT differed from PM values by 0.5%, 3.0%, 4.8%, and 6.5% for adipose, spongiosa, muscle, and skin, respectively while MVCT differed from PM values by 1.5%, 4.9%, 7.6%, and 11.4% for adipose, spongiosa, muscle, and skin, respectively. kVCT was more accurate than MVCT in determining w_{HA} for skin, muscle, and adipose due to the higher accuracy in calculating physical density values (used as a surrogate for determining calcium content) from kVCT versus MVCT in these tissues. The average physical density estimated from kVCT was within 2% of PM values for all tissues except spongiosa, which differed by 9.8%. The average physical density estimated from MVCT was within 1.5% for muscle and spongiosa, but differed by 3.3% and 4.5% for

adipose and skin, respectively. The average total hydrogen content (h_{tot}) estimated from ZTE MRI was within 1.5% using physical density derived from both kVCT and MVCT. The average organic hydrogen content (h_{org}) was within 2% of PM values derived from both kVCT and MVCT, with the exception of muscle, which deviated from PM values by approximately 20% for kVCT and 10% for MVCT. The high deviation in h_{org} values estimated for muscle are due to a single outlier phantom which had a higher average w_{H2O} value derived from the Dixon water image versus the other eight containers containing muscle. Despite this outlier, results of I_m and SPR estimated for muscle (shown below) retain accuracy due to the robustness of the multimodal imaging method, as demonstrated in the sensitivity analysis of I_m and SPR described below.

Results of relative electron density, mean ionization potential, and relative stopping power ratio computed separately for both kVCT and MVCT, along with SPR determined from the stoichiometric method for comparison, are shown in Figure 4.3. Points falling on the diagonal line indicate a match to physical measurements.

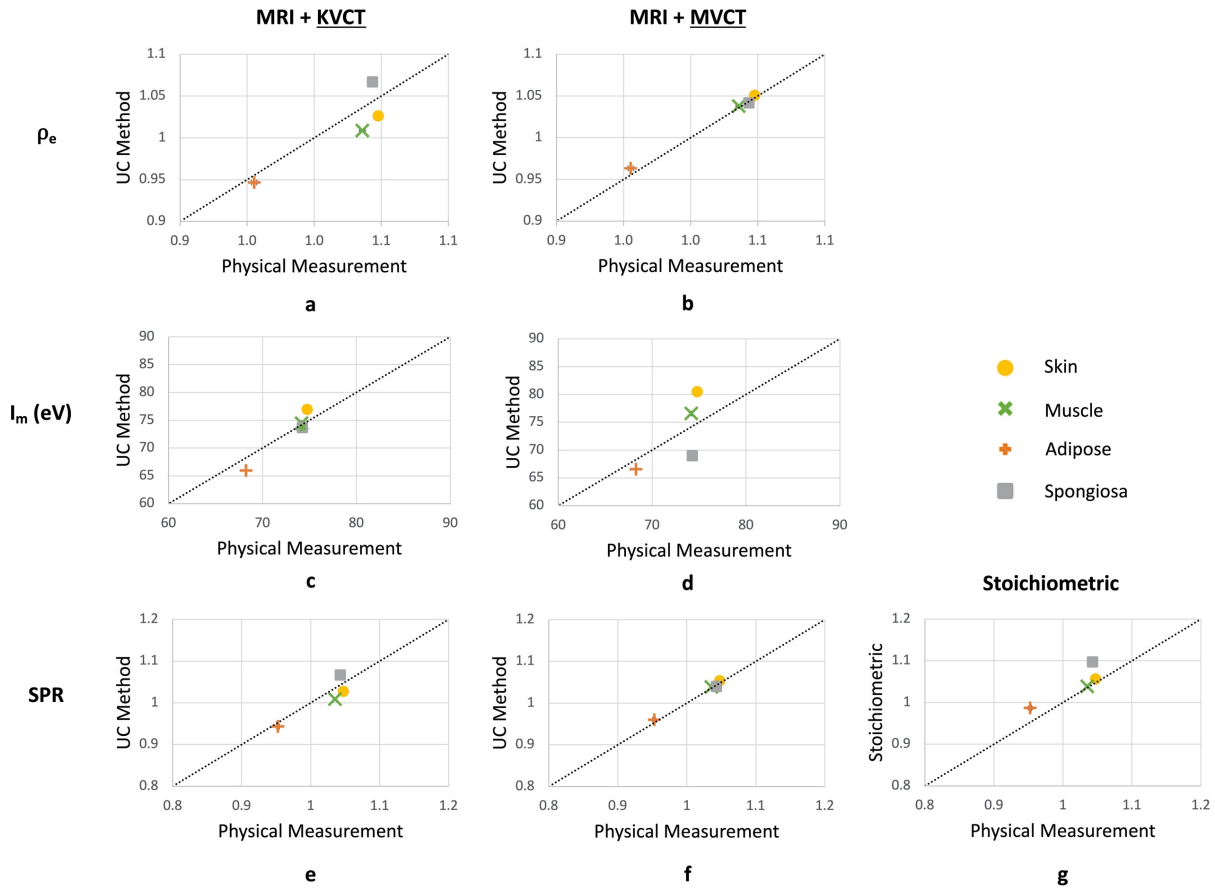

Figure 4.3: Results of multimodal imaging framework
Results of relative electron density, mean ionization potential, and stopping power ratios computed using kVCT (left) and MVCT (middle) compared to physical measurements. SPR from stoichiometric method is shown for comparison (panel g).

As expected, MVCT outperformed kVCT in determining relative electron density for all four tissues studied (Figure 4.3a,b), with average values within physical measurements by 2.6% for kVCT versus 0.9% for MVCT. Average ρ_e values differed from physical measurements by an average of 2.1%, 2.6%, 0.9%, and 2.2% for kVCT and an average of 0.3%, 0.2%, 0.9%, and 0.2% for MVCT in skin, muscle, adipose, and spongiosa, respectively. kVCT outperformed MVCT in determining mean ionization

potential (Figure 4.3c,d), with average values differing from PM calculations for kVCT by 2.9%, 0.4%, 3.4%, and 0.8% versus MVCT by 7.6%, 3.3%, 2.5%, and 7.2% for skin, muscle, adipose, and spongiosa, respectively. Deviations in I_m values computed from MVCT are predominantly due to errors in computing physical density as described above. Our multimodal imaging approach using both kVCT and MVCT was in close agreement to physical measurement SPR values for all tissue substitutes evaluated (Figure 4.3e,f) and provided higher accuracy than the stoichiometric method (Figure 4.3g) in the cases of adipose and spongiosa bone. Because SPR varies linearly with relative electron density and logarithmically with mean ionization potential, MVCT provided higher accuracy SPR values than kVCT due to the higher accuracy of MVCT for computing relative electron density. In skin, average SPR values were within 2.0% from kVCT, 0.5% from MVCT, and 0.8% from the stoichiometric method. In muscle, average SPR values were within 2.6% from kVCT, 0.3% from MVCT, and 0.3% from the stoichiometric method. In adipose, average SPR values were within 1.0% from kVCT, 0.8% from MVCT, and 3.6% from the stoichiometric method. In spongiosa bone, average SPR values were within 2.2% from kVCT, 0.4% from MVCT, and 5.1% from the stoichiometric method.

Table 4.2: Sensitivity analysis
Tabulated values for the percentage change to I_m and SPR from errors in kVCT and MVCT of magnitude ($\pm$5 HU) and water content and hydrogen density of magnitude ($\pm$10%).

			kVCT ($\pm$ 5 HU)		MVCT($\pm$ 5 HU)		w_{H2O} ($\pm$10%)		h_{tot} ($\pm$10%)	
	I_m	SPR	%I_m	%SPR	%I_m	%SPR	%I_m	%SPR	%I_m	%SPR
Skin	77.8	1.05	-0.20, 0.20	-0.49, 0.49	-0.35, 1.18	-0.47, 0.46	-1.76, 1.77	-0.05, 0.05	3.65, -3.80	0.01, -0.11
Muscle	76.5	1.04	-0.20, 0.20	-0.50, 0.50	-0.18, 0.18	-0.45, 0.45	-1.77, 1.79	-0.05, 0.05	3.70, -3.82	0.01, -0.11
Adipose	67.7	0.95	-0.18, -0.63	-0.52, 0.36	-0.20, 0.20	-0.52, 0.52	-1.75, 1.78	-0.05, 0.05	3.77, -3.68	0.01, -0.11
Spongiosa	72.4	1.04	-0.42, 0.42	-0.16, 0.16	-0.18, 0.18	-0.44, 0.44	-1.78, 1.81	-0.05, 0.05	3.81, -3.69	0.01, -0.11

Table 4.3: Sensitivity analysis with density
Tabulated values for the percentage change to I_m and SPR from errors in physical density and relative electron density.

			ρ (±10%)		ρ_e (± 10%)	
	I_m	SPR	%I_m	%SPR	%I_m	%SPR
Skin	77.8	1.05	-3.73, 3.43	-0.11, 0.09	0.0, 0.0	-9.54, 9.54
Muscle	76.5	1.04	-3.66, 3.34	-0.10, 0.09	0.0, 0.0	-9.65, 9.65
Adipose	67.7	0.95	-3.76, 2.69	-0.11, 0.07	0.0, 0.0	-10.47, 10.47
Spongiosa	72.4	1.04	-3.66, 3.54	-0.10, 0.10	0.0, 0.0	-9.58, 9.58

Results of the model sensitivity analysis demonstrating changes in mean ionization potential and stopping power ratios with variations in parameters estimated from CT HU, water fraction, and hydrogen density are shown in Table 4.2. Variations in CT number of ± 5 HU resulted in changes to I_m of up to 1.2% and SPR of up to 0.5%, with similar overall trends between kVCT and MVCT HU and subtle differences stemming from differences in their respective calibration curves. Variations in water content of $\pm 10\%$ led to changes in I_m of approximately $\pm 1.8\%$ and SPR of approximately $\pm 0.05\%$. With regards to parameters extracted from MRI, our model was more sensitive to variations in hydrogen density than water content, with variations in hydrogen density of $\pm 10\%$ leading to changes in I_m of approximately $\pm 3.8\%$ and SPR of approximately $\pm 0.1\%$. However, we found that the ZTE sequence used in this study for determining hydrogen density was quite accurate, with hydrogen density values well within 10% of physical measurements, as shown in the third column of Table 4.2.

Results of the model sensitivity with physical density and relative electron density are shown in Table 4.3. In our SPR determination methodology, mean ionization potential is sensitive to physical density uncertainties because physical density is used to calculate both mineral content (through calcium content, as shown in equation 21) and hydrogen density, which are parameters used to compute I_m. Variations in physical density led to changes in SPR only through changes in I_m, as physical density is not a parameter used directly to calculate SPR. Relative electron density, on the other hand, is not

used in determining I_m, therefore had no impact on this parameter. However, because SPR scales linearly with relative electron density (as shown in equation 9), variations of $\pm$10% in relative electron density correlated with $\pm$10% variations in SPR. As HU values were independently fit to physical and, separately, electron density, errors/uncertainties in the fitting of HU to either of these densities would manifest in SPR calculation very differently than errors/uncertainties in the HU value itself (which would affect both mass and electron density) as in Table 4.2.

4.5 Discussion

The proposed multimodal imaging framework provides improved accuracy in estimating stopping power ratio. In the proposed methodology, MRI is used to estimate the water content using a Dixon water-fat separation technique and hydrogen density using a ZTE technique. Either kVCT or MVCT can be used to estimate physical density, hydroxyapatite content, and relative electron density, with both modalities demonstrating accurate results of SPR.

Larger discrepancies in physical density values for spongiosa were seen for kVCT versus MVCT. Differences in the spongiosa bone are more prominent than for the other tissue substitute materials because the physical density of "standard" tissues are encoded into the calibration curves and, for our phantom, the density of the spongiosa phantom differed from typical human spongiosa bone. In practice, impacts of these differences may be seen in mineralized tissues which have a physical density different than standard values for bone, for example, in cases of lower bone mineral density such as osteoporosis. Notably, while this may be problematic if using kVCT, MVCT retains accuracy of electron density even in cases of non-standard density tissues. Of note, the proposed technique using a combination of MRI and MVCT provided results closer to physical measurements than the stoichiometric method, which has been demonstrated to provide high accuracy in SPR for soft tissues with composition close to water but reported to give errors in adipose of up to 3%[68] and bone of greater

than 3.5%[69]. With regards to hydroxyapatite composition, a potential solution to improve accuracy in deriving w_{HA} from physical density could include a segmentation which differentiates soft tissue from bone and setting non-bone w_{HA} values to zero. While more accurate methods of determining composition are under investigation, an important advantage of the UC Method for determining mean ionization potential is that the technique is robust to deviations in content of water, organic tissue, and hydroxyapatite[59] as indicated by the accuracy in estimations of SPR using either kV or MVCT. Intuitively, the robustness of the proposed technique appears to stem from the fact that data is being characterized by its individual components of water, organic material, or hydroxyapatite, which are extracted from the appropriate choice of imaging technique. For example, a water-specific MR scan (as in this proposed work) provided the data for water composition, limiting the error for this component to a relatively smaller set of possibilities relative to elemental/molecular composition and thus I_m. The robustness of the SPR determination methodology was demonstrated in the sensitivity analysis, with variations in CT number of ± 5 HU and variations in water content and hydrogen density of $\pm 10\%$ leading to changes of I_m within a couple percent and SPR within one percent. In contrast to the findings by Paganetti[15] for a kVCT-based stoichiometric calibration and SPR determination, our methodology limits the error due to I_m to a sub-1% level with electron density errors now dominating SPR determination.

Despite demonstrating improved accuracy in computing SPR with the proposed framework, future investigations would provide further validation on the method's accuracy. In particular, a final validation of proton range accuracy of our methodology would necessitate proton irradiation of actual tissues as imaging studies alone can provide significant though limited amounts of information. Additionally, improvements to the imaging protocols utilized may provide increase accuracy. For example, imaging pure water for quantitative analysis can cause errors in MRI. As shown in Figure 3.1c, the water container (center of phantom) does not produce the highest signal, likely due to the

ultra-long T_1 value of pure distilled water that would not be found biologically. In this study, tissue percent water was calculated by taking the ratio of intensity values to the highest image signal intensity, which was typically skin. Other methods that may result in more accurate water-fat separation (e.g. multi-peak models) may be useful to investigate.[70-72] In general, the proposed methodology can be applied to human biological tissues that exhibit similar physical characteristics to those provided in Tables 7 and 8, though not all materials fall within these criteria. For example, the hydrogen density of very dense bone is approximately 3-7%[46], falling outside the hydrogen density range of calibration materials used in this study (7-11%). To account for this, the calibration curve shown in Figure 4.1 would require additional calibration materials which span the full range of hydrogen density of materials being evaluated. Furthermore, materials with both extremely short T_2* values and high hydrogen content (for example, plastic) may produce erroneous results in the ZTE MRI scan as the short T_2* value may result in very little signal intensity despite having high hydrogen content. Metal implants commonly produce streaking artifacts and/or signal voids in both CT and MRI, posing challenges in accurate characterization of quantitative parameters. An additional MRI source of uncertainty for clinical implementation of this technique is signal uniformity and inhomogeneity. In this study, scans were acquired using a quadrature birdcage type head coil known for its high signal homogeneity, yet current MRI technology has developed away from these types of coils due to their lower signal-to-noise ratio compared with multi-channel receive-only phased array coils. While there exist numerous techniques (including the commonly used N4 technique[73]) to compensate for this B_1 inhomogeneity, image-based corrections operate under a set of assumptions (specifically that inhomogeneities are of low spatial frequency) with errors and biases occurring in situations outside of those. While this study serves as a proof of concept in four tissue substitute materials, a comprehensive investigation of the proposed methodology in non-biological and additional biological materials is required prior to clinical implementation.

A current limitation in the proposed framework stems from errors related to acquiring images using multiple modalities, for example, MR-to-CT image registration errors and differences in anatomical positioning of the patient between image acquisition. With perfect registration, we can expect to achieve high accuracy SPR estimation. In non-ideal cases (particularly for sites outside of the head/brain), we may expect that image registration errors between MR and CT are larger than those seen in this work. Active areas of research are being pursued to provide a framework for MRI-only dose calculation which would reduce the dependence/need for high-accuracy image registration. For SPR dose calculation in particular, an MRI-only framework may be achievable if MRI could be used to accurately quantify physical and electron density. Highly accurate MRI-only SPR calculation is challenging for a number of reasons with there being specific challenges due to ultrashort transverse relaxation time T_2/T_2^* in bone. Techniques such as UTE and ZTE have been developed to image materials with very short T_2/T_2^*,[74,75] however, studies have indicated that while these sequences have improved contrast in bone versus air, there are still instances of mislabeling[45]. Using calcium or phosphorus MRI can, in theory, be used as a surrogate for HA content in mineralized tissues knowing the chemical form of HA is $Ca_5(PO_4)_3(OH)$. However, the abundant isotope of calcium, ^{40}Ca, has no magnetic moment and is therefore not conducive to producing an MRI signal. Imaging the abundant isotope of phosphorus, ^{31}P, is currently an active area of research[76-79] but has not been well established on clinical systems and has limited spatial resolution. These methods, and others, are under investigation and could potentially allow for the implementation of an MRI-only framework.

The proposed imaging methodology using MRI and MVCT has provided a proof-of-concept for achieving sub-percent accuracy in the calculation of SPR for the four types of tissue substitutes evaluated while assuming perfect image registration. While there are several potential sources of error leading to the phenomenon of proton range uncertainty, the primary sources are those resulting from the conversion of CT number to stopping power ratio, contributing about 2.5%[13,15]. Reducing the beam

range uncertainty caused by this conversion from 2.5% to 1% would translate to a decrease in additional margin of several millimeters for a treatment beam of moderate energy. Along with decreased dose to healthy tissues, a practical application to this improvement would be in the selection of beam angles, which are often chosen such that the distal range is not directed at a nearby critical structure because the additional margins added to account for range uncertainty result in prohibitively high dose to these structures. Reducing margins by several millimeters would potentially allow for more optimal beam arrangements and a reduction in dose to nearby healthy critical structures.

4.6 Conclusion

We have demonstrated the methodology for improved SPR computation at the voxel level using the proposed multimodal imaging framework. We were able synthesize several tissue substitutes and compare results from our multimodal imaging framework to values computed directly from physical measurements. Using MRI and MVCT, SPR values derived using our method were within 1% of physical measurements and were more accurate than the stoichiometric method, which is currently the clinical standard for SPR estimation.

Chapter 5 – MRI to synthetic MVCT

Of the techniques used to generate synthetic CT from MRI, all have focused on the synthesis of CT images acquired in the kilovoltage (70-140kVp) energy range[52,80]. In light of the advantages provided by MV imaging for dose calculation and availability of MV hardware-equipped MRI-linacs, we propose a clinical pipeline in which MRI-derived synthetic MVCT (sMVCT) be used for initial treatment planning and the already present high-energy MV photon beam for online dose adaptation on the MRI-linac. This paradigm shift would offer a solution that eliminates the need for kV imaging systems for both initial treatment planning and online plan adaptation. Furthermore, undesirable characteristics of MV imaging, such as low SNR and CNR, are less important if the CT scan is being used purely for photon dose calculation rather than providing anatomical detail, as proposed in this workflow. The aim of this study was to provide proof-of-concept for generating synthetic MVCT images from MRIs using a 3D U-Net deep learning model and demonstrate the feasibility of our proposed workflow whereby the MRI is utilized for anatomical delineation while the MRI-derived sMVCT is used for radiotherapy dose calculation.

