Reducing Uncertainty in Proton Therapy Treatment Planning

Calculation of the Proton Stopping Power using Single and Dual Energy Computed Tomography

Annette Høisæter

Supervisors: Helge E. S. Pettersen, Kristian S. Ytre-Hauge, Kirsten N. Bolstad

Thesis for the degree of Master of Science in Physics

Departement of Physics and Technology
University of Bergen

December 2020

So remember to look up at the stars and not down at you feet. Try to make sense of what you see and hold on to that childlike wonder about what makes the universe exist.

Stephen Hawking

Abstract

Purpose: The purpose of this study was to investigate if, and to what degree, the use of dual energy CT can reduce the uncertainty in range and stopping power calculations, in proton therapy treatment planning, compared to single energy CT.

Methods: A theoretical investigation on existing studies of dual energy CT for proton therapy treatment planning was done. Three methods of predicting the relative stopping power of tissue substitutes were experimentally implemented and compared. These experiments were performed on CT images acquired specifically for this project. The current state-of-the-art method, a stoichiometric calibration on single energy CT (SECT) data was experimentally verified and used for comparing two newer calibration methods. These were a stoichiometric calibration using pseudo monoenergetic CT (MonoCT) data derived from a dual energy CT scan, and a direct RSP calibration, called RhoSigma, using dual energy CT (DECT) data directly. The accuracy of the calibration methods were compared using the mean absolute error (MAE) between the experimental RSP values and the theoretical reference RSP values. The root mean square error (RMSE) of the water equivalent path length (WEPL) differences was also used for comparing the methods. The calibration methods were also compared based on their ease of use.

Results: Across all the calibration methods, the DECT based RhoSigma calibration, provided the most accurate RSP prediction for all the tissue substitutes used. The SECT calibration proved more accurate than the calibration based on MonoCT. The MAE of the DECT, SECT, and MonoCT calibrations were: 1.3%, 2.3%, and 2.8%, respectively. Showing that the DECT calibration is notably more accurate than the two other calibration methods. The RMSE of the WEPL difference maps was calculated to be +3.2 mm (comparing DECT to SECT) and -1.4 mm (comparing MonoCT to SECT).

Conclusions: Based on the results of this study, the RhoSigma calibration method is concluded to improve range calculations in proton therapy treatment planning from 2.3% using SECT to 1.3% using DECT. This conclusion is in agreement with on existing literature. Regarding the MonoCT calibration, conclusions can not be made as the results of this study disagree with existing literature. The work done in this thesis has shown that the use of DECT has a clear potential to improve range uncertainties in proton treatment planning.

Acknowledgements

I would like to express my gratitude to my main supervisor Helge Egil Seime Pettersen, for his guidance and expertise during this past year. Thank you for assisting me with programming and for being available whenever I needed help or encouragement. I could not have done this without you.

Thank you to my co supervisors Kristian Smeland Ytre-Hauge and Kirsten Hansine Helene Nygård Bolstad for supporting me throughout this project. Thank you for being helpful, encouraging and understanding. Your feedback has been invaluable.

I would also like to thank Vicki Trier Taasti for being so welcoming of my questions and taking the time to help me understand how they have implemented MonoCT calibration in Aarhus.

A special thank you goes to my fellow students Ingrid Mossige, Stian Maurseth, Jon Sverre Dyrkolbotn, and Jon Asgeir Torsvik, for being there for me for the past 5 years. You have helped me succeed when the odds were against me. I would not be where I am today without your friendship and support.

Thank you so much to my parents and my family, for supporting all of my decisions and for helping me reach my goals. I especially want to thank my father, Rolf Gerhardt Høisæter, for teaching me to be sceptical, curious, and exited about science. Your many lectures about language, history and how everything works, have played a big role in my accomplishments.

Lastly I would like to thank my fiancee, Jørgen Guttorm Kristiansen. Thank you for always cheering me on, and believing in me, even when I don't believe in myself. I am the luckiest person in the world, and forever grateful, for having you by my side.