5.1 U-Net architecture and model

The dataset of 120 paired MRI-MVCT H&N images described in Chapter 3.3 was divided into an 80%-5%-15% train-validation-test split for the deep learning model with 18 test datasets. A deep neural network based on a fully convolutional 3D residual U-Net architecture was implemented to map MR images to sMVCT images as shown in Figure 5.1. Input to the model were volumetric patches (64x64x32) generated from 100 paired MRI and MVCT datasets. The encoding path consisted of the repeated application of two 3x3x3 convolutions, rectified linear unit (ReLU), group normalization, and 2x2x2 max pooling. The decoding path consisted of two up-sampling 3x3x3 convolutions, ReLU, and

group normalization. Skip connections were used to feed outputs of encoding layers to corresponding decoding layers to recover spatial information lost during down-sampling.

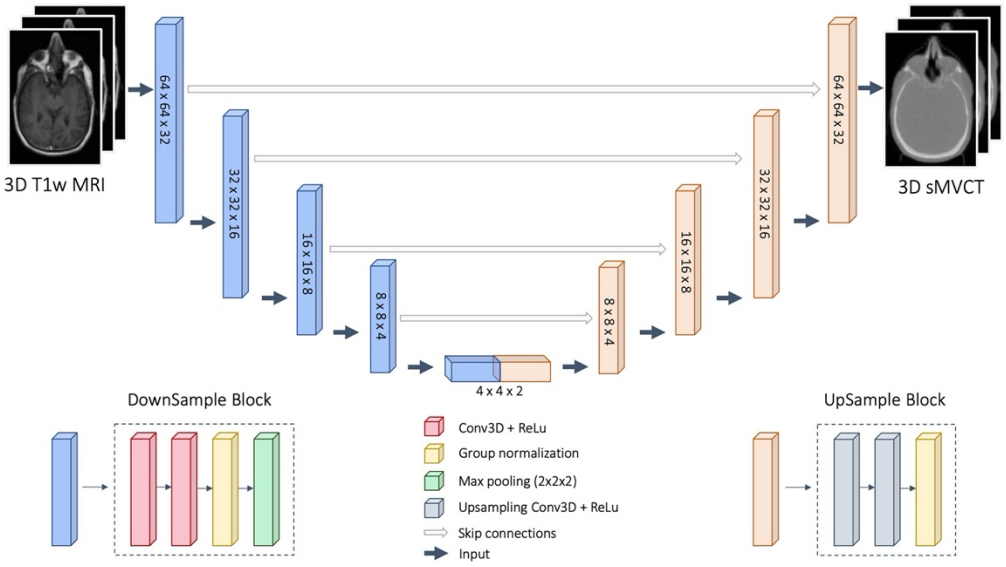

Figure 5.1: U-Net architecture (MRI to sMVCT)
U-Net architecture used in this study. Input to the model were 3D T_1w MRIs and output were 3D sMVCTs.

The U-Net was initialized with random parameters and trained for 100 epochs using the Adam optimizer (initial learning rate 10^{-4}) on an objective that minimizes the MAE (Equation 12). In this case for 3D MVCT HU maps and MRI-derived sMVCT HU maps, Equation 12 translates to a loss function represented by Equation 23:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|MVCT(i) - sMVCT(i)| \tag{23}$$

where N is the total number of image voxels and MVCT(i) and sMVCT(i) are the i^{th} voxels in the actual and synthetic MVCTs, respectively. Network and hyperparameter validation were performed on the validation dataset. During validation, model optimization was run long enough to qualitatively produce realistic sMVCT images and minimize the MAE between real and synthetized MVCT images in the validation dataset. Model accuracy was evaluated on images in the test dataset that were withheld during model training. The U-Net was implemented in PyTorch and trained on a single NVIDIA Titan

RTX 24GB GPU using a batch size of 60. Training the U-Net model took approximately 120 hours and reconstructing each 3D image dataset using the fully trained model took under 1 minute.

5.2. Evaluation of sMVCT images

Qualitative differences were assessed through visual side-by-side inspection of sMVCT versus MVCT images. The qualitative factors evaluated included overall shape of the anatomy, similarity in tissue contrasts, and the ability to distinguish air from bone, which remains an ongoing challenge in MR-to-CT synthesis given that both manifest as signal voids in most MR images[81]. Quantitative differences were assessed in different tissue types by segmenting images into whole body, soft tissue, and bone volumes. All images were masked to remove material outside of the patient, including air and setup devices. The whole body was defined as all material within the mask, which covered tissues inside the head-and-neck field-of-view for each patient. Within the whole body, the soft tissue volume was defined as voxels in the -120 to 100 HU range. Within the whole body, the bone volume was defined as voxels with HU value greater than 100 HU. For each tissue type, MAE (mean and standard deviation) was calculated between sMVCT and MVCT images (equation 23). Mean error was also computed and displayed as violin plots to visualize the error distribution for different tissue types in test patient datasets.

Figure 5.2 shows T_1-weighted MRI, MR-derived synthetic MVCT, actual MVCT, and HU difference maps for four example test datasets. Qualitatively, synthetic MVCT images have similar contrast to actual MVCT images with structural details including interfaces between air, bone, and tissue well preserved in the synthetic images. For example, bony regions between synthetic and actual images are well correlated in the mandible (Figure 5.2a), skull (Figure 5.2b,c), and vertebral bodies (Figure 5.2a,d). Structures containing air, such as the trachea (Figure 5.2a,d), were also well preserved.

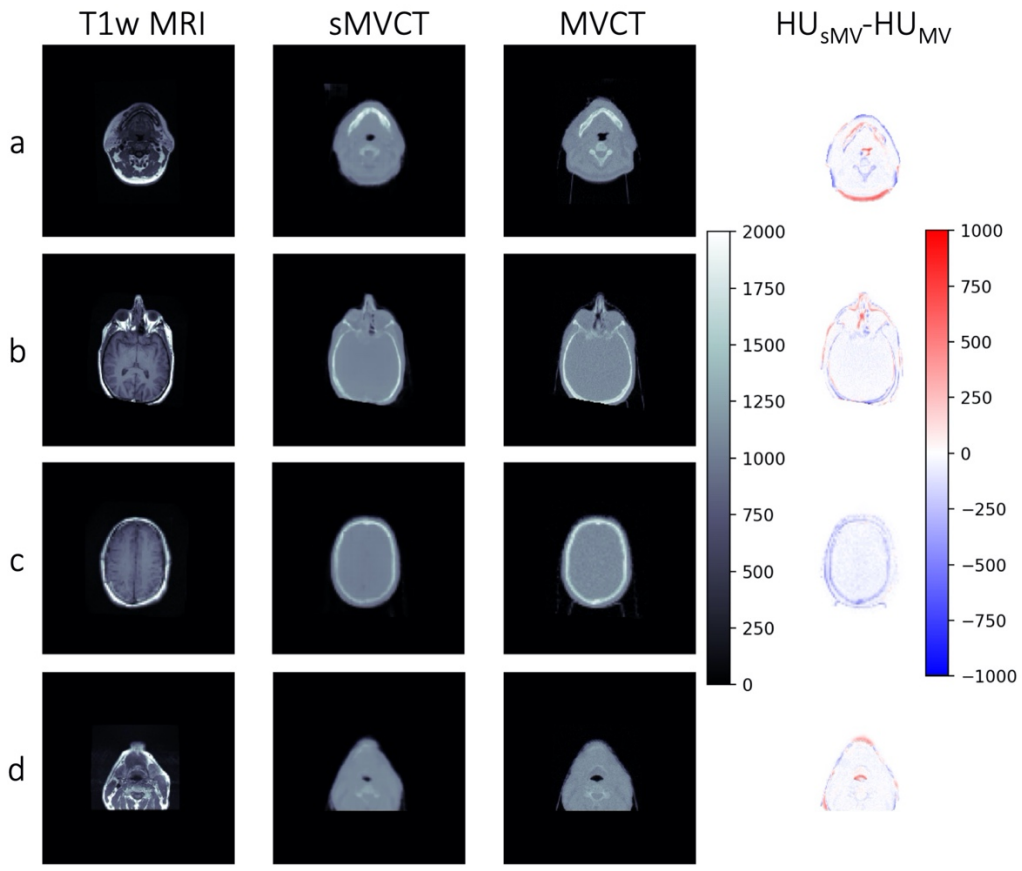

Figure 5.2: Results of MRI to synthetic MVCT
Axial slices of four representative T_1-weighted MRI, sMVCT, MVCT, and HU difference map. All images are displayed using a range of 0 to 2000 for visualization. Figures 11b and 11d are cropped due to limited field-of-view in the MRI.

Across the 18 test patients, the MAE (mean ± standard deviation) between sMVCT and MVCT averaged over all voxels in the whole body, soft tissue, and bone volumes was 93.3 ± 27.5, 78.2 ± 35.5, and 138.0 ± 43.4 HU, respectively. Mean error between sMVCT and MVCT images are displayed in Figure 5.3 for the four test patients used for radiotherapy treatment planning analysis (four left-most panels), as well as the average error over all 18 test patients (right-most panel). As expected, error in soft tissue volumes was the lowest while error in bone was highest for all test datasets.

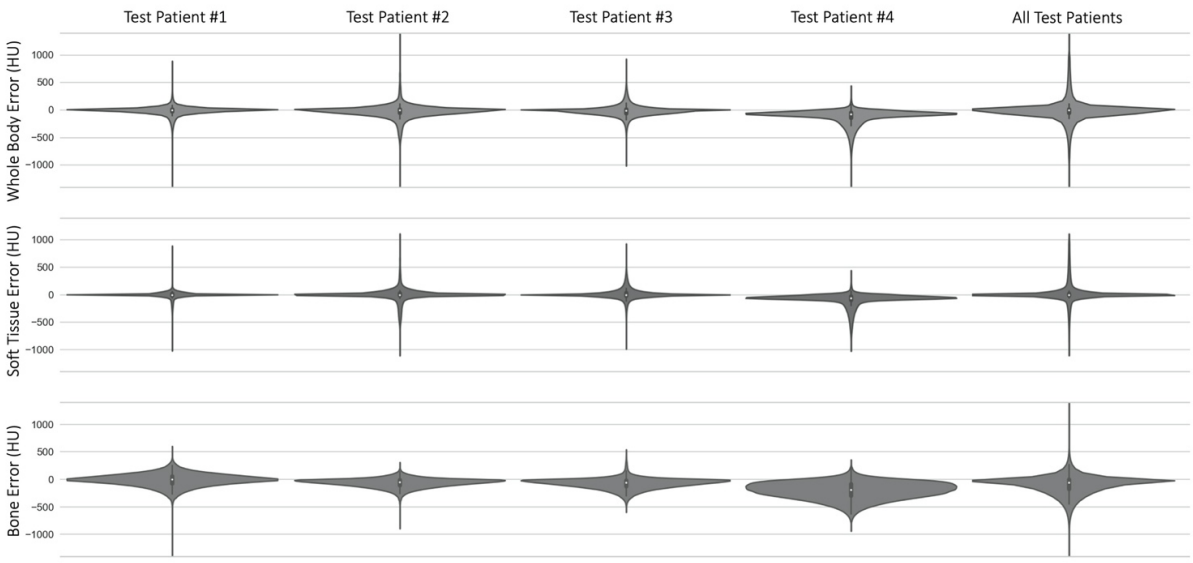

Figure 5.3. Violin plots of MRI to synthetic MVCT
Violin plots depicting the error in whole body, soft tissue, and bone voxels between sMVCT and MVCT in test patients 1-4 (four leftmost panels), in addition to all error average over all 18 test patients (rightmost panel).

Differences between sMVCT and MVCT images could occasionally be seen when the tissue positioning or patient anatomy were different between the MRI and MVCT in our dataset, as shown in Figure 5.4. For example, the top panel of Figure 5.4 shows changes in the nasal cavity filling while the bottom example highlights differences caused by immobilization devices such as oral stents and thermoplastic head-and-neck masks which are often used in radiotherapy but not in diagnostic MRI.

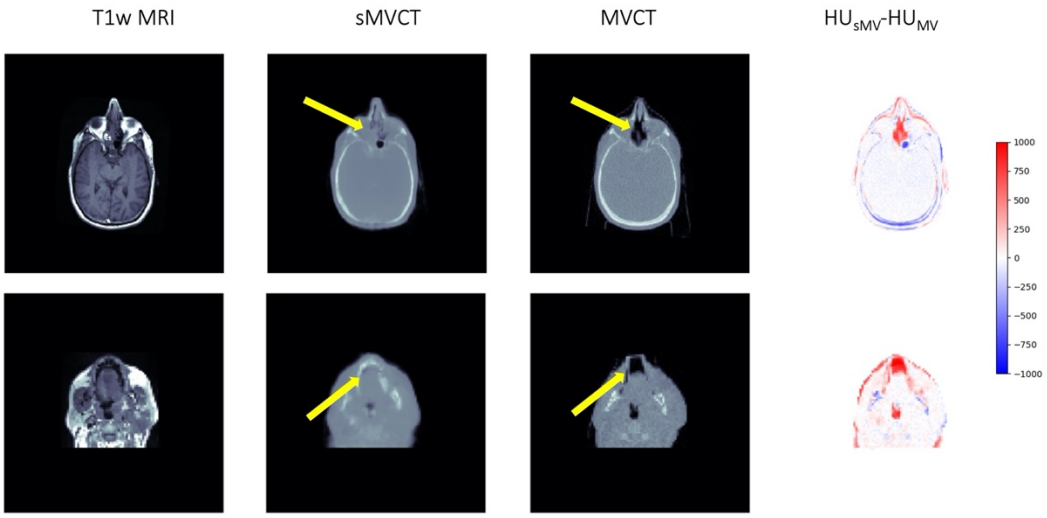

Figure 5.4. Dataset limitation examples
Axial slices of two representative T_1-weighted MRI, sMVCT, MVCT, and HU difference map for cases demonstrating limitations in the dataset including anatomical differences and presence or absence of patient immobilization devices as shown by the yellow arrows.

5.3 Treatment planning evaluation

Of the 18 test datasets, four representing a diversity of H&N sites including base-of-tongue, thyroid, neck, and nasal cavity were used to generate radiotherapy treatment plans to assess dosimetric impacts of using sMVCT images. Treatment planning was performed in RayStation (v.7.0 RayStation Laboratories, Stockholm, Sweden) using 6MV photon coplanar volumetric modulated arc therapy (VMAT) and an MVCT HU-to-electron density CT calibration curve for dose calculation. Contours were copied from the original CT planning dataset to the co-registered MVCT and sMVCT datasets. In cases where the MRI field-of-view was limited, contours were edited to be fully covered within the external contour. Planned dose, organ-at-risk (OAR) constraints, and planning target volume (PTV) coverage were reproduced based on the treatment plans used clinically for each patient. Dose was normalized to 95% of the PTV receiving 100% of the prescription. Plans generated on the MVCT datasets were recalculated onto sMVCT datasets using fixed plan parameters. Dose distributions for

MVCT and sMVCT were compared by assessing several different clinically relevant dose-volume-histogram (DVH) endpoints and the gamma index[82] calculated at 3%/3mm and 2%/2mm locally at a 10% dose threshold.

Figure 5.5 shows example dose distributions calculated on actual and synthetic MVCT test image sets across four different H&N sites. Dose volume histograms for the PTV and relevant OARs are shown. Dose distributions are visually very similar between actual and synthetic CT images. The majority of DVHs directly overlay each other, indicating very close agreement in dose distributions between structures. The only exceptions were the PTV and eye structures of test patient #4, who was treated to a tumor volume within the nasal cavity. Dose distribution disagreements in the PTV were caused by tumor growth within the sinus cavity that occurred between MRI and MVCT image acquisition. At the time the MRI was acquired (about three weeks prior to the MVCT) the tumor was smaller, therefore, the MRI-predicted sMVCT exhibited less tissue within the sinus versus the MVCT, which can be seen in the bottom panel of Figure 5.5. Small dose distribution disagreements were also seen in the eye structures for this patient because the eyes were relatively small structures directly abutting the PTV, therefore their DVHs were especially sensitive to small deviations in dose distributions between the MVCT and sMVCT.

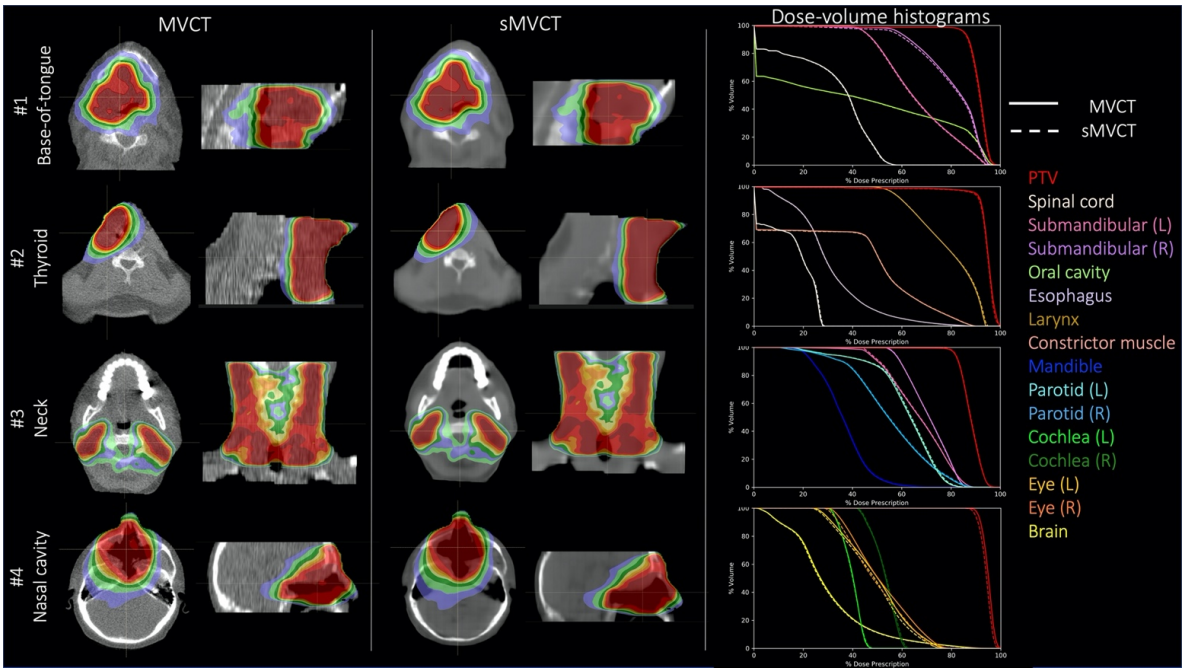

Figure 5.5. Dosimetric results
Axial dose distributions for base-of-tongue, thyroid, neck, and nasal cavity treatment plans computed on actual MVCTs and synthetic MVCTs with corresponding DVHs.

Table 5.1 shows results of gamma metrics calculated to compare sMVCT and MVCT dose distributions using metrics of 3%/3mm and 2%/2mm. Although all gamma values were within the 90% threshold typically used clinically for comparing dose distributions, investigating causes of lower gamma values is helpful for identifying and understanding dosimetric impacts of using a synthesized versus real image. For example, the lowest gamma values were seen in patients #2 and #4. In the case of patient #2, minor dose disagreement was attributed to small differences in the tissue surface due to the immobilization used for the MVCT but not MRI, which impacted this plan more than others due to the location of the target volume being at the patient surface. Discrepancies in patient #4 were due to changes in sinus filling as previously described.