Contents

List of Figures

List of Tables

Nomenclature

Abbreviations

ALARA	As Low as Reasonably Achievable
CT	Computed Tomography
DECT	Dual Energy Computed Tomography
HLUT	Hounsfield Look-up Table
HU	Hounsfield Unit
LET	Linear Energy Transfer
MonoCT	Pseudo monoenergetic CT image
OAR	Organ at Risk
PT	Proton Therapy
QA	Quality Assurance
RBE	Relative Biological Effectiveness
RCS	Relative Cross Section image
RED	Relative Electron Density image
RSN	Relative Stopping Number image
RSP	Relative Stopping Power
RT	Radiation Therapy
S	Stopping Power
SECT	Single Energy Computed Tomography
TPS	Treatment Planning System
WEPL	Water Equivalent Path Length

Symbols

β	Velocity in units of speed of light: v/c
$\hat{}$	Relative to water
μ	Linear Attenuation Coefficient
ρ	Mass Density
ρ_e	Electron Density
σ	Cross Section
A	Atomic Mass
L	Stopping Number
n	Number of atoms per unit volume: $\rho\frac{N_A}{A}$
N_A	Avogadro's Number: $6.02214 \cdot 10^{23}\ mol^{-1}$
Z	Atomic Number

1 INTRODUCTION

Cancer is a collective term for malignant diseases where abnormal growth and division of cells are happening at an uncontrolled rate, and with the potential to spread to other parts of the body. Where possible, the goal of cancer treatment involves killing or removing the cancerous cells so as to permanently stop the abnormal cell growth. Despite the apparent simplicity of this treatment, there are numerous obstacles that complicate the process of reaching this end goal. A major one is namely that in the process of killing the cancerous cells in a patient, it is unavoidable also to kill or damage some amount of healthy cells. Thus all methods of cancer treatment are at a constant tug-of-war between these two opposing goals; killing cancer cells while sparing healthy tissue. Even after more than a hundred years of research, with medical and, technological advancements in the field, cancer is still one of the leading causes of death worldwide [12], and the leading cause of death in Norway as of 2018 [13]. This fact should serve as a reminder of the complexity and severity of cancerous diseases and the work still ahead of us.

However, it is important to acknowledge just how far we have actually come. From the early 1980s until today, the 5 year relative survival rates, for all cancer types, for men and women in Norway has increased from 42% and 52% to 74% and 74% respectively [14]. This improvement is owed to the research and development of many different treatment methods in oncology. Methods within radiotherapy (RT) have evolved tremendously over the past few decades making radiotherapy a cornerstone in cancer treatment. Either as a stand alone treatment, or in combination with other treatments such as surgery, chemotherapy, and in later years immunotherapy.

1.1 The History of Radiotherapy

Attempts to treat cancer using radiotherapy began promptly after the discovery of x-rays in 1895 by Wilhelm Conrad Röntgen and radioactivity in 1896 by Marie and Pierre Curie. Although the underlying mechanisms were not understood at the time, experimental medical trial and error revealed that radiation could be used to treat cancer. However, early methods were primitive and with limited applicability, often with severe complications and side effects. One main problem was the inability to produce high energy x-rays with the traditional x-ray tubes, resulting in shallow penetration and high skin dose. Gamma radiation from radium had the ability to treat deeper tumor sites, but radium prices were very high, limiting availability.

Big strides were made with the development of the linear accelerator during the period from 1930 through 1950, in which Norwegian physicist Rolf Widerøe played a big role. His idea of accelerating particles with radiofrequency pulses greatly improved the feasible beam energy and his accelerators became the forerunners of modern particle accelerators and colliders. Many ideas were proposed in this era, e.g. treatment with neutron and proton beams.

In the same period progress was made in radiobiology and dosimetry, with growing understanding of oncology and safety standards being developed. In the 1970s and 1980s computers were introduced to assist in treatment planning and dose monitoring. Several modalities, e.g. Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET), but most importantly Computed Tomograpy (CT) imaging started being incorporated in the treatment course. An important consequence of incorporating CT based treatment plans, was the transition from a standardized, to a patient specific treatment beam setup. Today the state of the art x-ray radiotherapy is highly sophisticated with outcomes and precision incomparable to its humble beginnings. It has become essential in modern oncology.

Trials with proton beams, first only conducted at research facilities, yielded promising results, and in 1989 and 1990 the first hospital based particle centres opened at Clatterbridge Centre for Oncology in the UK and Loma Linda in USA. The appeal of using protons, or heavier ions, is especially that the dose deposition is highest at the end range of the protons. This is the famous Bragg peak, a characteristic that allows for better sparing of healthy tissue, named after William Henry Bragg who discovered this characteristic of charged particles already in 1904 [15]. As of July 2020, there are 104 operating particle therapy facilities worldwide, with 66 more under construction or planning [16]. In 2018 the Norwegian government decided to build two proton therapy centers. This means that Norway has taken an important step forward in the treatment of cancer patients. These facilities are currently under construction.

1.2 Objective of this Thesis

The precise dose deposition of protons is in essence the advantage of proton therapy (PT), however it also makes the method more vulnerable to incorrect range calculations. Due to the finite range and narrow high dose deposition of protons, an error in range calculation leads to a high risk of under-dosage of the tumor or over-dosage of healthy tissue or organs at risk. To account for these uncertainties and ensure full target coverage, appropriate safety margins are added to the target, increasing the target volume. This increases the amount of irradiated healthy tissue, reducing the benefit of proton therapy. These margins are typically in the order of 3.5% + 1 mm of the beam range, which becomes around 8 mm at 200 mm depth [17]. A significant amount of healthy tissue will therefore receive the same dose as prescribed for the tumor. For these reasons, more accurate range prediction is essential to fully exploit the benefits of proton therapy.

The range of the protons in a proton beam is determined by their energy and the relative stopping power (RSP) of the matter they penetrate. In radiotherapy (RT) this matter is the various tissues inside a patient, proximal to the tumor, and the tumor itself. Since the energy of the proton beam is controlled, the only unknown parameter for range calculation is the patient tissue composition. However, acquiring an accurate mapping of the patient RSP is actually the crux that makes up most of the range uncertainty in proton therapy [17].

The current method of obtaining the patient RSP is using single energy CT (SECT) scans. The problem with this approach is that there is no one-to-one correspondence between how photons and protons deposit their energy in matter. Thus the accuracy of the RSP calculation is limited by the physical interaction differences of photons and protons. Two types of tissues that attenuate photons in a similar way may have the same CT number, but they might not stop protons in the same way and thus have different RSP values. Using SECT such differences can not be detected, thus limiting the ability to accurately differentiate tissues.

Several studies point to dual energy CT (DECT) as a tool to reduce the range uncertainties [18, 19, 20, 21]. The additional information obtained from imaging with two different photon spectra, makes it possible to differentiate tissues much more accurately than with SECT. The theory of this thesis focuses on understanding the underlying physics that makes DECT a superior tool for accurate RSP calculation, over SECT.

As treatments at the new proton therapy centers in Bergen and Oslo are planned to begin in 2024/2025, it is important to find and implement the best possible methods of imaging and calibration for proton treatment planning. The main objective in this thesis was to investigate to what degree the use of DECT data can reduce the uncertainty in range and RSP calculations, compared to SECT data.

Calibrations with both modalities was compared, using relevant measurements of accuracy such as water equivalent path length (WEPL) and mean absolute error (MAE).

THEORY 2

In this chapter, relevant background information is explained. Firstly we look at the physical interactions of photons with matter, after which CT imaging and DECT is explained. Thereafter we look at interactions of protons with matter and the range of the proton beam. Lastly how CT imaging provides information that can be used for proton range prediction in patients is investigated.

2.1 Radiobiology

Radiobiology is a branch of science that describes the mechanisms and effects of ionizing radiation on biological tissues and living organisms [1]. What happens during, and immediately after, ionizing radiation interacts with biological structures is well understood. Yet why some patients will develop secondary cancer many years after radiation treatment while others will not, is unknown. A simplified way of looking at it is that the cells that are exposed to radiation, but survive, are altered in some way that leads to malfunction after some time. No threshold amount of dose has been proven to exist where this phenomenon comes into effect, meaning that any small amount of dose might increase the risk of secondary cancer. This is the basis for the ALARA (As Low As Reasonably Achievable) principle. A radiation safety principle of avoiding any amount of unnecessary radiation exposure with all reasonable means [22]. Any development in RT that leads to a reduction in dose to healthy tissue while still killing all tumor tissue, might reduce the number of patients that develop secondary cancer. Research furthering this goal is therefore especially impactful for patients with a long life expectancy after treatment, i.e. children and young adults [23].