Table 5.1. Gamma passing rates
Gamma passing rates between treatment plans computed with fixed plan parameters onto sMVCT and MVCT images and calculated using 10% local dose threshold at 3%3mm and 2%/2mm.

	3%/3mm	2%/2mm
#1 Base-of-tongue	99.3	98.3
#2 Thyroid	97.5	93.8
#3 Neck	99.9	99.6
#4 Nasal cavity	98.9	95.4

5.4 Discussion

Our results suggest that synthetic MVCT images can be generated from T_1-weighted gradient echo MRIs using a 3D U-Net for H&N photon radiotherapy treatment planning. Synthetic MVCT images predicted by the model exhibited similar qualitative features to real MVCT images. We measured MAE values between sMVCT and MVCT across whole body, soft tissue, and bone volumes as 93.3 ± 27.5, 78.2 ± 35.5, and 138.0 ± 43.4 HU, respectively, which is similar to deep learning approaches previously reported for synthetic kVCT. For example, Gupta et al reported MAE values of 81.0 ± 14.6, 17.6 ± 3.4, and 193.1 ± 38.3 HU for all whole body, tissue, and bone regions, respectively, when learning the relationship between T_1 MR and kVCT[83]. For H&N-based MR-derived sCT specifically, Wang et al reported MAE values of 97 ± 13, 131 ± 24, and 357 ± 44 HU for soft tissue, whole body, and bone, respectively, for T_2 MR to kVCT learning[84]. In a systematic literature review of synthetic CT generation for MRI-only radiotherapy planning, voxel-based techniques using standard MRI sequences yielded MAE results between 11-135HU and 91-99.9% gamma passing rate depending on MR technique, body site, and gamma criteria[85]. Of note, all previous studies in this realm investigated kV imaging, whereas our study investigated MV imaging. CT scans acquired at kV and MV energies yield different HU values for higher density materials such as bone, where the bilinear kV and linear MV CT calibration curves deviate from one another[28]. For example, typical bony tissue with relative

electron density of 1.2 will have HU values of ~400 in a kVCT and ~250 in an MVCT, which contributed to why our MV approach yielded comparable or lower MAE values in bone relative to prior kV approaches. In our work, dose distributions computed on synthetic images were very similar to those computed on real MVCT images, with gamma metrics >97% and >93% when computed at clinical thresholds of 3%/3mm and 2%/2mm, respectively. Our dosimetric results are similar to those previously described in actual versus synthetic CT comparisons, with Gupta et al[83] and Wang et al[84] also demonstrating overlapping DVH curves for treatment plans calculated on actual versus synthetic images.

While many studies have been performed to investigate the translation between MRIs and traditional kVCT images for purposes of MR-only radiotherapy planning, this work is the first to synthesize MVCT images from MRI. While MVCT has several advantages relative to its kV counterpart, such as fewer beam hardening and streaking artifacts, these come at the expense of lower SNR and CNR. Techniques have been implemented to minimize noise in MV imaging through statistical methods[86,87] and deep learning approaches[88,89], however, reduced SNR and CNR inherent to MV imaging becomes less important if the CT scan is being used purely for photon dose calculation rather than providing anatomical detail, as proposed in this workflow.

Our results indicate that high quality synthetic CT images can be achieved even when derived from MR images acquired with some variability between scanners and acquisition parameters. A limitation of this study was due to the diagnostic nature of the MR image datasets, which resulted in different patient set-up and, oftentimes, a time delay between MR and MVCT acquisitions. Patients being treated for H&N cancer are particularly susceptible to experiencing changes in weight and/or tumor volume over the course of their treatment[83,90,91], suggesting that treatment planning images should be acquired as close to the beginning of treatment as possible to avoid differences between scans caused by anatomical changes in tumor volume, sinus filling, and/or weight. We anticipate these

limitations will be largely mitigated through the utilization of radiation oncology-dedicated MR scanners for initial treatment planning, which will consist of standardized acquisition protocols, identical immobilization devices and patient setup as used for treatment, and a shorter time duration between simulation and treatment (ideally no more than 1-2 weeks). To better distinguish between bone and air, MRI protocols used for this purpose may include specialized techniques such as ultrashort UTE or ZTE MRI which have been developed to image materials with ultrashort transverse relaxation time (e.g. bone and other solids)[74,75]. Furthermore, using onboard MV imaging for online dose adaptation, as in our proposed future workflow, will mitigate issues caused by anatomical changes occurring between original treatment planning images and onboard pre-treatment MRIs, eliminating the need for co-registration of these datasets.

While adaptive radiotherapy planning using combined MRI and MVCT for photon radiotherapy was the focus of this work, MVCT can also provide additional advantages in proton radiotherapy. Range uncertainty remains an ongoing challenge in proton radiotherapy, with uncertainties related to kV HU-to-SPR calibration curves necessitating additional treatment margins on the order of 1-2% the proton range[16,17]. In the energy range utilized for kV imaging acquisition, photons interact primarily though the photoelectric effect which is highly sensitive to the material molecular composition. For kVCT scans, the relationship between HU and electron density (to which SPR is linearly related) is not one-to-one and can exhibit uncertainties as large as 7%[28,92,93]. In contrast, the primary mode of photon interaction for high energy x-rays is Compton scattering, which is predominantly dependent on electron density, resulting in a one-to-one relationship between HU and electron density at higher energies[94,95]. MVCT has been demonstrated to provide higher accuracy for mapping both electron density and SPR[28,92] and deriving stoichiometric HU-to-SPR calibration curves[57]. When it comes to both CT image synthesis for treatment planning and adaptive planning in proton radiotherapy, synthetic MVCTs may provide a more accurate alternative to synthetic kVCTs.

The U-Net is one of the most popular architectures for image-to-image domain translation and was used in this study to map MRI to synthetic CT using supervised learning[96,97]. Other architectures incorporating weakly-supervised and unsupervised learning, such as generalized adversarial networks (GANs) and its variants (e.g. CycleGAN) have also demonstrated promising results for sCT generation[98,99], though initial evaluations of deep learning methods for medical image synthesis suggest that the supervised U-Net can achieve lower MAE, higher structural similarity, and higher peak signal-to-noise than synthetic CT images provided by traditional image transformation networks[100]. Nevertheless, the method presented here is equally compatible with a GAN approach, which could be implemented by adding an additional adversarial loss to the objective in Equation 23. Overall, it is expected that the incorporation of additional loss components, modifications to the network backbone (3D U-Net), and future developments within the rapidly evolving field of image-to-image synthesis will further improve the results presented here.

5.5 Conclusion

In summary, this study presents a novel technique of synthesizing MVCT images from T_1w MR images using a deep learning U-Net model in a dataset of 120 patients treated for H&N cancer. MAE values between sMVCT and MVCT datasets were 93.3 ± 27.5, 78.2 ± 35.5, and 138.0 ± 43.4 HU for whole body, soft tissue, and bone volumes, respectively. Dose distributions evaluated for four test cases showed very close agreement between synthetic MVCT and actual MVCT images when evaluated using DVHs and gamma dose metrics, which averaged to 98.9 ± 1.0% and 96.8 ± 2.6% analyzed at 3%/3mm and 2%/2mm, respectively. These results demonstrate the feasibility of using MRI-derived sMVCT in an MR-only treatment planning workflow.

Chapter 6 – kVCT to synthetic MVCT

In the treatment planning stage of radiation therapy, kVCT scans serve two main purposes: 1) identification of normal and diseased tissues for contouring and 2) calculation of mass/electron densities and/or stopping power ratio required for dose calculation. For the former purpose, kVCT scans possess favorable characteristics in terms of soft and bony tissue contrast relative to noise and imaging dose. For the latter purpose, kVCT may have greater uncertainty in the determination of electron density and proton stopping power ratio. MVCT HU, on the other hand, exhibits higher accuracy when mapping HU to ρ_e and SPR. In terms of dose calculation accuracy, a methodology to create a more accurate relationship between kVCT HU and electron density (in the case of photon radiotherapy) or SPR (in the case of proton radiotherapy) would be ideal. In this study, we hypothesized that deep learning can be used to improve the mapping between kVCT HU and ρ_e or SPR by converting kVCT HU to synthetic images that inherently map these values more accurately.

In recent years, new image processing techniques have been developed to translate the characteristics of one scan to mimic the characteristics of another. In particular, machine learning techniques involving convolutional neural networks (CNNs) have grown in prominence and popularity. Involving radiation oncology volumetric scans for dose calculation, two main use cases have emerged: conversion of MR data to kVCT HU data[101-103] for MRI-only radiotherapy planning and improving the image quality of cone-beam CT (CBCT) images from kVCT data[104-106]. Generally, these deep learning techniques have found achievable errors in the range of 50-100 HU mean absolute error, with results improving with incremental developments in machine learning technology. In both discussed use cases, the imaging dataset that is being translated towards is fan beam kVCT HU. As there does not exist a one-to-one relationship between single scan kVCT HU and electron density or SPR, it is expected that translation of imaging data to synthesize kVCT HU will include any of the

inherent electron density or SPR uncertainties in this process. As kVCTs exhibit more contrast and less noise than MVCTs, we hypothesized that it was possible to accurately translate (through deep machine learning) trends in HU values from kVCT to MVCT. In doing so, we anticipated that synthetic MVCT images would naturally exhibit higher accuracy of electron density and stopping power ratio, thus promising improvements for dose calculation accuracy. In this work, we report on the first known application of deep machine learning to reduce uncertainty in the relationship between kVCT HU and ρ_e and SPR through learning of MVCT HU data. Through this learning, we are able to show the differences/uncertainties in ρ_e and SPR as calculated by kVCT HU compared with those calculated by sMVCT/MVCT HU.

After creation of the CycleGAN machine learning model, the clinical workflow proposed through this work requires only subsequent kVCT data for the conversion to sMVCT data, potentially alleviating the need for a MVCT machine or scan. In this workflow, we propose that the original kVCT images, with the high contrast and low noise intrinsic to kV imaging, be leveraged for localization and contouring of normal and diseased tissues while the sMVCT generated from the machine learning model be used for dose calculation in the background, as shown in Figure 6.1.

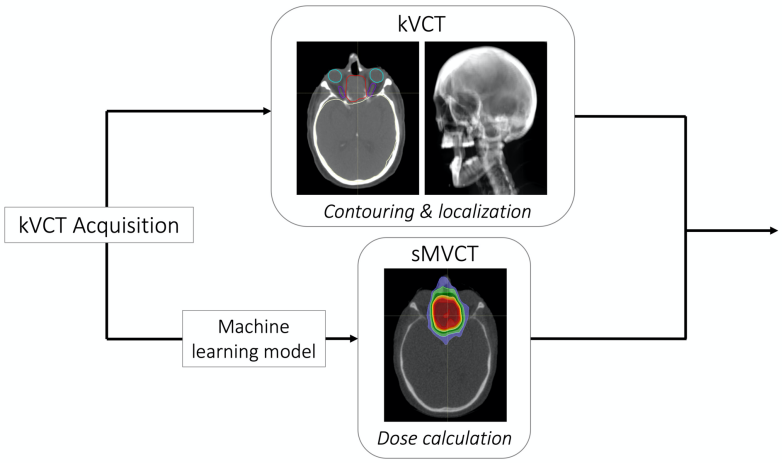

Figure 6.1. Workflow for kVCT to synthetic MVCT
The proposed clinical workflow requires a single kVCT scan, which is subsequently used for contouring and localization (top panel) and converted into a sMVCT using the machine learning model for dose calculation (bottom panel).

6.1 Evaluation of tissue-mimicking phantom

To motivate and demonstrate the increased accuracy of ρ_e and SPR provided by MVCT over kVCT, the in-house tissue-mimicking phantom described in Chapter 3.1 was used to compare physical measurements of ρ_e and SPR with those determined from kV and MV CT imaging. Results of relative electron density and stopping power ratio for the four tissue-mimicking phantoms that were measured and calculated through kV and MV CT calibration curves described in Chapter 3.5 are shown in Table 6.1. Percent differences between kV and MV imaging and physical measurements (% diff) are also shown.

Table 6.1. MVCT and kVCT values for synthetic tissue phantom
Results of ρ_e and SPR values calculated using the kV and MV CT calibration curves for skin, muscle, adipose, and spongiosa tissue mimicking phantoms.

	Relative electron density (g/cc)			SPR $(115\ MeV)$		
	Measured	kVCT (% diff)	MVCT (% diff)	Measured	kVCT (% diff)	MVCT (% diff)
Skin	1.048 ± 0.002	1.026 (-2.10)	1.051 (0.29)	1.049 ± 0.002	1.055 (0.62)	1.052 (0.29)
Muscle	1.036 ± 0.002	1.009 (-2.61)	1.038 (0.19)	1.036 ± 0.002	1.038 (0.13)	1.037 (0.11)
Adipose	0.955 ± 0.002	0.947 (-0.84)	0.963 (0.84)	0.953 ± 0.002	0.978 (2.58)	0.970 (1.70)
Spongiosa	1.044 ± 0.002	1.067 (2.20)	1.042 (-0.19)	1.044 ± 0.002	1.090 (4.38)	1.042 (-0.22)

Electron density values computed using MVCT were within 0.29%, 0.19%, 0.84%, and -0.19% versus those computed using kVCT of -2.10%, -2.61%, -0.84%, and 2.20% for skin, muscle, adipose, and spongiosa, respectively. In the case of stopping power ratio, both kVCT and MVCT provided accuracy to within 1% for skin and muscle tissues. In the case of adipose, physical measurements of SPR differed by 2.58% for kVCT and 1.70% for MVCT. In the case of bone, SPR values computed with MVCT were within 0.22% of physical measurements while kVCT deviated by more than 4%.

6.1 CycleGAN architecture and model

In this study, a generative adversarial network[15] (GAN) of CNNs was used to automatically translate between kVCT and MVCT images. The CycleGAN[107] machine learning model consists of a pair of parallel opposed GANs to learn the relationship from image X to image Y (i.e. MVCT to kVCT) and from image Y to image X (i.e. kVCT to MVCT). Figure 6.2 provides a view of the CycleGAN loss architecture for one of the GAN pairs. These two GANs work in opposition to each other to generate and detect increasingly more accurate fake images. The role of the generator is to generate data that as closely as possible matches the real data while the discriminator's role is to accurately distinguish between real and "fake" generator produced data. In general, the roles of the generator and discriminator are in competition with each other as the generator learns to produce increasingly realistic images that the discriminator learns to distinguish as "fake" for the purpose of model improvement. The learning of both GANs are coupled with a cyclic loss which penalizes the degree to which the composed forward and backward transformation varied from the identity function. If $G: X \rightarrow Y$ is the generator producing kVCT images from MVCT images and $F: Y \rightarrow X$ is the generator producing MVCT images from kVCT images, then the cyclic loss can be written as

$$L_{cyc} = \frac{1}{N}\sum_{i=1}^{N}\left\|F\left(G(x_i)\right) - x_i\right\|_1 + \left\|G\left(F(y_i)\right) - y_i\right\|_1 \quad (24)$$

where $\|\cdot\|_1$ is the L1 norm of a vector, x_i is an MVCT training example, and y_i is a kVCT training example.

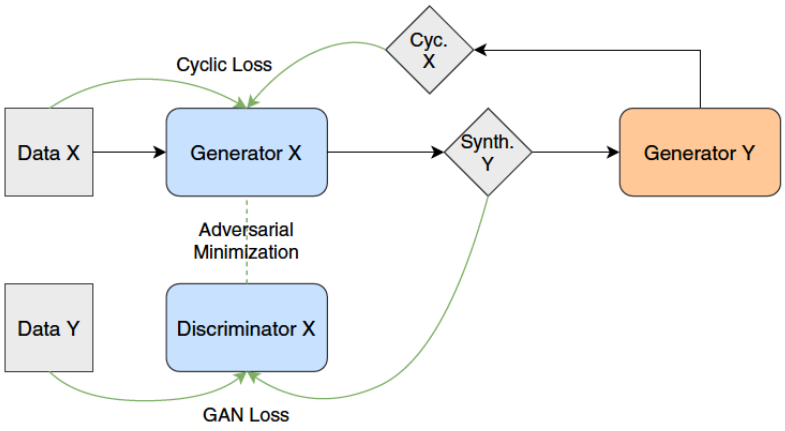

Figure 6.2 (courtesy of Luciano Vinas[88]). Loss architecture of CycleGAN Loss architecture for the X data-type generator and discriminator. Loss architecture for the Y type generator and discriminator is a mirror image but with X and Y switched.

The CycleGAN deep learning algorithm was implemented in PyTorch using Zhu *et al*'s architecture[16]. The generators consisted of a U-Net architecture with a 2x2 bottleneck and the discriminators consisted of a four-layer PatchGAN as described work previously published by our group[88]. Our model was trained using the Adam optimizer with a learning rate of 0.00025, β_1 of 0.5, and ϵ of 1.0×10^{-8}. Training was run for 210 epochs as it was found, during validation, that going past this point negligibly improved generator expressivity while adding to adversarial noise and potential overtraining. Skip connections between encoding layers and corresponding decoding layers were used to recover spatial information lost during down-sampling. Our methodology is analogous to prior CycleGAN machine learning studies of CT and on-board treatment machine CT imaging[108,109]. The U-Net architecture used for both is visualized in Figure 6.3.

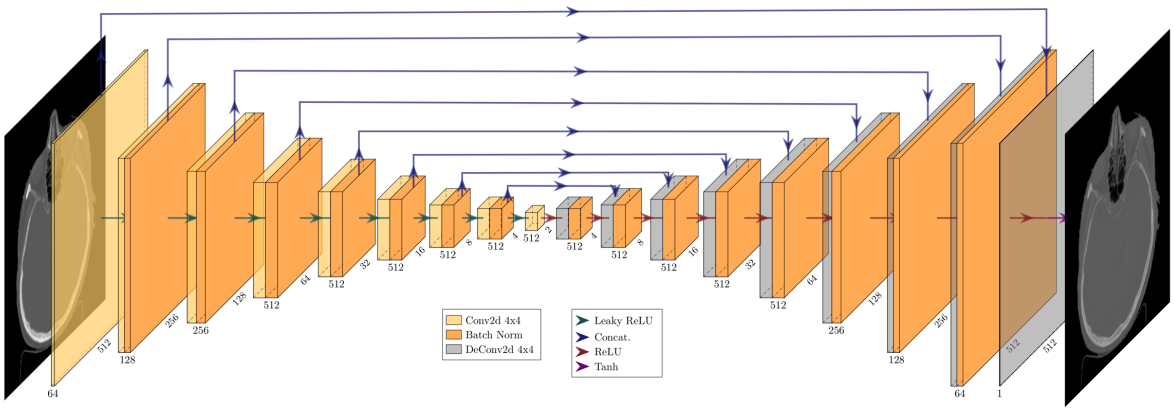

Figure 6.3. U-Net architecture (kVCT to sMVCT)
Schematic diagram of the U-Net 256 generator architecture used in this work. This figure shows the image transformation pipeline from kVCT to sMVCT image.

6.2 Evaluation of *in vivo* sMVCT images

CycleGAN machine learning of kVCT images to MVCT images was used to learn the relationship between kVCT HU and MVCT HU using the H&N database of 120 images described in Chapter 3.3. 100 paired kVCT-MVCT image datasets were used for training and validation while 20 datasets were used for testing. During final conversion and analysis, data were rescaled back to HU and then converted to electron density and stopping power ratio using the linear piecewise interpolations of calibration data shown in Figures 3.2 and 3.4, respectively.