2.2 Ionizing Radiation

Radiation deposits energy in small volumes of tissue corresponding to single cells or parts of cells depending on the type of radiation. In order to kill a cell, sufficient damage needs to be dealt to the cell's DNA molecule where genetic information needed for mitosis (cell division) is stored. This is mostly achieved when both strands of the DNA molecule is broken [24]. For this to be possible, the radiation needs to carry enough energy to remove electrons in the target material from their bound state in atoms and molecules, creating ions and free radicals. This type of radiation is therefore called ionizing radiation. There are two main types of ionizing radiation; directly, and indirectly ionizing.

Indirectly ionizing radiation consists of electrically neutral, high energy particles photons (x-rays, and gamma-rays) and neutrons. Despite these particles' ability to ionize, most of the ionizations in the absorbing material is done by secondary directly ionizing particles liberated by the primary radiation. This is why in a depth-dose curve for a photon beam, the peak energy deposition will be, not at the surface, but after an initial buildup of free energetic electrons a few cm into the absorbing material (see figure 2.11). Although neutrons are also indirectly ionizing, they are more harmful, due mainly to their reacting with the nuclei of the absorbing material. The most common interactions are inelastic collisions, neutron capture, and fission, resulting in a combination of emitted gamma-rays, x-rays, beta particles, and radioactive fragments. The secondary charged particles released by the neutral indirectly ionizing particles then go on to cause DNA damage. [1]

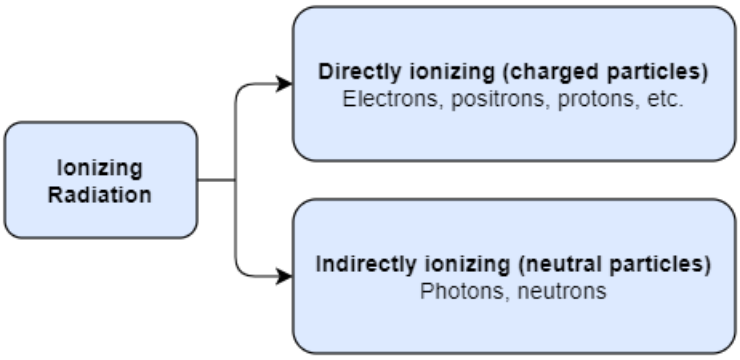

Figure 2.1: Types of ionizing radiation. (Adapted from [1] fig. 1.1)

Directly ionizing radiation consists of charged particles like electrons, protons, and other heavier ions. These particles interact with matter primarily through Coulomb forces and can directly damage the DNA of several cells while traveling through matter until all their kinetic energy is lost.

2.3 Photon-Matter Interactions

Photons in the energy range relevant for medical imaging (50keV to 150keV) interact with matter through three main mechanisms: the photoelectric effect, Compton scattering, and Rayleigh scattering, see figure 2.2.

Photoelectric Effect
An incident photon, with energy close to the binding energy of the electron, is absorbed by an atom which releases an electron with energy equal to the photon, see figure 2.2a. This effect takes place mainly at relatively low photon energies, when the photon energy is close to that of the binding energy of the electron. The likelihood of this interaction is proportional with Z^3/E^3 [25]. Thus it increases with increasing atomic number and with decreasing energy. This characteristic is highly useful in dual energy CT (DECT) imaging, as it can be utilized for separating the effects of the energy dependence via the

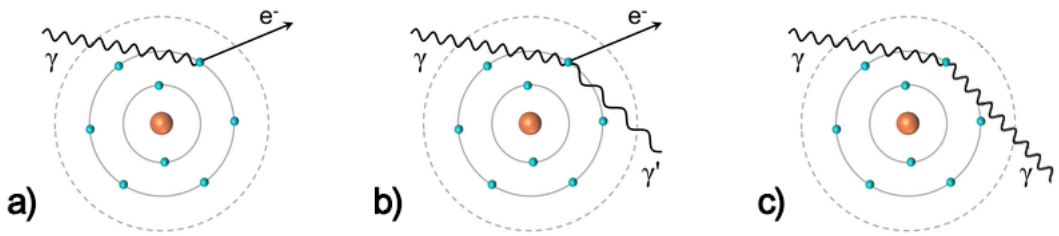

Figure 2.2: Photon interactions with matter. a) Photoelectric effect. b) Compton Scattering. c) Rayleigh scattering. (Adapted from [2] fig. 3)

two energy spectra (more on this in section 2.4).