Patient kVCT and sMVCT data (HU, ρ_e , and SPR) were compared with MVCT data as the reference. Figure 6.4 shows exemplar images for kVCT HU, sMVCT HU, MVCT HU, and difference heatmaps of kVCT HU and sMVCT HU compared to MVCT HU for four different subject scans from the test group. In the difference heatmaps (two rightmost columns), there exhibits some trends in value differences between different tissue types. As expected from the tissue mimicking phantom results, bony tissues exhibit the largest difference between kVCT and MVCT, with kVCT images exhibiting higher HU values than those exhibited by corresponding MVCT images, leading to an overestimation

of electron density and SPR values mapped from its CT calibration curve. Overall, there appears to be learning of some characteristics of MVCT images beyond its tissue-independent relationship to electron density. sMVCT images exhibit increased noise compared with equivalent kVCT images, an unintended consequence of the learning process.

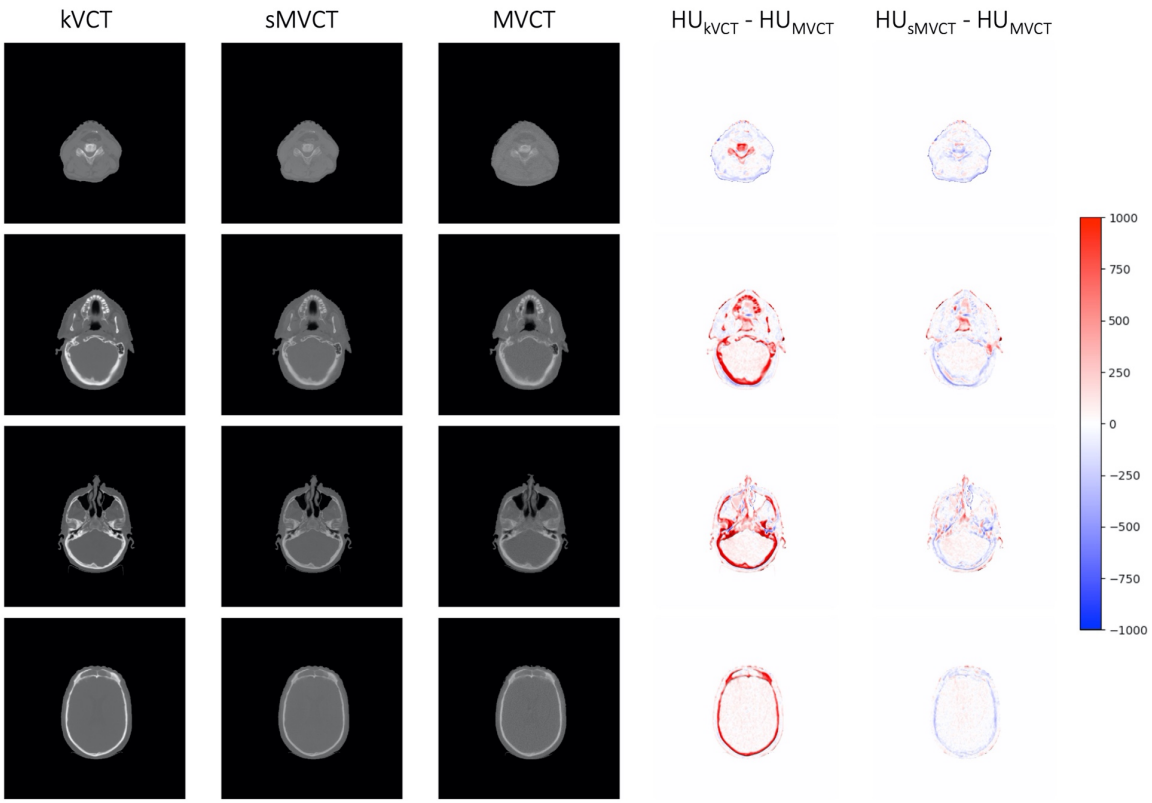

Figure 6.4. Results of kVCT to synthetic MVCT
Exemplar images from four different subject scans of kVCT HU (1st column), sMVCT HU (2nd column), MVCT HU (3rd column), and HU difference heat maps for kVCT versus MVCT (4th column) and for sMVCT versus MVCT (5th column). The HU scale used in the difference maps is shown on the right.

Within the 20 test patient datasets, additional analysis was completed to assess the changes that occurred through the deep learning sMVCT creation process. Tissue segmentation was implemented to compare electron density and stopping power ratio accuracy for soft tissue and bony tissue at standard HU thresholds. A total of three tissue masks were created from the kVCT data: a body mask,

a bony tissue mask, and a soft tissue mask. The body mask was created by truncating all HU values less than -400 and then segmenting the different connected regions. The largest of these connected regions was defined to be the body region and refined to a mask through a binary fill operation which preserved any cavities (including nasal and sinus cavities) as well as the lungs of the body. All voxels inside the body mask with HU values greater than 150 HU were considered to be part of the bony tissue mask. The set subtraction between body mask and bony tissue mask produced the soft tissue mask. ME, MAE, SNR, PSNR, MSE, RMSE, and SSIM (equations 11-16) were computed to assess differences between the imaging datasets were calculated at the voxel level.

Quantitatively, similar trends can be seen between kVCT/sMVCT HU and MVCT HU values as seen in qualitative results shown in Figure 6.4. A summary of similarity metrics between kVCT HU, sMVCT HU and MVCT HU are presented in Table 6.2. Following CycleGAN learning, sMVCT HU values were markedly more similar to MVCT HU values than kVCT values. In particular, mean ME and RMSE HU values improved for both soft tissue and bone with the largest improvement in bone. MAE HU values did not numerically show this same improvement with only bone MAE HU between sMVCT and MVCT showing a large decrease. A potential reason for MAE values not improving may be due to the large amount of noise inherent in MVCT and sMVCT scans. To assess the contribution of noise on MAE, the voxel MVCT HU data of a center ROI within a uniform water container was measured and the data fit to a Gaussian distribution ($R^2 = 0.96$) with standard deviation of 55 HU. Assuming a Gaussian distribution for the soft tissue/water CT noise, the MAE due to only CT noise would then be equivalent to the mean of a half-normal distribution, $\sigma\sqrt{2}/\sqrt{\pi}$. For an MVCT HU Gaussian noise distribution of 55 HU (1σ), the lowest achievable MAE would then be 44 HU assuming only noise. While the kVCT scans do not possess significant noise (< 10HU for 1σ), sMVCT scans will attempt to mimic both the contrast and noise profile of MVCT scans. To assess the noise properties of these synthetic scans, SNR values for kVCT and sMVCT divided by the respective MVCT SNR

values were calculated and tabulated in Table 6.2. Generally, kVCT SNR values were roughly an order of magnitude higher than MVCT SNR values. sMVCT SNR values were similar (albeit slightly lower) to MVCT SNR values. Thus, the lowest achievable MAE for two scans with MVCT like noise (in soft tissue/water) rises to 62 HU assuming uncorrelated noise between the scans. In both cases, noise would be a dominant contributor to MAE values with a noisier sMVCT (relative to kVCT) potentially overshadowing improvements.

Table 6.2. Similarity metrics for kVCT and sMVCT (versus MVCT).

	kVCT	sMVCT (CycleGAN)
ME soft tissue (HU)	11	2
ME bone (HU)	327	3
MAE soft tissue (HU)	66	61
MAE bone (HU)	340	97
RMSE soft tissue (HU)	118	105
RMSE bone (HU)	270	160
ME soft tissue (ρ_e)	-0.012 (1.2%)	-0.0035 (0.35%)
ME bone (ρ_e)	0.057 (3.8%)	0.0020 (0.13%)
MAE soft tissue (ρ_e)	0.057	0.057
MAE bone (ρ_e)	0.099	0.095
RMSE soft tissue (ρ_e)	0.16	0.10
RMSE bone (ρ_e)	0.35	0.13
ME soft tissue (SPR)	0.030 (3.0%)	0.0024 (0.24%)
ME bone (SPR)	0.106 (7.0%)	0.0024 (0.16%)
MAE soft tissue (SPR)	0.068	0.062
MAE bone (SPR)	0.131	0.106
RMSE soft tissue (SPR)	0.11	0.11
RMSE bone (SPR)	0.16	0.16
SNR vs MVCT SNR soft tissue (ratio)	9.1	0.95
SNR vs MVCT SNR bone (ratio)	11.1	0.95
SSIM	0.856	0.862
PSNR	25.8	29.8

A heatmap plot of HU, ρ_e, and SPR values at each voxel in a given kVCT or sMVCT scan were plotted against corresponding values at the same voxel in the equivalent MVCT image as shown in

Figure 6.5. Within this type of plot, the ideal correlation of the two scan's HU values would be along a 45° diagonal line. In reality, differences in patient positioning and anatomy between the two scans along with intrinsic voxel level specific noise creates a spread away from this line even if there were no inherent systematic differences in HU values.

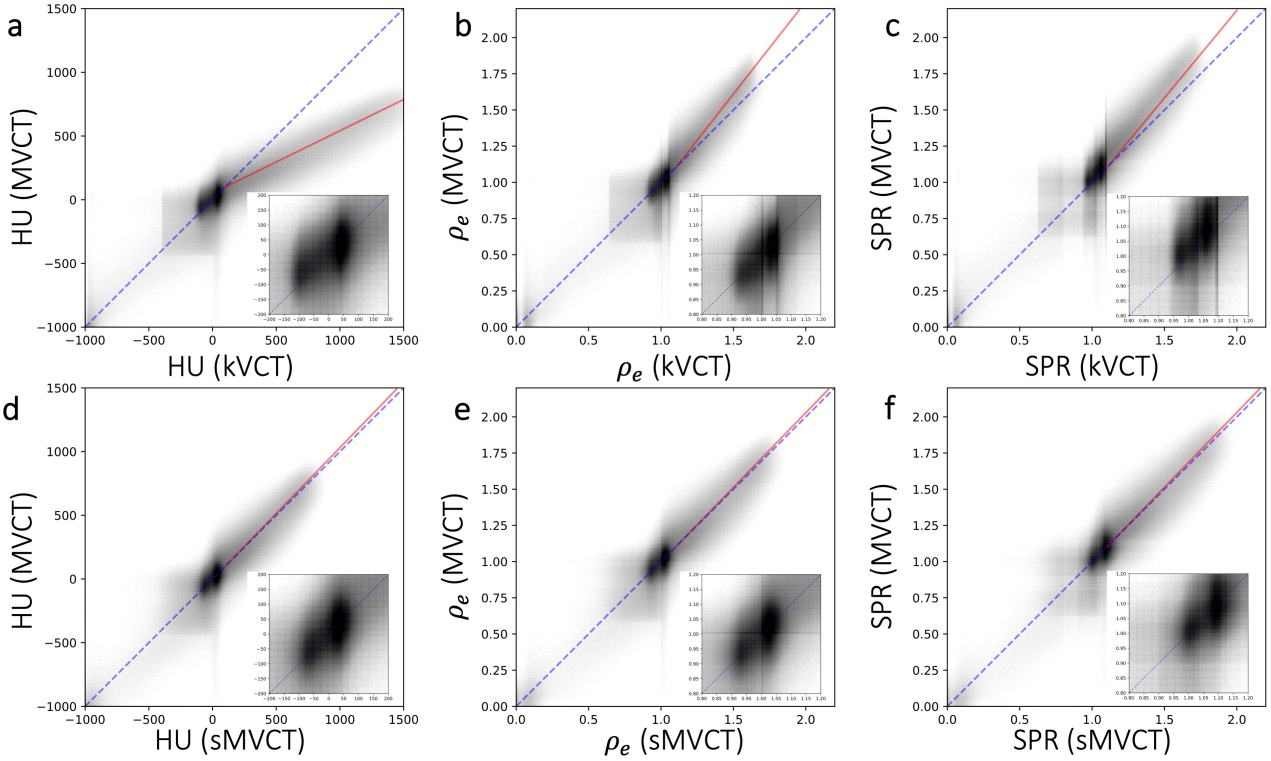

Figure 6.5. Heatmap plots
A heatmap plot of HU (left-most column), ρ_e (middle colum), and SPR (right-most column) values at each voxel in a given CT scan plotted against HU values at the same voxel in a different CT scan for kVCT vs MVCT (top row) and b) sMVCT vs MVCT (bottom row). Grey scale values were plotted on a log scale with numbers of voxel. The dotted blue diagonal line in each panel represents a 45° reference line. The solid red line represents a linear fit of the bone data between the two image sets. In the lower righthand corner of each panel is an inset image of the -200 to 200 HU, 0.8 to 1.2 ρ_e values, and 0.8 to 1.2 SPR values expanded for greater clarity.

As predicted, the relationship between kVCT HU and MVCT HU (as in Figure 6.5a) is non-linear with what appears to be at least 3 distinct regions with differing slopes: 1) low density tissues, 2) soft tissues, and 3) bony tissues. For the bony region, the deviation away from the reference line is about 19° (shown

as a solid red line in the figure). Qualitatively, the relationship between kVCT HU and MVCT HU resembles that of the kVCT calibration curve from Figure 3.2. In Figure 6.5d, sMVCT HU value are plotted against MVCT HU values and appear to agree closely with one another, as indicated by the coincidence between red and blue lines, demonstrating that kVCT HU can be mapped to a corresponding sMVCT HU map that agrees well with MVCT HU. In the bony region, the deviation away from the reference line is minimal (< 1°) relative to the 19° deviation seen for kVCT images, and is shown as a solid red line in the figure. The overall trend in the data is consistent with there being an atomic number dependent effect, as *in vivo* tissues with higher Z will exhibit a higher HU on kVCT and a lower HU on MVCT.

HU values from each kVCT or MVCT scan were converted to ρ_e with the HU-to-electron density calibration curve depicted in Figure 3.2. Heatmap plots of ρ_e (Figure 6.5, middle column) at each voxel in a given kVCT or sMVCT scan were plotted against corresponding ρ_e values at the same voxel in the analogous MVCT scan (equivalent slice). In Figure 6.5b (middle, top), we see noticeable differences between ρ_e values of kVCT versus MVCT. Specifically, ρ_e(kVCT) values differ from ρ_e(MVCT) the most for higher density bone with this deviation appearing to increase with increasing electron density values. Quantitatively, soft tissues with ρ_e between 0.95 and 1.05 exhibited an average deviation away from the 45° reference line of < 0.01. For higher density tissues, there appears to be a large jump in the average deviation away from the 45° reference line. In the bony region, the red fit line in Figure 3.3b shows a 7° deviation away from the 45° reference line as ρ_e increases. ME values (tabulated in Table 6.2) between ρ_e(kVCT) and ρ_e(MVCT) were -0.012 (1.2%) for soft tissues and 0.057 (3.8%) for bone. MAE values (tabulated in Table 6.2) between ρ_e(kVCT) and ρ_e(MVCT) were 0.057 for soft tissues and 0.099 for bone. RMSE values (tabulated in Table 6.2) between ρ_e(kVCT) and ρ_e(MVCT) were 0.16 and 0.35 for soft tissue and bone, respectively. In Figure 6.5e, ρ_e (sMVCT)

values were plotted against ρ_e (MVCT) values, representing the ability of the developed CycleGAN machine learning model to accurately convert kVCT scans to synthetic MVCT scans. As expected, there is uniform and marked improvement in the alignment of ρ_e (sMVCT) values with ρ_e (MVCT) values. The heatmap, as a whole, appears to better align with the 45° reference line than in Figure 6.5b. In the bony region, the red fit line in Figure 6.5b shows a < 1° deviation away from the 45° reference line. Quantitatively, average deviations of ρ_e from the 45° reference line were all < 0.01. ME values (tabulated in Table 6.2) between ρ_e(sMVCT) and ρ_e(MVCT) were considerably lower at -0.0035 (0.35%) for soft tissues and 0.0020 (0.13%) for bone. MAE values (tabulated in Table 6.2) between ρ_e(sMVCT) and ρ_e(MVCT) were 0.057 for soft tissues and 0.095 for bone. RMSE values (tabulated in Table 6.2) between ρ_e(sMVCT) and ρ_e(MVCT) improved, relative to ρ_e(kVCT), to 0.10 and 0.13 for soft tissue and bone, respectively.

HU values from each kVCT or MVCT scan were converted to SPR with the HU-to-SPR calibration curve depicted in Figure 3.4. Heatmap plots of SPR (Figure 6.5, right column) at each voxel in a given kVCT or sMVCT scan were plotted against corresponding SPR values at the same voxel in the analogous MVCT scan (equivalent slice). In Figure 6.5c (right, top), we see noticeable differences between SPR values of kVCT versus MVCT. Specifically, SPR (kVCT) values differ from SPR (MVCT) the most for higher density bone with this deviation appearing to increase with increasing SPR values. Quantitatively, soft tissues with SPR between 0.95 and 1.05 exhibited an average deviation away from the 45° reference line of < 0.01. Similar to the trends seen for ρ_e, for higher density tissues there appears to be a large jump in the average deviation away from the 45° reference line. In the bony region, the red fit line in Figure 6.5b shows a >5° deviation away from the 45° reference line as SPR increases. ME values (tabulated in Table 6.2) between SPR (kVCT) and SPR (MVCT) were 0.030 (3.0%) for soft tissues and 0.106 (7.0%) for bone. MAE values (tabulated in Table 6.2) between SPR

(kVCT) and SPR (MVCT) were 0.068 for soft tissues and 0.131 for bone. RMSE values (tabulated in Table 6.2) between SPR (kVCT) and SPR (MVCT) were 0.11 and 0.16 for soft tissue and bone, respectively. In Figure 6.5f, SPR (sMVCT) values were plotted against SPR (MVCT) values, representing the ability of the developed CycleGAN machine learning model to accurately convert kVCT scans to synthetic MVCT scans. As expected, there is uniform and marked improvement in the alignment of SPR (sMVCT) values with SPR (MVCT) values. The heatmap, as a whole, appears to better align with the 45° reference line than in Figure 6.5c. In the bony region, the red fit line in Figure 6.5b shows a < 1° deviation away from the 45° reference line. Quantitatively, average deviations of SPR from the 45° reference line were all < 0.01. ME values (tabulated in Table 6.2) between SPR (sMVCT) and SPR (MVCT) were considerably lower at 0.0024 (0.24%) for soft tissues and 0.0024 (0.16%) for bone. MAE values (tabulated in Table 6.2) between SPR (sMVCT) and SPR (MVCT) were 0.062 for soft tissues and 0.106 for bone. RMSE values (tabulated in Table 6.2) between SPR (sMVCT) and SPR (MVCT) were 0.11 and 0.16 for soft tissue and bone, respectively.

6.3 Evaluation of anthropomorphic head phantom

To independently verify the accuracy of bone electron density and SPR of kVCT, MVCT, and synthetic MVCT images produced by the developed CycleGAN model, electron density and SPR were analyzed in the anthropomorphic head phantom described in Chapter 3.2 and shown in Figure 6.6. kVCT and MVCT images were acquired of the head phantom using the same protocols as described for the *in vivo* data. The kVCT data was run through the CycleGAN model to generate a synthetic MVCT scan. Bone was segmented using intensity-based thresholding in the kVCT and copied to the co-registered MVCT and sMVCT datasets. HU values (mean and standard deviation) for cortical bone were extracted and mapped to corresponding ρ_e and SPR values using the CT calibration curves shown in

Figures 3.2 and 3.4, respectively. Values of ρ_e and SPR estimated from kVCT, MVCT, and sMVCT were compared to the physical measurements tabulated in Table 6.3.