Compton Scattering
Compton scattering happens when an incident photon ejects an atomic electron, and is itself deflected at an angle, see figure 2.2b. The photon loses some of its energy in the interaction. The amount of energy lost, depends on the angle of the scattering, with the highest energyloss at 180° deflection. This effect is more or less constant at diagnostic energies.

Rayleigh Scattering
The photon is scattered at a small angle of an atomic electron without transferring energy to the electron, see figure 2.2c.

The scattering interactions contribute to image noise if the scattered photons reach the detector, and to dose to persons standing close to the patient if not.

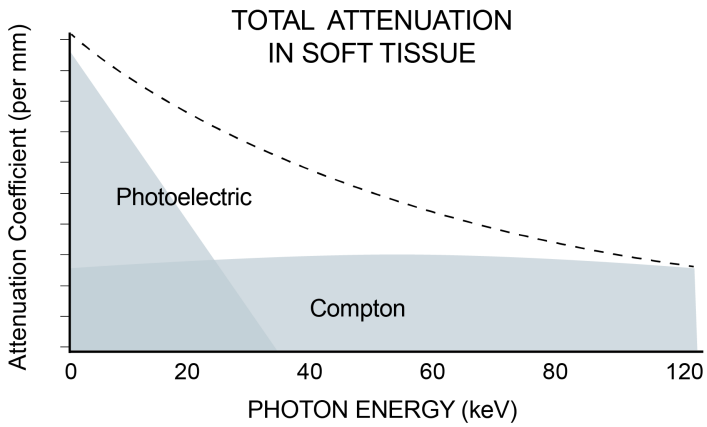

Figure 2.3: Attenuation dependency on photon energy. The photoelectric effect varies strongly within imaging energy range while Compton scattering is more or less constant. (Adapted from [3])

2.3.1 Photon Attenuation

The above mentioned photon-matter interactions remove photons from the beam as they travel through a material. The intensity of the beam, or the number of photons that make it all the way through, decreases exponentially with depth, as shown in fig. 2.4.

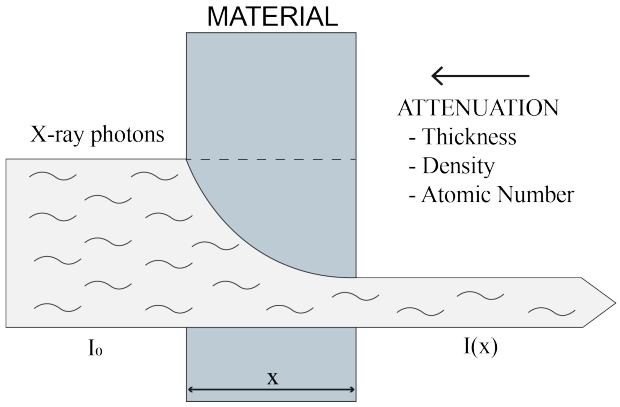

Figure 2.4: Factors affecting x-ray attenuation. I_0 is the incident intensity, x is the thickness of the attenuating material. (Adapted from [3] fig. 11-1)

The attenuation of the beam intensity is described by

$$I(x) = I_0 \, e^{-\mu x} \tag{2.1}$$

where I_0 is the initial photon intensity, x is the thickness of the material traversed and μ is the linear attenuation coefficient. The linear attenuation coefficient can be described as the number of atoms per unit volume n [26], or as the electron density ρ_e [27], times the photon absorption cross section σ.

$$\mu = n \, \sigma \tag{2.2a}$$

$$\mu = \rho_e \sigma \tag{2.2b}$$

where σ is the cross-section for photon-matter interaction. σ is a function of the photon energy, making μ a function of energy as well: $\mu(E) = n\sigma(E)$. The total cross section can be described as the sum of the cross sections for the different photon-matter interactions that contribute [28]. Compton and Rayleigh scattering, and the photoelectric effect are denoted subscripts C, R, and p:

$$\sigma(E) = \sigma_C(E) \, + \, \sigma_R(E) \, + \, \sigma_p(E). \tag{2.3}$$

This is relevant for converting CT-numbers to proton stopping powers (section 2.7). Pair production is not included, because the effect only occurs at energies higher than what is relevant for medical imaging.