Figure 6.6 shows exemplar images for kVCT HU, sMVCT HU, MVCT HU, and difference heatmaps of kVCT HU and sMVCT HU compared to MVCT HU for axial planes at the level of the jaw and skull for the anthropomorphic head phantom. Qualitatively, trends seen in the head phantom were similar to *in vivo* data, where the largest differences were seen between kVCT and MVCT in bony tissues. As expected, kVCT images exhibited higher HU values than those obtained from corresponding MVCT images, leading to an overestimation of electron density and SPR. sMVCT images, on the other hand, appeared qualitatively more similar to MVCT images, as shown in the 5th and 6th columns of Figure 6.6. While the HU distribution of the sMVCT is more similar to MVCT than the kVCT is, there remains some level of overestimation of the sMVCT HU values that exists in the anthropomorphic phantom results but not the *in vivo* results. Because the model was trained on *in vivo* data, it may not be ideally suited for phantom studies despite the bone material being reasonably mimicking of *in vivo* bone composition and density.

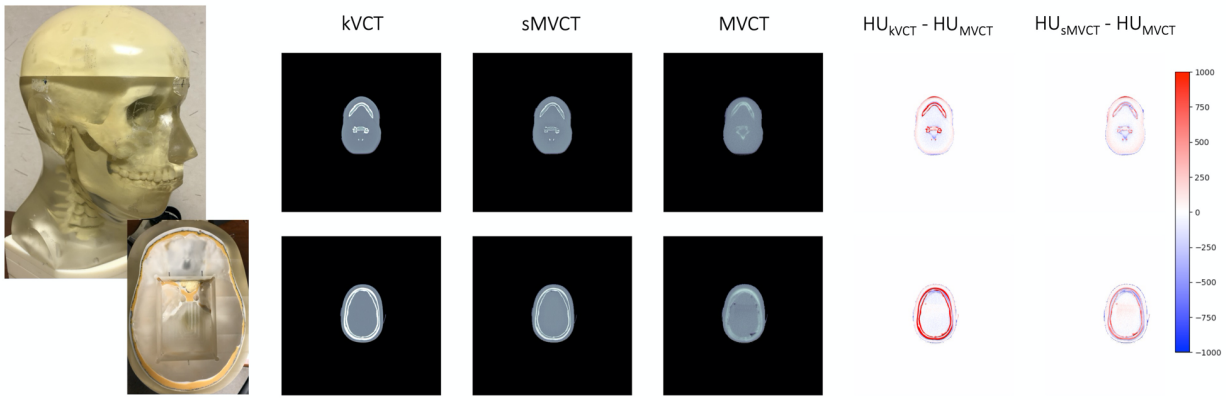

Figure 6.6. Anthropomorphic head phantom results
Exemplar images of the anthropomorphic head phantom (1st column) imaged using kVCT (2nd column), sMVCT HU (3rd column), MVCT HU (4th column), and HU difference heat maps for kVCT versus MVCT (5th column) and sMVCT versus MVCT (6th column) shown at axial levels of the skull (top row) and jaw (bottom row). The HU scale used in the difference maps is shown on the right.

Quantitative results of relative electron density and stopping power ratio calculated from physical measurements and estimated through kV and MV CT calibration curves for bone material in the anthropomorphic head phantom are shown in Table 6.3. Percent differences between kV and MV imaging and physical measurements (% diff) are also shown.

Table 6.3: MVCT and kVCT values for anthropomorphic head phantom
Values of relative electron density and stopping power ratio derived from physical measurements of bone phantom versus kVCT, MVCT, and sMVCT imaging.

	Relative electron density				Stopping power ratio (115 MeV)			
	Measured	kVCT (% diff)	MVCT (% diff)	sMVCT (% diff)	Measured	kVCT (% diff)	MVCT (% diff)	sMVCT (% diff)
Bone (Head Phantom)	1.12	1.167 ± 0.10 (4.16%)	1.122 ± 0.13 (0.18%)	1.131 ± 0.20 (0.98%)	1.13	1.198 ± 0.09 (6.05%)	1.139 ± 0.11 (0.80%)	1.147 ± 0.17 (1.50%)

These results were consistent with those seen in the tissue mimicking phantom. Electron density values of the bone computed from kVCT images deviated from physical measurements by 4.16%. This deviation decreased to 0.18% and 0.98% when estimated from MVCT and sMVCT, respectively. Stopping power ratio values of the bone computed from kVCT images deviated from physical measurements by 6.05%. This deviation decreased to 0.80% and 1.50% when estimated from MVCT and sMVCT, respectively.

6.4 Discussion

The relationships between kVCT HU and ρ_e or SPR have historically been modeled as piecewise linear. In reality, the relationships between kVCT HU and ρ_e or SPR are more complicated due to the intrinsic non-linearity and no explicit one-to-one relationship between these values when imaged at kV energies. In this manuscript, we used a tissue mimicking phantom to demonstrate that higher accuracy mapping between HU to ρ_e or HU to SPR can be achieved when using MVCT instead of kVCT. We

then used a method to empirically learn the relationship between kVCT HU and MVCT HU through CycleGAN deep learning. In doing so, we intrinsically linearize and improve the accuracy of the relationship between kVCT HU with ρ_e and kVCT HU with SPR through an MVCT surrogate. Finally, the cycleGAN model was independently validated in an anthropomorphic head phantom containing a tissue mimicking human bone material.

For *in vivo* data, the ρ_e value differences between the two scans (kVCT and MVCT) were spatially and organ/tissue dependent with ME values of 0.012 (1.2%) and 0.057 (3.8%) for soft tissue and dense bone, respectively. The ρ_e values for tissues with high water content (such as orbits and parotids) exhibited little change in ρ_e (< 0.01) from kVCT to sMVCT with ρ_e values being close (within 0.005) to the expected value of 1.0 in both cases. It is possible that because liquid water was a material used for generating CT calibration curves that these exact HU to ρ_e and SPR relationships may have reduced uncertainty associated with them. Similarly, subcutaneous fat tended to have, on average, subtle changes in ρ_e (< 0.01) from kVCT to sMVCT. In contrast, a significant portion of soft tissues/organs with higher protein levels such as skin exhibited relatively larger negative changes in ρ_e of about -0.03 (from kVCT to sMVCT). For kVCT HU > 0, we see that there is greater uncertainty in the relationship between kVCT HU and ρ_e (as in Figure 6.5b) and kVCT HU and SPR (as in Figure 6.5c) with increasing non-linearity in this region in both curves. Similarity metrics (shown in Table 6.2) for ME, MAE, RMSE, SNR, PSNR, and SSIM showed overall improvements in agreement with MVCT compared with initial kVCT. In our work, SSIM and PSNR values were found to be 0.862 and 29.8, respectively. Our unpaired results are comparable to the work of Liang *et al*[108] who achieved 0.85 and 30.65 for SSIM and PSNR, respectively, for kV cone-beam CT to kVCT image translation.

In this study, we used a tissue mimicking phantom to evaluate the accuracy in predicting ρ_e/SPR from MVCT HU conversion compared to kVCT HU conversion. For all tissue mimicking phantom materials examined in this work, MVCT derived ρ_e/SPR values agreed with physical

measurements to within 1%, with the exception of SPR in adipose which differed by 1.7%. While still more accurate than kVCT derived SPR (which differed by 2.58%), both kV and MV SPR errors were higher than their corresponding ρ_e errors. In the Bethe-Bloch equation there are only two parameters that are tissue specific: ρ_e and I_m. In cases where ρ_e is accurate but SPR is inaccurate, there is the implication that uncertainties in I_m are the cause. In our prior work using the same tissue mimicking phantom[56], stoichiometric calibration curves were created using I_m values published in 1982[53]. More recently, I_m values optimized for use with the Bragg additivity rule were published by Bär et al[54] and differed from previous values by ~3% in adipose tissues and ~4.5%. This work demonstrated that outdated I_m values may lead to overestimation of range uncertainty and should be reassessed in clinical practice. This finding is consistent with the work of Paganetti *et al*[15] which reported that I_m was a major contributor to the uncertainties in SPR determination, with uncertainty of I_m values lowering the achievable accuracy of CT derived SPR.

In bony tissues specifically, we see the greatest amount of differences in kVCT relative to MVCT. In the spongiosa bone phantom, errors in ρ_e/SPR values were reduced from 2.20%/4.38% to 0.19%/0.22% for kVCT and MVCT, respectively. These findings are also in line with the bony tissue errors found in the anthropomorphic head phantom whereby ρ_e/SPR values significantly improved in accuracy from kVCT to sMVCT. This finding is consistent other groups who have investigated the uncertainty in proton therapy dose calculation due to kVCT-based SPR predictions, whereby they found that SPR in bone exhibited greater uncertainty than SPR in soft tissue[13]. It is also in line with our group's finding[28] that ambiguities in bony tissue elemental and molecular composition relative to nominal tissue compositions would lead to greater uncertainties in the prediction of ρ_e and SPR based on the conversion of kVCT HU. As this error/uncertainty in bone is systematic and increases with ρ_e/SPR, it is possible that the composition of the bone tissue surrogate materials may not closely mimic that of our subject population. As adults age, there is an age dependent bone change that results in

increasing mineral content with decreasing protein content[110] potentially accounting for this difference. Additional possible explanations for the large difference between kVCT and MVCT predicted ρ_e and SPR may be due to a systematic bias in the overall elemental composition of the bone tissue surrogate used in our study. In the work of Woodard and White[111], they found that bone composition varies significantly across bone type. White *et al*[55] further showed that even within just cortical bone, the composition varies across age. Within the reference material provided by tissue surrogate manufacturer, it is noted that the cortical bone reference is based on ICRU 44[50]. It is likely that the bone composition of tissue surrogates is compositionally different in terms of atomic composition from adult cancer patient head and neck bones, and possibly biased towards the composition of other bones in the body or for other age groups.

Uncertainties in SPR are especially important for charged particle therapies such as proton therapy, where the particle range uncertainty is directly proportional to the SPR uncertainty. For the *in vivo* human data, we showed that kVCT SPR on average deviated from MVCT SPR by 0.030 (3.0%) and 0.106 (7.0%) for soft and bony tissues, respectively, which would directly translate to charged particle range uncertainties of 3.0% and 7.0%. These results are consistent with the clinical range uncertainty budgets (typically an average over all tissues types of 2.5-3.5% the proton beam range) used by proton therapy groups to account for these kVCT-to-SPR uncertainties. In comparison, sMVCT SPR deviated from MVCT SPR with substantially lower ME value differences of 0.0024 (0.24%) and 0.0024 (0.16%) for soft and bony tissues, respectively. In the case of prostate proton therapy treatments, a range uncertainty margin of as much as 1cm is typical when using kVCT-based SPR prediction methods. Utilizing the proposed method to reduce CT-based SPR prediction errors from 2.5-3.5% to <0.5% could result in significant reductions in normal tissue irradiations and potentially allow for dose escalation of tumors. With respect to margins used for prostate radiotherapy, results from Bortfelt et al[112] suggest that total safety margins can be reduced to about 6mm when using

DECT-based SPR predictions in treatment planning. With a mean range prediction accuracy of 0.0%±0.5% demonstrated using their DECT-based technique, we would expect similar margin reductions using our approach. Overall, our proposed methodology provides a unique perspective into areas of greater SPR uncertainty and potentially increased accuracy of SPR determination when using sMVCT.

Another potential benefit of translating the relationship between kVCT imaging to MVCT imaging is in the reduction of beam hardening artifacts caused by high-Z materials such as dental implants. Reducing these artifacts could have significant benefits for proton therapy given that dosimetric impacts caused by artifacts or metallic objects can cause more pronounced uncertainties[113,114]. In Figure 6.7, we show example images (MVCT, sMVCT, and kVCT) from a patient with dental implants. The sMVCT image better mimics the MVCT image in that there is a reduced streaking artifact in both the dental area itself and in the surrounding soft tissue. Of note, because dental artifacts were not found in all patient scans, it is possible that this behavior was observed in the training dataset but not fully learned, resulting in only partial reduction of beam hardening/streaking artifacts in the sMVCT images. While the machine learning code could be further optimized to better correct for streaking artifacts, this development is outside the scope of the current work.

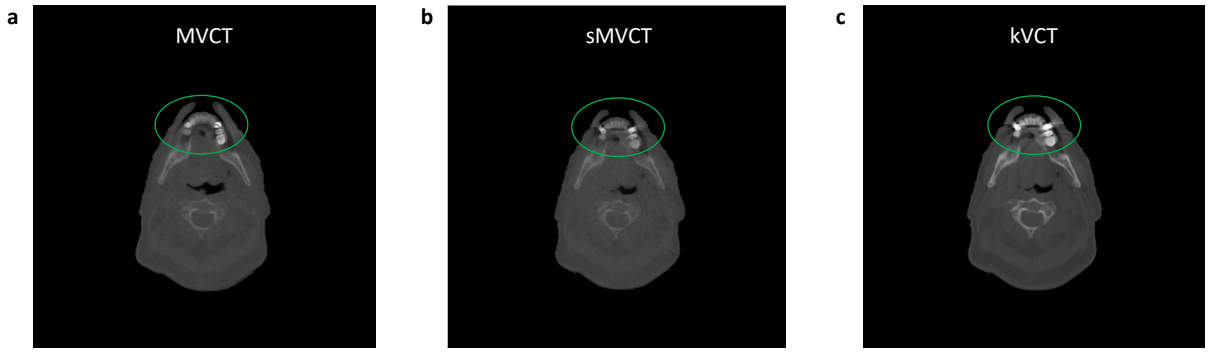

Figure 6.7. Streaking artifacts
Exemplar images of a streaking/beam hardening artifact around a dental implant in (a) MVCT HU (1st column), (b) sMVCT HU (2nd column), and (c) kVCT HU (3rd column. A linear grayscale colormap was used with -1000 HU as the minimum (black) and the maximum HU of the kVCT image as the maximum (white) and applied to the kVCT, MVCT, and sMVCT images.

While results of the anthropomorphic phantom illustrate that sMVCT images can estimate ρ_e and SPR more accurately than kVCT images, there are several limitations inherent to phantom analysis used in this context. The first limitation is that physical measurements of the phantom are derived from small samples whose exact elemental composition may not reflect the variability in material distribution or imperfections caused during manufacturing. Commercially available phantoms are meant to generally mimic material electron density for purposes of dose calculation and geometric quality assurance, however, it is not routine practice to rely on composition at the elemental level. The head phantom is an over-simplified model containing only a small number of tissue types, and while its shape resembles a human head and includes a reasonable bone surrogate material, the non-bone phantom materials did not mimic molecular properties of biological tissues. Therefore it is unlikely that a model trained on *in vivo* data will show identical trends when applied to a phantom made of non-biological material. CycleGAN (and CNNs in general) learn spatial patters and relative local intensity distributions and it is conceivable that nearby non-biological material may bias local distributions, which may have contributed to the slight overestimation of HU values in sMVCT images relative to

MVCT images. An in-depth analysis validating machine learning models using phantoms is outside the scope of this work, however, phantom development remains an important and active area of research where detailed understanding and accurate characterization is imperative as synthetic images become used more frequently in the clinic.

There are some additional considerations for groups considering clinical implementation of this method for kVCT HU-to-ρ_e or HU-to-SPR conversion. Currently, TPSs accept CT HU as input for dose calculation with conversion of these HU values to ρ_e or SPR through calibration curves similar to those shown in Figures 3.2 and 3.4. This process could be further simplified if the input datasets were direct maps of ρ_e or SPR rather than HU. With our methodology, no modification would have been necessary in our learning process to accomplish this. During the pre-processing stage of machine learning, all CT images (both kV and MV) were scaled from HU to values between [-1, 1]. As MVCT HU values were converted to ρ_e or SPR using a linear scale, the learning process would have been identical (as they would have scaled to equivalent values in this [-1, 1] scale). On the post-learning end, we could have created a linear mapping from [-1, 1] directly to ρ_e or SPR values. Within the TPS, the calibration curve would then be a simple linear mapping. Another consideration for clinical implementation of this method is generalizability of a developed model across kVCT models, scanner protocols, and kVCT energies. In general, developed CycleGAN models converting from kVCT HU to sMVCT ρ_e/SPR would be most accurate if created for a specific center/kVCT scanner and for a specific imaging protocol. For centers with a machine capable of creating MVCT scans and kVCT scans, it would be possible to create a dataset of matched patient kVCT and MVCT scans. This matched dataset could be used with deep learning (as in this work) to learn the relationship between kVCT and sMVCT ρ_e/SPR. For centers without MVCT capabilities, implementation of this methodology may be less straightforward as the created deep learning model would likely be specific (and therefore most accurate) for a given kVCT scanner and/or energy spectrum. As our patient population only received

kVCT scans from a specific CT scanner, we were not able to test the effect of scanner model/energy spectrum variability. Future work in this area should include tests on generalizability and stability across scanner/energy spectrum of the created CycleGAN model.

6.4 Conclusion

In this study, deep learning was used to map the relationship between kVCT HU and MVCT HU and an approach was developed to calculate more accurate values of ρ_e and SPR based on MVCT HU. The accuracy of ρ_e and SPR values calculated using MVCT and kVCT were quantified using tissue mimicking phantoms of known material composition. The deep learning model was also validated using an anthropomorphic head phantom with physical characteristics closely resembling human bone. Synthetic MVCT scans generated by our method appear to closely mimic acquired MVCT scans with even some mitigation of beam hardening artifacts. SPR values estimated from sMVCT data produced by the deep learning model resulted in substantially lower mean error values for soft and bony tissues versus kVCT data. If sMVCT scans are used in conjunction with the originally acquired kVCT, they, together, allow for an appealing scenario in which the advantages of each modality can be fully utilized.

Chapter 7 – Clinical evaluation of a commercial MRI-based synthetic CT

Patients being treated with SBRT to the prostate routinely undergo MRI for purposes of anatomical delineation. The soft tissue contrast provided by MRI is particularly well suited for visualizing prostatic lesions and critical structures such as the urethra[115,116] while also decreasing interobserver variability[39] versus CT alone. Replacing CT with MRI-based synthetic sCT for dose calculation would eliminate uncertainties related to multi-modality image registration and offer numerous workflow-related advantages. The goal of this work was to implement and assess a commercially available MRI-derived sCT protocol for treatment planning with two common modalities used to treat prostate cancer, namely,

VMAT delivered on a conventional c-arm linear accelerator and the CyberKnife system (Accuray, Sunnyvale, CA), a compact linear accelerator mounted onto a robotic base. This investigation was performed on the dataset of 10 patients treated using SBRT for prostate cancer at UCSF as described in Chapter 3.4. This study provides recommendations for optimizing the sCT calibration curve for consistency with the users' clinically commissioned CT scanner and is the first to report on utilizing sCT for treatment planning on CyberKnife.

7.1 Vendor solution

To provide a framework for MRI-only treatment planning, an FDA-approved protocol for generating synthetic CT images from MRI in the pelvis has been developed by Siemens[117,118]. The protocol requires a set of Dixon volumetric interpolated breath-hold examination (VIBE) in-phase and out-of-phase images that were acquired in the axial plane using a 3D T_1-weighted gradient echo sequence with 3D distortion correction applied. The bulk assignment algorithm classifies material as air, adipose, soft tissue, inner bone or outer bone corresponding to HU values of -1000, -75, 0, 204, and 1170, respectively. Details of the algorithm are kept proprietary, but the primary concept is that tissue and adipose are classified using spectral information, air using thresholding, and bone using a multi-atlas-based model based on the in-phase, out-of-phase, fat-only, and water-only MRIs produced by the Dixon sequence.

7.2 sCT versus CT comparison

To compare the sCT bulk tissue HU values with those seen *in vivo* for the 10 datasets evaluated, air, adipose, soft tissue, inner bone and outer bone were segmented on each patients' CT. Intensity-based thresholding was used to define each region, with a final binary mask created for each tissue type based on the initial thresholding and manual contouring in MIM, with an example shown in the top row of Figure 7.1. For each patient, the binary masks were used to define corresponding structures

representing adipose, soft tissue, inner bone, and outer bone, as shown in the bottom row of Figure 7.1. HU values (mean and standard deviation) were computed and averaged across all patient CT datasets.

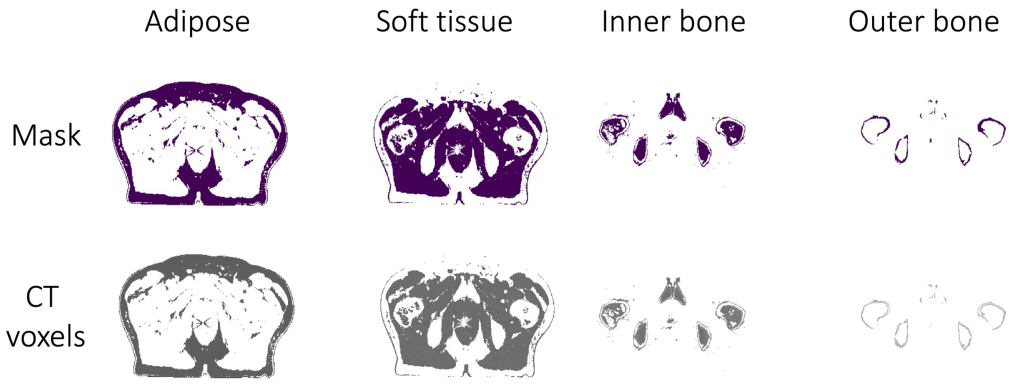

Figure 7.1: CT-derived segmentation
CT-derived segmentation of adipose, soft tissue, inner bone, and outer bone in an example patient. CT HU values (mean and standard deviation) were compared to sCT HU values.

HU values for air, adipose, soft tissue, inner bone, and outer bone provided in the literature[119], extracted from patient CT datasets, and MRI-generated sCT datasets are tabulated in Table 7.1.

Table 7.1: HU values (CT vs. sCT)
Tabulated HU values for air, adipose, soft tissue, inner bone, and outer bone provided in the literature119, extracted from in vivo CT datasets, and assigned automatically to sCT datasets.

	Literature	CT	sCT
Air	-1000	-1000	-1000
Adipose	-100	-96 ± 6.1	-75
Soft tissue	30-45	37 ± 3.9	0
Inner bone	200-800	219 ± 14	204
Outer bone	>1000	1000 ± 23	1170

Of note, sCT HU values for adipose and soft tissue disagree with those for CT and are outside the range of values provided in the literature. Furthermore, adipose and soft tissue structures typically make up the largest volume of the pelvis datasets, as shown in Figure 7.1, and mischaracterization of these HU values could lead to incorrect effective beam path lengths computed in the dose calculation algorithms

(demonstrated in section 7.4). To visualize these discrepancies, difference HU maps were computed by subtracting the sCT from its corresponding co-registered CT. These maps are illustrated in Figure 7.2, which shows three exemplary datasets of CT (first column), sCT (second column), the original HU difference map (third column), and a modified HU difference map that was created by manually overriding the sCT values for adipose and soft tissue to match the CT (fourth column). In other words, in the modified sCT HU difference map, the sCT bulk classification for air, adipose, soft tissue, inner bone or outer bone corresponded to HU values of -1000, -96 (instead of -75), 37 (instead of 0), 204, and 1170, respectively. As demonstrated in column D, these modified sCT datasets agreed more closely with *in vivo* CT datasets in soft tissue and adipose regions.

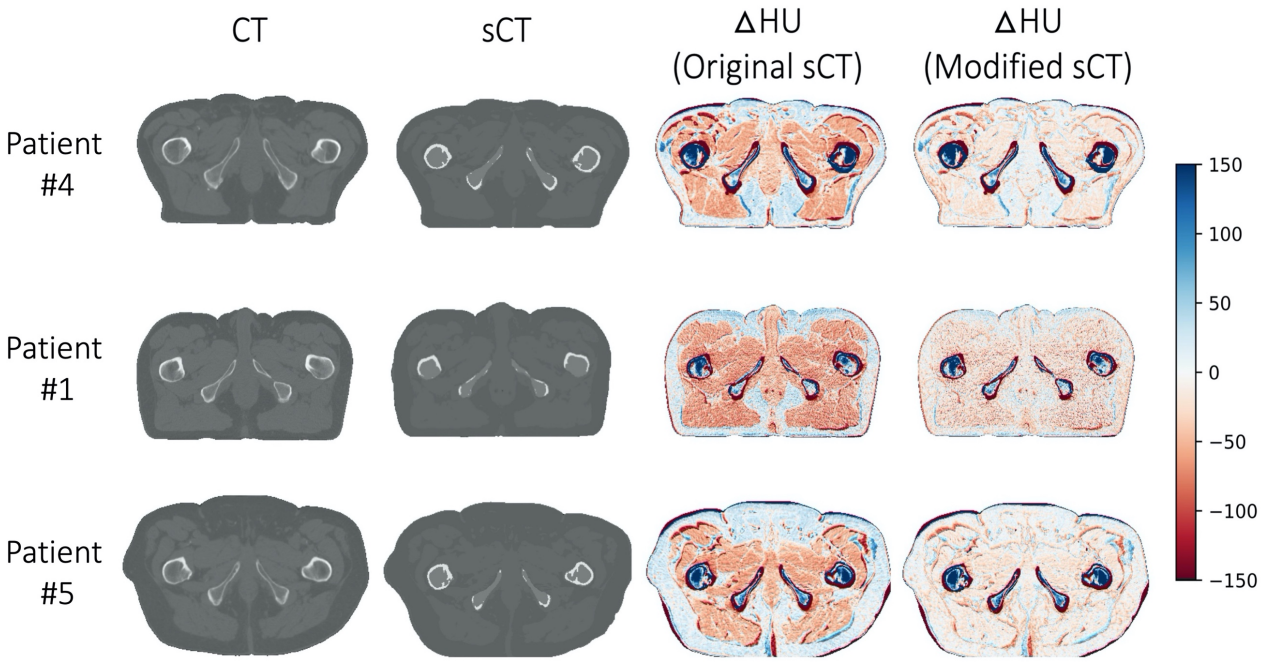

Figure 7.2: sCT versus CT with HU maps
Three exemplary datasets of CT (first column), sCT (second column), the original HU difference map (third column), and a modified HU difference that was created by manually overriding the sCT values for adipose and soft tissue to match the CT (fourth column). HU values are displayed from -150 to 150 to better visualize the differences in seen soft tissue and fat.

CT and sCT HU values were compared using mean absolute error (Equation 12). The total MAE (mean and standard deviation amongst all 10 datasets) was reduced from 106.9±17.9 to 93.9±20.1 HU for the original versus modified sCT, respectively. The soft tissue MAE was reduced from 76.5±7.6 to 50.8±8.1 HU when calculated for the original versus modified sCT, respectively. The adipose MAE was reduced from 96.2±30.2 to 93.6±32.2 HU when calculated for the original versus modified sCT, respectively. Because original sCT values for bone were within the range reported in the literature, this HU value was not modified in the sCT datasets and the total bone MAE is reported to illustrate the variation among patients.

Table 7.2: Tabulated values of MAE for original versus modified sCT
Tabulated values of MAE (HU) for HU difference maps using the original sCT versus the modified sCT for the 10 patients analyzed for this study.

Patient #	Total MAE (HU)		Soft tissue MAE (HU)		Adipose MAE (HU)		Total bone MAE (HU)
	Original	Modified	Original	Modified	Original	Modified	Original only
1	77.8	62.6	73.7	45.5	48.2	43.2	232.5
2	94.0	79.4	74.0	46.4	68.4	63.0	319.8
3	136.5	127.6	71.9	50.6	147.4	150.9	367.5
4	126.2	112.2	90.6	62.4	120.4	117.7	314.4
5	113.4	101.2	85.6	58.9	111.1	108.7	311.5
6	117.8	111.5	83.1	64.0	113.0	113.7	334.9
7	102.4	86.1	75.9	47.8	84.4	80.3	329.4
8	112.5	98.5	72.7	46.8	113.8	106.7	309.4
9	87.5	72.0	64.9	39.7	66.2	62.5	315.8
10	100.6	87.7	72.7	45.7	88.9	89.3	321.1

Areas of disagreement most commonly seen between sCT and CT datasets are depicted in Figure 7.3. For example, differences could be seen in instances of variable bowel and/or rectal gas filling due to anatomical changes between image acquisition (top row). Gold fiducial makers placed in the prostate for purposes of image alignment or calcifications within the prostate gland appeared as hyperintense on CT and hypointense on MRI (and therefore sCT), resulting in additional discrepancies seen between

datasets (middle row). Misalignment of bone due to small differences in pelvic flexion and/or leg positioning was also occasionally seen (bottom row).

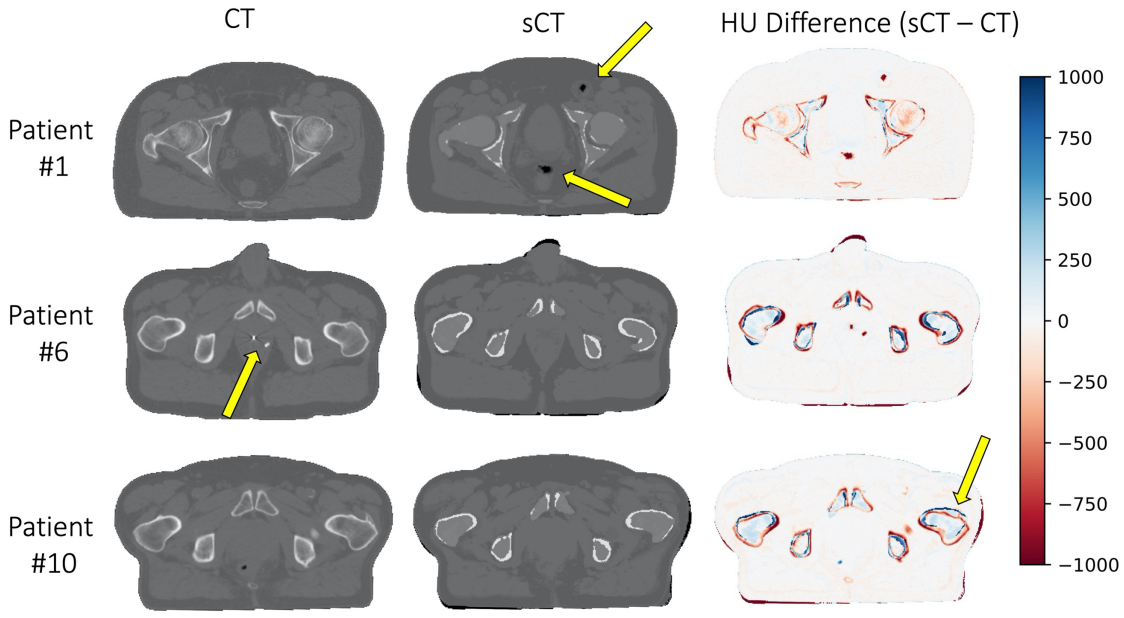

Figure 7.3: sCT versus CT differences
Example axial CT (left), sCT (middle), and HU difference maps (right) for three patients included in this study.

Mislabeling of bone remains an ongoing challenge when generating CT from MRI due to the lack of bone signal produced in most clinical MRI sequences. The most drastic deviation was seen in one patient whose right femoral head was completely displaced (shown in Figure 7.4), which may have been caused by the abnormal shape of the pelvic bones in this patient leading to problems in the atlas-based reconstruction.

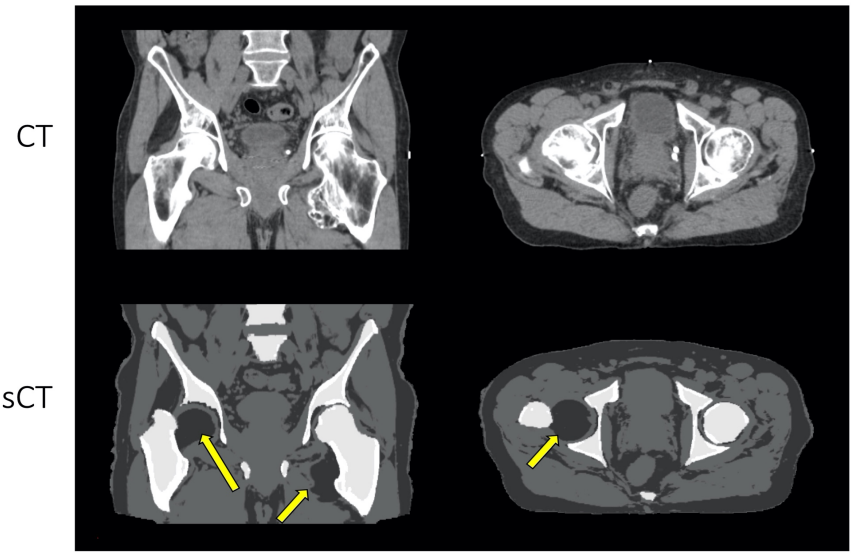

Figure 7.4: Femoral head displacement
Example sagittal and axial slices of CT (top row) and sCT (bottom row) datasets for Patient #9 whose datasets exhibited the largest deviation in bone segmentation, which may have been caused by the abnormal shape of the pelvic bones.

7.3 CT calibration curve

In clinical practice, it is not uncommon to perform manual HU overrides of materials in the treatment planning systems to better match the planning datasets to a patients' true anatomy. For example, this is routinely performed in cases where CTs exhibit artifacts caused by high-Z materials like dental implants[120], when a portion of the patient anatomy is cut off due to insufficient field-of-view[121] due to imaging a patient with large body habitus, or when metallic implants saturate the CT number scale and an appropriate HU value must be manually assigned according to the implant material[122]. In the case of synthetic CT, mitigating HU value differences can similarly be performed by manually assigning HU values of the sCT to better match CT values, as shown in Figure 7.2. However, performing this manual override on each patient dataset in practice would result in logistical challenges related to clinical streamlining and quality assurance. To circumvent these challenges, the same dosimetric outcome can be achieved by using a global sCT-specific CT calibration curve generated using HU vs

mass density values of -1000 vs 0, -96 vs 0.95, 37 vs 1.04, 205 vs 1.1, and 1170 s 1.7g/cc for air, adipose, soft tissue, inner bone, and outer bone, respectively. This optimized sCT calibration curve (with the original sCT curve for comparison) is shown in Figure 7.5.

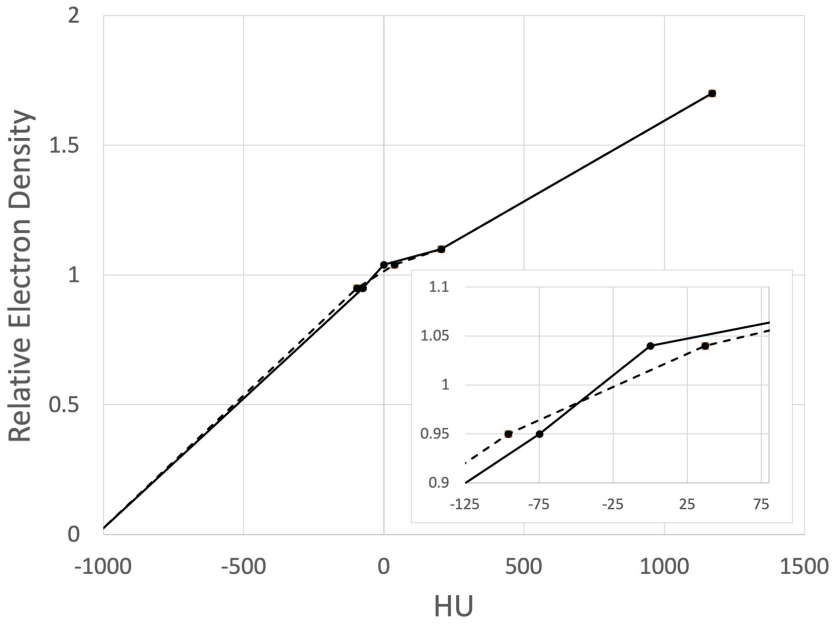

Figure 7.5: Modified sCT calibration curve
Original (solid line) and optimized (dashed line) sCT calibration curves used for dose calculation in the clinical treatment planning systems.

For purposes of treatment planning, the optimized sCT calibration curve was input into RayStation v7.0 (RaySearch Medical Laboratories AB, Stockholm, Sweden) and Accuray Precision v3.1 treatment planning systems.

7.4 Treatment plan evaluation

VMAT treatment plans were generated in RayStation using a two 6MV coplanar arc arrangement and dose calculation performed with a collapsed cone dose calculation algorithm. CyberKnife treatment plans were generated in Precision using a 6FFF non-isocentric beam arrangement and dose calculation performed with a ray-tracing dose calculation algorithm. All plans were generated according to

institutional practice and met clinical objectives. For each patient, treatment plans were generated on CT datasets and recomputed onto corresponding sCT datasets using identical plan parameters. For comparison, dose calculation was performed using both the original and the optimized sCT calibration curves. Dose distributions were compared using gamma analysis and dose-volume-histograms (DVH) of target and critical structures including the bladder, rectum, femoral heads, large bowel, and small bowel.

Results for all plans calculated using the original (non-optimized) sCT calibration curve demonstrated systematic discrepancies between DVHs in sCT versus CT datasets, as shown by non-coincident histograms in the top row of Figure 7.6. This is caused by inaccurate scaling of each beamlets' effective path length caused by incorrect HU values assigned to soft tissue and adipose. However, these discrepancies were resolved when the optimized sCT calibration curve was used, as shown by coincident histograms in the bottom row of Figure 7.6.

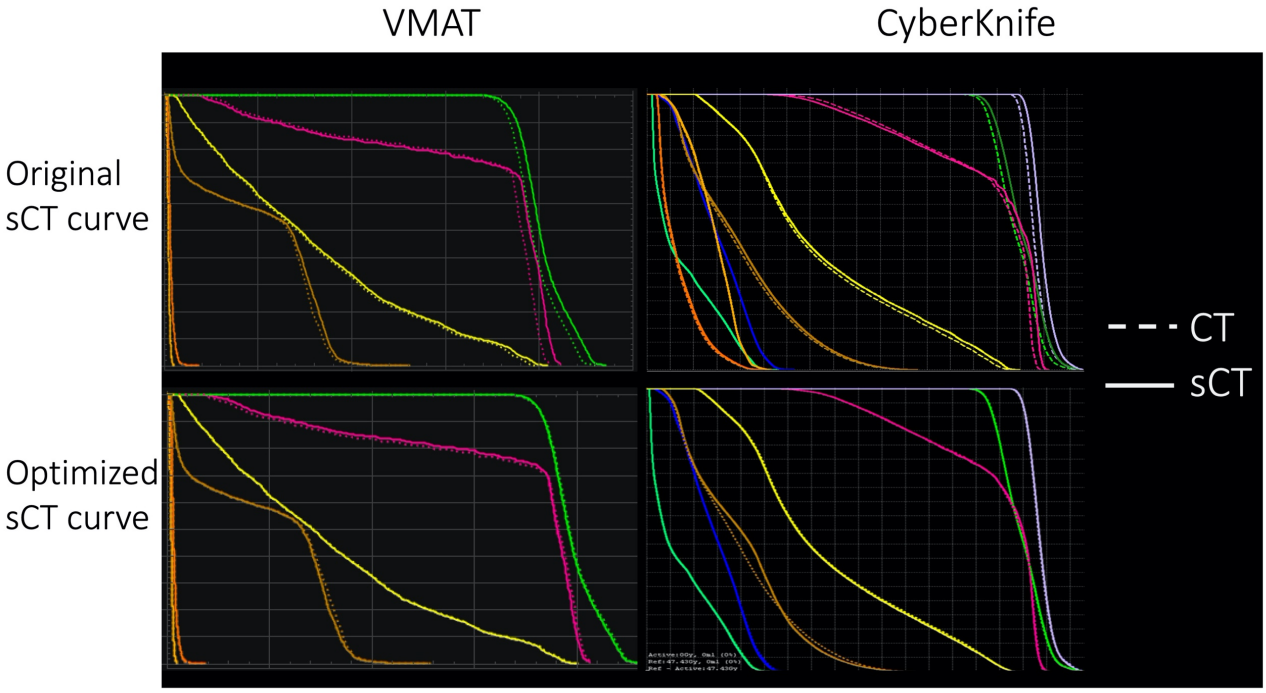

Figure 7.6: Example dose volume histograms
Example DVH plots in sCT (solid line) versus CT (dashed line) datasets demonstrating the systematic discrepancies when using the original CT calibration curve (top row) which are resolved when implementing the optimized CT calibration curve (bottom row) for VMAT (left column) and CyberKnife (right column) treatment plans.