2.3.2 The x-ray spectrum

X-rays are produced by accelerating electrons towards a high Z material. The accelerated electrons collide with the target, where they are decelerated and release their energy in the form of electromagnetic radiation, called x-rays or photons.

The energy of the x-rays is limited by the energy of the accelerated electrons. The electron energy is equal to the voltage, V, on the tube times the electron charge, e, e.g. 120 keV if the tube Voltage is 120 kV The resulting x-rays however are not mono energetic, but are distributed over a spectrum of energies up to 120 keV (see Figure 2.5).

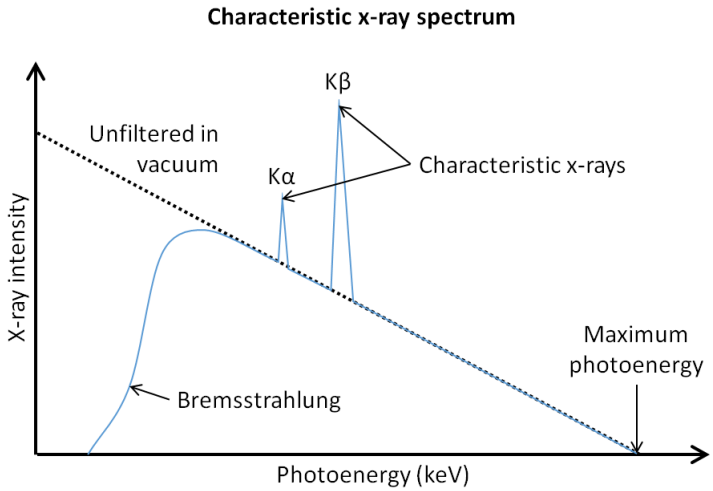

Figure 2.5: A typical x-ray spectrum. The continuous curve is from bremsstrahlung, while the spikes are from characteristic x-rays. The lower energy end of the spectrum is typically filtered out. Figure from [4]

The distinct shape of the x-ray spectrum is indicative of the two main mechanisms in which the electron energy is transformed into x-rays. Bremsstrahlung (or braking radiation) is responsible for the continuous curve while characteristic radiation produce peaks specific for the target atom. When referring to the energy of an x-ray source therefore, the ending kV_p, where p stands for peak, is used, indicating the highest possible energy of photons in the spectrum [25]. As seen in Figure 2.5 however, the effective energy of the photons in such a spectrum is notably lower, usually around 30% or 40% of the peak energy, depending on the shape of the spectrum. The term effective energy represents the energy of a supposed mono-energetic beam with the same penetrating ability as the original spectrum [3].

Note also that the lower end of the spectrum dips towards zero. In medical use, e.g. CT imaging, it is preferable to remove the lower energy photons as they would otherwise be absorbed in the patient while not contributing to the CT image. This is done by filtering the beam through a foil where some of the low energy photons are absorbed, also known as beam hardening. Additional filtering is commonly used in dual energy CT (DECT), to better separate the two energy spectra.

2.4 Computed Tomography

Unlike traditional x-ray imaging where the x-ray source is in a fixed position, CT imaging uses a moving x-ray source that quickly circles around the patient in a toroidal gantry. Detectors are placed directly opposite the x-ray source in the gantry. When imaging, the patient is placed on a table that moves through the gantry, as the x-ray source revolves around the patient in a helix, sending a wide x-ray beam through the patient, see 2.6. Energy spectra for CT imaging usually ranges from 80 to 140 kV_p. A typical imaging energy used in single energy CT (SECT) for adults is 120 kV_p.

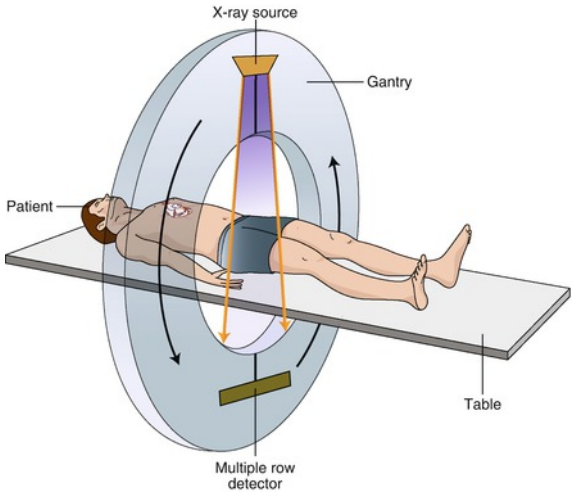

Figure 2.6: Basic principles of CT scanner design (Figure from [5])