Dose distributions for treatment plans created using the optimized sCT calibration curve were compared by calculating the gamma metric using 3%/3mm local criteria. Gamma metric results (mean and standard deviation amongst all 10 datasets) were 98.9±0.9% (97.1-100%) and 97.7±1.3% (95.3-99.3%) for VMAT and CyberKnife plans, respectively. All gamma results were >95%, indicating good dosimetric agreement between plans calculated on sCT versus CT datasets. Dose to 95% of PTV volume in sCT plans received 100.5±0.8% (99.8-102.5%) and 97.0±6.0% (81.9-103.0%) of the PTV volume in the CT plans for VMAT and CyberKnife plans, respectively. All values of 95% PTV coverage for sCT datasets were within 4% of corresponding CT datasets, with the exception of one

patient (#9) whose PTV in the CyberKnife plan received lower dose coverage in the sCT dataset versus CT dataset due to the mislabeling of the bone/femoral head as previously described.

Table 7.3: Gamma metrics and PTV coverage
Gamma metrics and relative PTV coverage for plans calculated on VMAT and CyberKnife plans for the 10 patients evaluated in this study.

Patient #	Gamma metric (3%/3mm)		Dose to 95% of PTV (sCT relative to CT)	
	VMAT	CyberKnife	VMAT	CyberKnife
1	99.96	99.27	100.1	98.8
2	99.39	96.02	102.5	96.6
3	97.96	98.91	101.0	98.5
4	99.63	98.60	100.1	99.4
5	99.13	97.26	100.4	101.7
6	97.13	95.27	100.1	98.3
7	99.05	98.56	100.2	99.0
8	98.00	96.5	100.5	103.0
9	99.71	98.01	99.8	81.9
10	98.91	98.28	100.6	93.0

Representative axial dose distributions for VMAT (top row) and CyberKnife (bottom row) plans calculated on a CT (left images) versus synthetic CT (right images) are shown in Figure 7.7. Close agreement between DVHs for sCT versus CT can be seen for both plan types, demonstrated by the indistinguishable dose volume histograms.

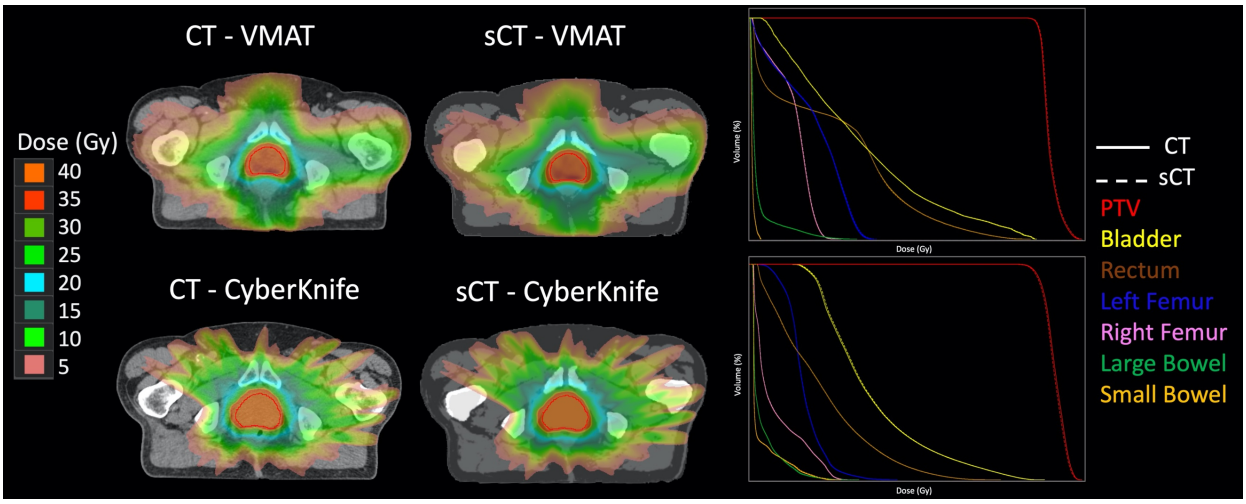

Figure 7.7: Linac-based and CyberKnife treatment plans
Representative axial dose distributions for VMAT (top row) and CyberKnife (bottom row) plans calculated on a CT (left images) versus synthetic CT (right images). DVH curves in each plot (right column) directly overlay one another, demonstrating close agreement between distributions for sCT versus CT.

7.5 Discussion

In this work, a commercially-available MRI-derived sCT protocol was implemented and assessed for treatment planning on a dataset of 10 patients treated for prostate cancer using VMAT and CyberKnife. sCT and CT HU values were compared and a sCT-specific calibration curve was generated to provide better agreement between sCT HU values and treatment plan distributions calculated on sCT versus CT datasets.

HU values for soft tissue and adipose in the original sCT datasets demonstrated disagreement with literature and *in vivo* values. For example, the HU value used for soft tissue was 0 HU, which is the value defined for water. Because the majority of soft tissue in the pelvis is muscle, a value of 30-45 HU is more appropriate. Similarly, with adipose, an HU value of approximately -100 HU is more appropriate than -75 HU. Using these values in sCT datasets results in better agreement between sCT and CT. The total MAE was reduced from 106.9±17.9 to 93.9±20.1 HU for the original versus modified sCT, respectively. The soft tissue MAE was reduced from 76.5±7.6 to 50.8±8.1 HU when calculated

for the original versus modified sCT, respectively. The adipose MAE was reduced from 96.2±30.2 to 93.6±32.2 HU when calculated for the original versus modified sCT, respectively. Overall, these MAE values are consistent with those reported in the literature for atlas-based MRI-derived sCT[123].

In this work, disagreement could be seen in instances of changing bowel/rectal gas filling between scans, fiducials or calcifications appearing in the CT but not in the sCT, and/or bone density distribution differences. Accurate bone delineation remains a challenge for reconstruction of MRI-based sCT, with a single outlier CyberKnife plan for patient #9 yielding a significant under-dose of the target volume ($D_{95\%}$=81.9%) occurring due to a mischaracterization of a femoral head as soft tissue. This case demonstrates the undesirable dosimetric consequences caused by misidentification of bone. When commissioning an sCT dose calculation framework in practice, a judicious comparison of the synthetic CT HU values is imperative. To mitigate any HU differences in practice, a sCT-specific calibration curve can be generated for purposes of sCT-based dose calculation. As demonstrated in this work, using an sCT-specific calibration curve will yield treatment plans of comparable quality to CT-based planning. However, mislabeling of bone can be problematic for treatment planning even when the sCT calibration curve is used.

This project was performed in close collaboration with Siemens scientists and findings have been reported to their research and development team. As of fall 2021, an updated algorithm has been developed, which will be made available to our group with investigations planned for the near future.

7.6 Conclusion

MRI-derived sCT using a sCT-optimized CT calibration curve shows good dosimetry agreement with conventional CT simulation, demonstrating the feasibility of using MRI-derived sCT for prostate SBRT treatment planning. Accurate delineation of bone remains a challenge for reconstruction of MRI-

based sCT and improvements in this space could potentially allow for the implementation of an MRI-only framework.

Chapter 8 – Summary and future directions

8.1 Summary of dissertation

This dissertation presents several novel methods for exploiting the unique advantages of MRI, kVCT, and MVCT to improve the accuracy of photon and proton dose calculation. We have demonstrated improved accuracy in estimating SPR using the UC Method with a combination of MRI and CT, showing that MVCT is superior to kVCT for deriving electron density and stopping power ratio. Expanding on the utility of MVCT for dose calculation in MRI-linac systems, the feasibility of a proposed workflow was demonstrated in which on-board MVCT are used for treatment planning and dose calculation by utilizing a deep learning-based model used to map MRI to sMVCT. In this scenario, the combined set of MRI/sMVCT scans potentially achieve the advantages of both imaging modalities with superior soft tissue contrast (from the MRI) and ρ_e (from the sMVCT). Finally, we used a deep learning model to convert kVCT to sMVCT such that the combined set of kVCT/sMVCT scans achieve the advantages of both imaging modalities with strong contrast and low noise (from the kVCT) and higher accuracy ρ_e and SPR determination (from the sMVCT).

8.2 Future research directions

8.2.1 Clinical implementation of image domain translation

Recent publications on the theory and applications of machine learning to radiation oncology including tumor classification, knowledge-based treatment planning, bioinformatics and modeling of normal tissue complication predictions[124]. Recent work by Mahadevaiah *et al*[125] provided recommendations for adopting artificial intelligence-based systems in the clinic, which included selecting the most appropriate system for the application of interest, testing the security, privacy, and safety of the system, optimization and customization, implementing workflows, and QA monitoring. While guidelines like

these provide excellent starting frameworks, there are very few recommendations for quality assurance and commissioning specific to image-domain translation. Points which must be taken into consideration include which approach is best for a given scenario (for example, an MR-sequence based reconstruction as described in Chapter 7 versus deep learning models as described in Chapters 5 and 6), what are appropriate metric thresholds for assessing synthetic images, how much data is required for model training and testing, how applicable is the model to different disease sites or other institutions, and many others. I look forward to contributing to this area in bringing image domain translation to clinical practice.

8.2.2 MRI-to-MVCT Commercialization

There has been a substantial increase in the use of MRI in radiotherapy due to the superior soft tissue contrast provided relative to CT. During treatment planning, the current paradigm is to acquire an MRI (for anatomy visualization) and CT (for dose calculation) for each patient. During radiation treatment on a combined MRI-linac the current paradigm is to use the MRI acquired on the treatment machine with a CT scan (acquired previously for treatment planning) to adapt a treatment plan to a patient's anatomy, which can change daily. Although MRI-linac technology has gained substantial popularity over the last several years, it remains a new and evolving space with regulatory approval only recently obtained in 2017. The innovation of our work is to shift from the current paradigm of using MRI and kVCT towards a new paradigm of using MRI and MVCT all while using imaging technology available directly on the treatment machines. This would necessitate fewer imaging systems and provide higher accuracy dose calculation. The core idea is composed of two parts (shown in the bottom panel of Figure 8.1):

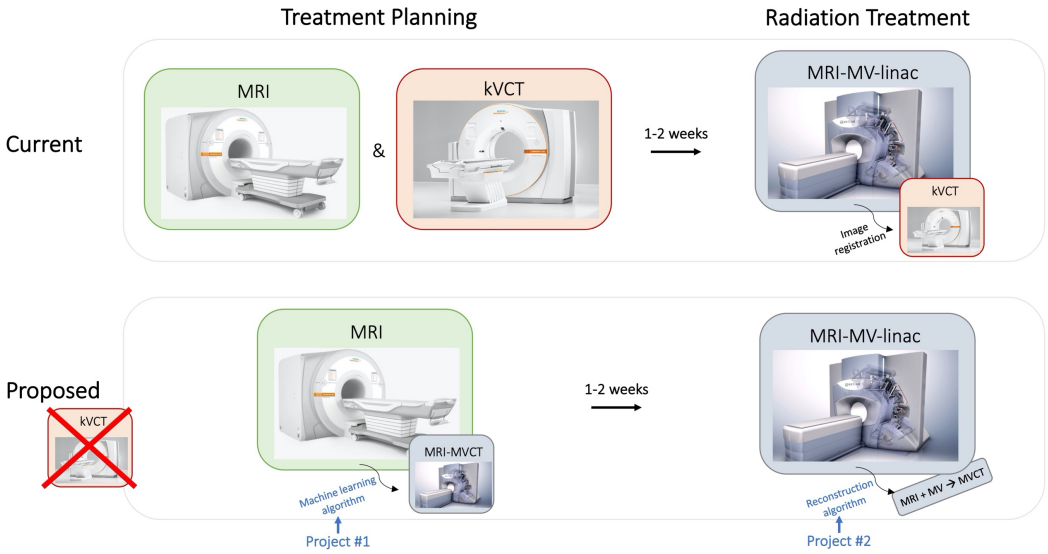

Figure 8.1: Future directions MRI to sMVCT
Current (top) paradigm using MRI and kVCT and proposed (bottom) paradigm for using MRI and MVCT with imaging technology available directly on treatment machines.

In Chapter 5 of this dissertation I described project #1, namely, to synthesize MVCT datasets from MRI using deep learning synthesize MVCT images from MRI using a deep learning model. We demonstrated the feasibility of utilizing MRI-derived synthetic CTs for radiotherapy treatment planning of head and neck cancer. This differs from the current paradigm which uses either an additional kVCT dataset or synthesizes kVCT images from MRI using deep learning or MR-sequence based approaches. For project #2, we propose to reconstruct an MVCT dataset (required for dose calculation) using MRI and 2D MV radiographs that are both acquired on the MRI-linac treatment machine[126]. This approach is more accurate and better streamlined than the current paradigm which

requires registering an MRI to an old kVCT scan. I look forward to partnering with MRI-linac vendors on these efforts.

8.2.3 Proton therapy SPR measurements

In Chapters 4 and 6, I described work that revolved around reducing proton beam range uncertainties through more accurate estimation of SPR. To assess these approaches, tissue substitute and anthropomorphic phantoms were used to compare physical "ground truth" measurements to values predicted from imaging. Through collaboration with proton centers, additional validation can be performed by measuring the SPR of materials through range measurements acquired in a proton beam. For example, the range of a proton beam can be estimated using the imaging approaches described in this work and compared to proton beam measurements acquired using a multilayer ionization chamber[127]. Furthermore, additional dosimetric comparisons can be performed using Monte Carlo simulations[128] and clinical proton treatment planning systems.

Funding

Research reported in this manuscript was supported by the NIBIB of the National Institutes of Health under award number R21EB026086. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Thariat J, Hannoun-Levi J-M, Myint AS, Vuong T, Gérard J-P. Past, present, and future of radiotherapy for the benefit of patients. *Nature reviews Clinical oncology.* 2013;10(1):52-60.
2. Schaue D, McBride WH. Opportunities and challenges of radiotherapy for treating cancer. *Nature reviews Clinical oncology.* 2015;12(9):527-540.
3. Khan FM, Gibbons JP. *Khan's the physics of radiation therapy.* Lippincott Williams & Wilkins; 2014.
4. Hendee WR, Ibbott GS, Hendee EG. *Radiation therapy physics.* John Wiley & Sons; 2013.
5. Salvat F, Fernández-Varea JM. Overview of physical interaction models for photon and electron transport used in Monte Carlo codes. *Metrologia.* 2009;46(2):S112.
6. Knöös T, Wieslander E, Cozzi L, et al. Comparison of dose calculation algorithms for treatment planning in external photon beam therapy for clinical situations. *Physics in Medicine & Biology.* 2006;51(22):5785.
7. Lu L. Dose calculation algorithms in external beam photon radiation therapy. *International Journal of Cancer Therapy and Oncology.* 2014;1(2).
8. Durante M, Loeffler JS. Charged particles in radiation oncology. *Nat Rev Clin Oncol.* 2010;7(1):37-43.
9. Smith AR. Vision 20/20: proton therapy. *Med Phys.* 2009;36(2):556-568.
10. Buzug TM. Computed tomography. In: *Springer handbook of medical technology.* Springer; 2011:311-342.

11. Bushberg JT, Boone JM. *The essential physics of medical imaging.* Lippincott Williams & Wilkins; 2011.

12. Huizenga H, Storchi PR. The use of computed tomography numbers in dose calculations for radiation therapy. *Acta Radiol Oncol.* 1985;24(6):509-519.

13. Yang M, Zhu XR, Park PC, et al. Comprehensive analysis of proton range uncertainties related to patient stopping-power-ratio estimation using the stoichiometric calibration. *Phys Med Biol.* 2012;57(13):4095-4115.

14. Ainsley CG, Yeager CM. Practical considerations in the calibration of CT scanners for proton therapy. *J Appl Clin Med Phys.* 2014;15(3):4721.

15. Paganetti H. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Phys Med Biol.* 2012;57(11):R99-117.

16. Schaffner B, Pedroni E. The precision of proton range calculations in proton radiotherapy treatment planning: experimental verification of the relation between CT-HU and proton stopping power. *Phys Med Biol.* 1998;43(6):1579-1592.

17. Moyers MF, Sardesai M, Sun S, Miller DW. Ion stopping powers and CT numbers. *Med Dosim.* 2010;35(3):179-194.

18. Han D, Siebers JV, Williamson JF. A linear, separable two-parameter model for dual energy CT imaging of proton stopping power computation. *Med Phys.* 2016;43(1):600.

19. Hunemohr N, Krauss B, Dinkel J, et al. Ion range estimation by using dual energy computed tomography. *Z Med Phys.* 2013;23(4):300-313.

20. Taasti VT, Petersen JB, Muren LP, Thygesen J, Hansen DC. A robust empirical parametrization of proton stopping power using dual energy CT. *Med Phys.* 2016;43(10):5547.

21. Tsukihara M, Noto Y, Hayakawa T, Saito M. Conversion of the energy-subtracted CT number to electron density based on a single linear relationship: an experimental verification using a clinical dual-source CT scanner. *Phys Med Biol.* 2013;58(9):N135-144.

22. Yang M, Virshup G, Clayton J, Zhu XR, Mohan R, Dong L. Theoretical variance analysis of single- and dual-energy computed tomography methods for calculating proton stopping power ratios of biological tissues. *Phys Med Biol.* 2010;55(5):1343-1362.

23. Bar E, Lalonde A, Royle G, Lu HM, Bouchard H. The potential of dual-energy CT to reduce proton beam range uncertainties. *Med Phys.* 2017;44(6):2332-2344.

24. Lee HHC, Li B, Duan X, Zhou L, Jia X, Yang M. Systematic analysis of the impact of imaging noise on dual-energy CT-based proton stopping power ratio estimation. *Med Phys.* 2019;46(5):2251-2263.

25. Li B, Lee HC, Duan X, et al. Comprehensive analysis of proton range uncertainties related to stopping-power-ratio estimation using dual-energy CT imaging. *Phys Med Biol.* 2017;62(17):7056-7074.

26. Mackie TR, Holmes T, Swerdloff S, et al. Tomotherapy: a new concept for the delivery of dynamic conformal radiotherapy. *Medical physics.* 1993;20(6):1709-1719.

27. Langen K, Meeks S, Poole D, et al. The use of megavoltage CT (MVCT) images for dose recomputations. *Physics in Medicine & Biology.* 2005;50(18):4259.

28. Sudhyadhom A. On the molecular relationship between Hounsfield Unit (HU), mass density, and electron density in computed tomography (CT). *PLoS One.* 2020;15(12):e0244861.

29. Meeks SL, Harmon JF, Jr., Langen KM, Willoughby TR, Wagner TH, Kupelian PA. Performance characterization of megavoltage computed tomography imaging on a helical tomotherapy unit. *Med Phys.* 2005;32(8):2673-2681.

30. Langen KM, Meeks SL, Poole DO, et al. The use of megavoltage CT (MVCT) images for dose recomputations. *Phys Med Biol.* 2005;50(18):4259-4276.

31. Ruchala KJ, Olivera GH, Schloesser EA, Hinderer R, Mackie TR. Calibration of a tomotherapeutic MVCT system. *Phys Med Biol.* 2000;45(4):N27-36.

32. Yadav P, Tolakanahalli R, Rong Y, Paliwal BR. The effect and stability of MVCT images on adaptive TomoTherapy. *J Appl Clin Med Phys.* 2010;11(4):3229.

33. Paudel MR, Mackenzie M, Fallone BG, Rathee S. Evaluation of metal artifacts in MVCT systems using a model based correction method. *Med Phys.* 2012;39(10):6297-6308.

34. Yang M, Virshup G, Mohan R, Shaw CC, Zhu XR, Dong L. Improving accuracy of electron density measurement in the presence of metallic implants using orthovoltage computed tomography. *Med Phys.* 2008;35(5):1932-1941.

35. Schulte RW, Penfold SN. Proton CT for Improved Stopping Power Determination in Proton Therapy, invited. *Trans Am Nucl Soc.* 2012;106:55-58.

36. Schneider U, Pedroni E. Proton radiography as a tool for quality control in proton therapy. *Med Phys.* 1995;22(4):353-363.

37. Li T, Liang Z, Singanallur JV, Satogata TJ, Williams DC, Schulte RW. Reconstruction for proton computed tomography by tracing proton trajectories: a Monte Carlo study. *Med Phys.* 2006;33(3):699-706.

38. Hashemi RH, Bradley WG, Lisanti CJ. *MRI: the basics: The Basics.* Lippincott Williams & Wilkins; 2012.

39. Villeirs GM, Van Vaerenbergh K, Vakaet L, et al. Interobserver delineation variation using CT versus combined CT+ MRI in intensity–modulated radiotherapy for prostate cancer. *Strahlentherapie und Onkologie.* 2005;181(7):424-430.

40. Geets X, Daisne J-F, Arcangeli S, et al. Inter-observer variability in the delineation of pharyngo-laryngeal tumor, parotid glands and cervical spinal cord: comparison between CT-scan and MRI. *Radiotherapy and oncology.* 2005;77(1):25-31.

41. Weltens C, Menten J, Feron M, et al. Interobserver variations in gross tumor volume delineation of brain tumors on computed tomography and impact of magnetic resonance imaging. *Radiotherapy and Oncology.* 2001;60(1):49-59.

42. Hricak H, Gatsonis C, Coakley FV, et al. Early invasive cervical cancer: CT and MR imaging in preoperative evaluation—ACRIN/GOG comparative study of diagnostic performance and interobserver variability. *Radiology.* 2007;245(2):491-498.

43. Lagendijk JJ, Raaymakers BW, Raaijmakers AJ, et al. MRI/linac integration. *Radiotherapy and Oncology.* 2008;86(1):25-29.

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

45. Hsu SH, 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):8419-8435.

46. White DR, Booz J, Griffith RV, Spokas JJ, Wilson IJ. Report 44. *Journal of the International Commission on Radiation Units and Measurements.* 1989;os23(1):NP-NP.

47. International Commission on Radiological Protection. Task Group on Reference Man. *Report of the Task Group on Reference Man : a report.* Oxford ; New York: Pergamon Press; 1975.

48. J. H. Hubbell SMS. Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients from 1 keV to 20 MeV for Elements Z = 1 to 92 and 48 *NIST Standard Reference Database 126.* July 2004.

49. ICRU 1989 Tissue Substitutes in Radiation Dosimetry and Measurement. *(Bethesda, MD: International Commission on Radiation Units and Measurements)*.

50. ICRU. *Tissue substitutes in radiation dosimetry and measurement.* Bethesda, Md., U.S.A.: International Commission on Radiation Units and Measurements; 1989.

51. Nyúl LG, Udupa JK, Zhang X. New variants of a method of MRI scale standardization. *IEEE transactions on medical imaging.* 2000;19(2):143-150.

52. Schneider U, Pedroni E, Lomax A. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Phys Med Biol.* 1996;41(1):111-124.

53. Janni JF. Energy Loss, range, path length, time-of-flight, straggling, multiple scatting, and nuclear interaction probability: In Two Parts. Part I. For 63 Compounds Part 2. For Elements 1<Z<92. *Atomic Data and Nuclear Data Tables.* 1982;27(4-5):341-529.

54. Bär E, Andreo P, Lalonde A, Royle G, Bouchard H. Optimized I-values for use with the Bragg additivity rule and their impact on proton stopping power and range uncertainty. *Physics in Medicine & Biology.* 2018;63(16):165007.

55. White DR, Widdowson EM, Woodard HQ, Dickerson JW. The composition of body tissues (II). Fetus to young adult. *Br J Radiol.* 1991;64(758):149-159.

56. Scholey JE, Chandramohan D, Naren T, Liu W, Larson PEZ, Sudhyadhom A. Technical Note: A methodology for improved accuracy in stopping power estimation using MRI and CT. *Med Phys.* 2020.

57. De Marzi L, Lesven C, Ferrand R, Sage J, Boule T, Mazal A. Calibration of CT Hounsfield units for proton therapy treatment planning: use of kilovoltage and megavoltage images and comparison of parameterized methods. *Phys Med Biol.* 2013;58(12):4255-4276.

58. Das IJ, Cheng CW, Cao M, Johnstone PA. Computed tomography imaging parameters for inhomogeneity correction in radiation treatment planning. *J Med Phys.* 2016;41(1):3-11.

59. Sudhyadhom A. Determination of mean ionization potential using magnetic resonance imaging for the reduction of proton beam range uncertainties: theory and application. *Phys Med Biol.* 2017;62(22):8521-8535.

60. ICRU 1993 Stopping Powers and Ranges for Protons and Alpha Particles. *(Bethesda, MD: International Commission on Radiation Units and Measurements)*.

61. Ma J. A single-point Dixon technique for fat-suppressed fast 3D gradient-echo imaging with a flexible echo time. *J Magn Reson Imaging.* 2008;27(4):881-890.

62. Zhou H, Keall PJ, Graves EE. A bone composition model for Monte Carlo x-ray transport simulations. *Med Phys.* 2009;36(3):1008-1018.

63. Wu Y, Ackerman JL, Chesler DA, Graham L, Wang Y, Glimcher MJ. Density of organic matrix of native mineralized bone measured by water- and fat-suppressed proton projection MRI. *Magn Reson Med.* 2003;50(1):59-68.

64. Gruber B, Froeling M, Leiner T, Klomp DWJ. RF coils: A practical guide for nonphysicists. *J Magn Reson Imaging.* 2018.

65. 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.

66. Press RH, Shu HG, Shim H, et al. The Use of Quantitative Imaging in Radiation Oncology: A Quantitative Imaging Network (QIN) Perspective. *Int J Radiat Oncol Biol Phys.* 2018;102(4):1219-1235.

67. Gurney-Champion OJ, Mahmood F, van Schie M, et al. Quantitative imaging for radiotherapy purposes. *Radiother Oncol.* 2020;146:66-75.

68. Jiang H, Seco J, Paganetti H. Effects of Hounsfield number conversion on CT based proton Monte Carlo dose calculations. *Med Phys.* 2007;34(4):1439-1449.

69. Goma C, Almeida IP, Verhaegen F. Revisiting the single-energy CT calibration for proton therapy treatment planning: a critical look at the stoichiometric method. *Phys Med Biol.* 2018;63(23):235011.

70. Yu H, McKenzie CA, Shimakawa A, et al. Multiecho reconstruction for simultaneous water-fat decomposition and T2* estimation. *J Magn Reson Imaging.* 2007;26(4):1153-1161.

71. Yu H, Shimakawa A, McKenzie CA, Brodsky E, Brittain JH, Reeder SB. Multiecho water-fat separation and simultaneous R2* estimation with multifrequency fat spectrum modeling. *Magn Reson Med.* 2008;60(5):1122-1134.

72. Kulikova S, Hertz-Pannier L, Dehaene-Lambertz G, Poupon C, Dubois J. A New Strategy for Fast MRI-Based Quantification of the Myelin Water Fraction: Application to Brain Imaging in Infants. *PLoS One.* 2016;11(10):e0163143.

73. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging.* 2010;29(6):1310-1320.

74. Springer F, Martirosian P, Schwenzer NF, et al. Three-dimensional ultrashort echo time imaging of solid polymers on a 3-Tesla whole-body MRI scanner. *Invest Radiol.* 2008;43(11):802-808.

75. Siu AG, Ramadeen A, Hu X, et al. Characterization of the ultrashort-TE (UTE) MR collagen signal. *NMR Biomed.* 2015;28(10):1236-1244.

76. Frey MA, Michaud M, VanHouten JN, Insogna KL, Madri JA, Barrett SE. Phosphorus-31 MRI of hard and soft solids using quadratic echo line-narrowing. *Proc Natl Acad Sci U S A.* 2012;109(14):5190-5195.

77. Seifert AC, Li C, Rajapakse CS, et al. Bone mineral (31)P and matrix-bound water densities measured by solid-state (31)P and (1)H MRI. *NMR Biomed.* 2014;27(7):739-748.

78. Seifert AC, Wehrli FW. Erratum to: Solid-State Quantitative 1H and 31P MRI of Cortical Bone in Humans. *Curr Osteoporos Rep.* 2016;14(4):159-161.

79. Seifert AC, Wehrli FW. Solid-State Quantitative (1)H and (31)P MRI of Cortical Bone in Humans. *Curr Osteoporos Rep.* 2016;14(3):77-86.

80. Thomas SJ. Relative electron density calibration of CT scanners for radiotherapy treatment planning. *The British journal of radiology.* 1999;72(860):781-786.

81. Mastrogiacomo S, Dou W, Jansen JA, Walboomers XF. Magnetic resonance imaging of hard tissues and hard tissue engineered bio-substitutes. *Molecular imaging and biology.* 2019;21(6):1003-1019.

82. Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Medical physics.* 1998;25(5):656-661.

83. Gupta D, Kim M, Vineberg KA, Balter JM. Generation of synthetic CT images from MRI for treatment planning and patient positioning using a 3-channel U-net trained on sagittal images. *Frontiers in oncology.* 2019;9:964.

84. Wang Y, Liu C, Zhang X, Deng W. Synthetic CT generation based on T2 weighted MRI of nasopharyngeal carcinoma (NPC) using a deep convolutional neural network (DCNN). *Frontiers in oncology.* 2019;9:1333.

85. 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. *International Journal of Radiation Oncology* Biology* Physics.* 2018;100(1):199-217.

86. Liu Y, Yue C, Zhu J, et al. A megavoltage CT image enhancement method for image-guided and adaptive helical TomoTherapy. *Frontiers in oncology.* 2019;9:362.

87. Sheng K, Gou S, Wu J, Qi SX. Denoised and texture enhanced MVCT to improve soft tissue conspicuity. *Medical physics.* 2014;41(10):101916.

88. Vinas L, Scholey J, Descovich M, Kearney V, Sudhyadhom A. Improved contrast and noise of megavoltage computed tomography (MVCT) through cycle-consistent generative machine learning. *Medical Physics.* 2021;48(2):676-690.

89. Ozaki S, Kaji S, Nawa K, et al. Training deep cross-modality conversion models with a small amount of data and its application to MVCT to kVCT conversion. *arXiv preprint arXiv:210705238.* 2021.

90. Langius JA, van Dijk AM, Doornaert P, et al. More than 10% weight loss in head and neck cancer patients during radiotherapy is independently associated with deterioration in quality of life. *Nutrition and cancer.* 2013;65(1):76-83.

91. Bando R, Ikushima H, Kawanaka T, et al. Changes of tumor and normal structures of the neck during radiation therapy for head and neck cancer requires adaptive strategy. *The Journal of Medical Investigation.* 2013;60(1.2):46-51.

92. Scholey JE, Chandramohan D, Naren T, Liu W, Larson PEZ, Sudhyadhom A. Technical Note: A methodology for improved accuracy in stopping power estimation using MRI and CT. *Medical Physics.*n/a(n/a).

93. Yang M, Virshup G, Clayton J, Zhu XR, Mohan R, Dong L. Theoretical variance analysis of single-and dual-energy computed tomography methods for calculating proton stopping power ratios of biological tissues. *Physics in Medicine & Biology.* 2010;55(5):1343.

94. Parker R, Hobday PA, Cassell K. The direct use of CT numbers in radiotherapy dosage calculations for inhomogeneous media. *Physics in Medicine & Biology.* 1979;24(4):802.

95. Yang M, Virshup G, Mohan R, Shaw CC, Zhu XR, Dong L. Improving accuracy of electron density measurement in the presence of metallic implants using orthovoltage computed tomography. *Medical physics.* 2008;35(5):1932-1941.

96. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Paper presented at: International Conference on Medical image computing and computer-assisted intervention2015.

97. Han X. MR-based synthetic CT generation using a deep convolutional neural network method. *Medical physics.* 2017;44(4):1408-1419.

98. Emami H, Dong M, Nejad-Davarani SP, Glide-Hurst CK. Generating synthetic CTs from magnetic resonance images using generative adversarial networks. *Medical physics.* 2018;45(8):3627-3636.

99. Koike Y, Akino Y, Sumida I, et al. Feasibility of synthetic computed tomography generated with an adversarial network for multi-sequence magnetic resonance-based brain radiotherapy. *Journal of radiation research.* 2020;61(1):92-103.

100. Li Y, Li W, Xiong J, Xia J, Xie Y. Comparison of Supervised and Unsupervised Deep Learning Methods for Medical Image Synthesis between Computed Tomography and Magnetic Resonance Images. *BioMed Research International.* 2020;2020.

101. Li W, Li Y, Qin W, et al. Magnetic resonance image (MRI) synthesis from brain computed tomography (CT) images based on deep learning methods for magnetic resonance (MR)-guided radiotherapy. *Quant Imaging Med Surg.* 2020;10(6):1223-1236.

102. Florkow MC, Zijlstra F, Willemsen K, et al. Deep learning-based MR-to-CT synthesis: The influence of varying gradient echo-based MR images as input channels. *Magn Reson Med.* 2020;83(4):1429-1441.

103. Torrado-Carvajal A, Vera-Olmos J, Izquierdo-Garcia D, et al. Dixon-VIBE Deep Learning (DIVIDE) Pseudo-CT Synthesis for Pelvis PET/MR Attenuation Correction. *J Nucl Med.* 2019;60(3):429-435.

104. Barateau A, De Crevoisier R, Largent A, et al. Comparison of CBCT-based dose calculation methods in head and neck cancer radiotherapy: from Hounsfield unit to density calibration curve to deep learning. *Med Phys.* 2020.

105. Chen L, Liang X, Shen C, Jiang S, Wang J. Synthetic CT generation from CBCT images via deep learning. *Med Phys.* 2020;47(3):1115-1125.

106. Lei Y, Tang X, Higgins K, et al. Learning-based CBCT correction using alternating random forest based on auto-context model. *Med Phys.* 2019;46(2):601-618.

107. Zhu JY, Park T, Isola P, Efros AA. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. Paper presented at: IEEE International Conference on Computer Vision (ICCV)2017.

108. Liang X, Chen L, Nguyen D, et al. Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy. *Phys Med Biol.* 2019;64(12):125002.

109. Vinas L, Scholey J, Descovich M, Kearney V, Sudhyadhom A. Improved contrast and noise of megavoltage computed tomography (MVCT) through cycle-consistent generative machine learning. *Med Phys.* 2020.

110. Boskey AL, Coleman R. Aging and bone. *J Dent Res.* 2010;89(12):1333-1348.

111. Woodard HQ, White DR. Bone models for use in radiotherapy dosimetry. *Br J Radiol.* 1982;55(652):277-282.

112. Berthold J, Khamfongkhruea C, Petzoldt J, et al. First-In-Human Validation of CT-Based Proton Range Prediction Using Prompt Gamma Imaging in Prostate Cancer Treatments. *International Journal of Radiation Oncology* Biology* Physics.* 2021.

113. Jäkel O. Ranges of ions in metals for use in particle treatment planning. *Physics in Medicine & Biology.* 2006;51(9):N173.

114. Jäkel O, Reiss P. The influence of metal artefacts on the range of ion beams. *Physics in Medicine & Biology.* 2007;52(3):635.

115. Van Schie MA, Dinh CV, Van Houdt PJ, et al. Contouring of prostate tumors on multiparametric MRI: evaluation of clinical delineations in a multicenter radiotherapy trial. *Radiotherapy and Oncology.* 2018;128(2):321-326.

116. Dinh CV, Steenbergen P, Ghobadi G, et al. Magnetic resonance imaging for prostate cancer radiotherapy. *Physica Medica.* 2016;32(3):446-451.

117. MR-only RT planning for the brain and pelvis with synthetic CT. *Siemens Healthineers MReadings: MR in RT.* 2019.

118. Kündel; DTCWDMUGM, Müller KNDZDWA-C. Synthetic CT Generation for the Pelvic Region Based on Dixon-MR Sequences: Workflow, Dosimetric Quality and Daily Patient Positioning. *Siemens Healthineers MReadings: MR in RT.* 2019;73(5).

119. Wegener OH. *Whole Body Computed Tomography 2nd edn.* Cambridge, MA: Blackwell; 1993.

120. Giantsoudi D, De Man B, Verburg J, et al. Metal artifacts in computed tomography for radiation therapy planning: dosimetric effects and impact of metal artifact reduction. *Physics in Medicine & Biology.* 2017;62(8):R49.

121. Cheung JP, Shugard E, Mistry N, Pouliot J, Chen J. Evaluating the impact of extended field-of-view CT reconstructions on CT values and dosimetric accuracy for radiation therapy. *Medical physics.* 2019;46(2):892-901.

122. Mullins JP, Grams MP, Herman MG, Brinkmann DH, Antolak JA. Treatment planning for metals using an extended CT number scale. *Journal of applied clinical medical physics.* 2016;17(6):179-188.

123. 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. *Medical physics.* 2018;45(11):5218-5233.

124. El Naqa I, Li R, Murphy MJ. *Machine learning in radiation oncology: theory and applications.* Springer; 2015.

125. Mahadevaiah G, Rv P, Bermejo I, Jaffray D, Dekker A, Wee L. Artificial intelligence-based clinical decision support in modern medical physics: selection, acceptance, commissioning, and quality assurance. *Medical physics.* 2020;47(5):e228-e235.

126. Roberts DA, Sandin C, Vesanen PT, et al. Machine QA for the Elekta Unity system: A Report from the Elekta MR-linac consortium. *Medical physics.* 2021.

127. Dhanesar S, Sahoo N, Kerr M, et al. Quality assurance of proton beams using a multilayer ionization chamber system. *Medical physics.* 2013;40(9):092102.

128. Paganetti H. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Physics in Medicine & Biology.* 2012;57(11):R99.

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Jessica Scholey
67DFAA085656407...
Author Signature

12/8/2021
Date