The x-rays that pass through the patient are measured by the detector as transmission readings, or sums of the x-ray attenuation coefficients, μ, that pass through the patient. One full rotation around the patient is reconstructed into one axial slice image, using mathematical algorithms to solve for μ at all points inside the patient. This allows for seeing organs and anatomy inside the patient without invasive surgery (see figure 2.7). The scan also provides energy averaged information on the density and composition of tissues inside the patient, according to equations 2.2 and 2.3. This information is used when estimating proton stopping powers, see section 2.7. The pixel values are represented in scaled μ-values, or CT numbers in Hounsfield Units (HU), defined as [26]:

$$\mathscr{H} = (\hat{\mu} - 1) \cdot 1000HU \ , \tag{2.4}$$

where $\mathscr{H}$ is the CT number, $\hat{\mu}$ is the attenuation coefficient relative to the one of water[1]. With this definition the CT number of water is always zero, independent of the x-ray spectrum.

[1]A hat on a variable is used throughout this work to refer to that unit relative to the same unit for water

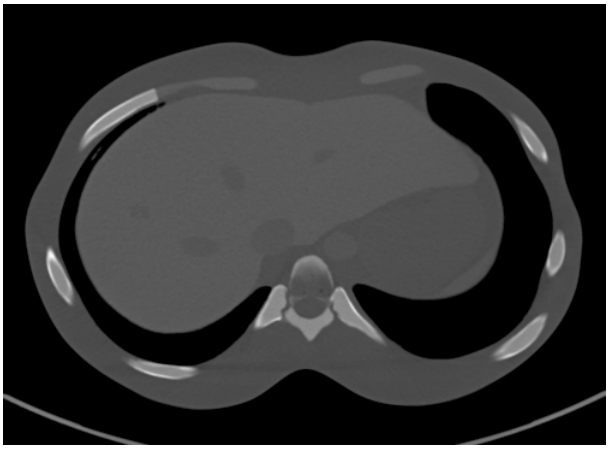

Figure 2.7: Axial slice of an anthropomorphic abdomen phantom. Darker/lighter grays indicate lower/higher densities.

The x-ray beam is hardened before entering the patient (see section 2.3.2), but further beam hardening also occurs naturally as more of the low energy photons are absorbed while traveling through the patient. This means that tissues closer to the center of the patient will receive a higher mean energy spectrum than tissues closer to the surface as more of the lower energy photons are absorbed while traversing the patient volume. The measured attenuation and CT numbers of identical tissues inside the patient, will therefore vary depending on their position inside the patient and on the patient size.[29]

2.4.1 Dual Energy CT

Dual energy CT, or DECT, is a modality which allows for CT imaging with two different x-ray spectra. This produces two sets of images, one for each energy. Ideally the two image sets should overlap perfectly to avoid blurring and to ensure that each corresponding pixel in the two images represents the same tissue. This is only possible if the image sets are acquired simultaneously and at the same angle. In practice this is only possible to achieve with a dual layer DECT scanner (see figure 2.8c) where only one x-ray source is used. The two layers of detectors detect different parts of the energy spectrum, creating two acquisitions with different attenuation information. DECT images can also be acquired in different ways, some of which can be seen illustrated in figure 2.8.

In rapid kV-switching, the system uses one x-ray source that rapidly ($\sim$ ms) switches between the high and low energies. Dual source DECT uses two sources fixed orthogonally in respect to each other in the gantry. In both of these techniques the discrimination of the energy spectra happens in the x-ray source. Energies can also be separated in the detector, by using a layered detector where each layer is sensitive to different photon energies. Other approaches also exist [6], with photon counting CT being one promising relevant modality [30].

The main goal is the same: the two images should be geometrically as identical as possible, and spectrally as different as possible. The different energy spectra provide

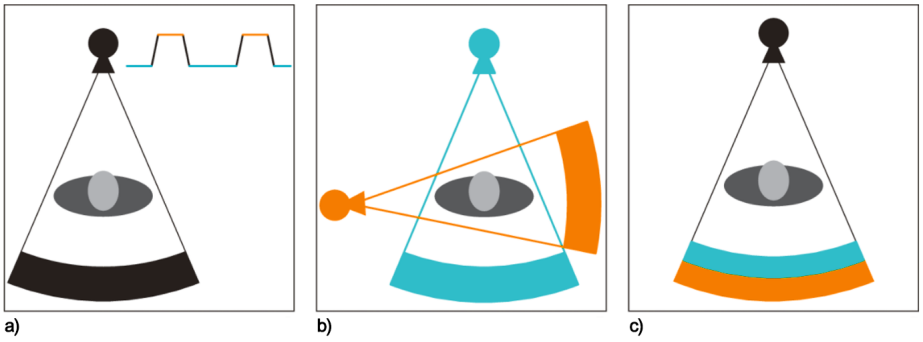

Figure 2.8: Technical approaches to DECT. a) Rapid kV-switching, using one x-ray source. b) Dual source, using two source-detector pairs mounted orthogonaly in the gantry. c) Dual layer, using a single x-ray source and separating the energies in the layered detector. Figure from [6]

different information on the tissues inside the patient, therefore, the more they overlap the less is to gain from using two separate spectra. The lowest reasonable and highest possible energies are thus used in DECT. A common energy pair is 80 kV_p and 140 kV_p with additional tin filtration is used for the highest spectrum to further increase the spectrum separation (see figure 2.9). This additional filtration of one of the x-ray spectra is only possible with dual source DECT, as the filters can not be added/removed at the same pace as the voltage is switched.

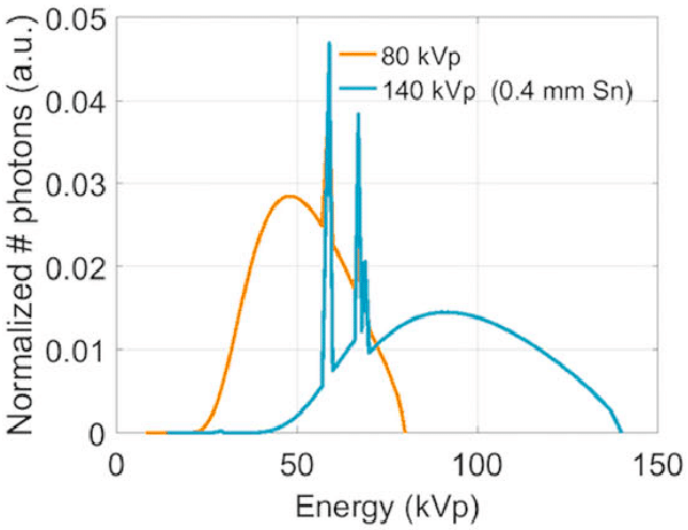

Figure 2.9: The energy spectra of 80 kV_p and 140 kV_p (plus 0.4 mm additional tin (Sn) filtration) x-rays. (Figure from [7])

The CT number gives the average μ over the x-ray spectrum. For this reason two tissues with slightly different μ, can have the same CT number [31] in SECT. These same tissues will however rarely have the same average μ for two different x-rays pectra, making them differentiable in DECT. This improved ability to differentiate tissues with DECT has many advantages in medical applications. One of them is the possibility to

create pseudo monoenergetic CT images, referred to as MonoCT in this thesis. MonoCT images can be made from a superposition of the two CT numbers derived from DECT; $\mathscr{H}_L$ and $\mathscr{H}_H$ from the low and high energy spectra respectively [32]:

$$\mathscr{H}_{Mono}(\alpha) = \alpha\,\mathscr{H}_L + (1-\alpha)\mathscr{H}_H \tag{2.5}$$

where α is the blending factor. By varying the value of the α, the contribution from the two spectra can be changed, e.g. Compton scattering is more dominant in the high energy spectrum and the photoelectric effect is more dominant in the low energy spectrum (see figure 2.3). Thus from one DECT scan one can produce different MonoCTs for different purposes. To name some, an $\alpha > 0$ yields a high contrast image, useful for contouring an delineation, while an $\alpha < 0$ results in an image with reduced metal artifacts [18]. These calculated CT numbers are similar to the CT numbers that would result if a monoenergetic x-ray beam had been used. The MonoCT images should therefore in theory not be as affected by beam hardening as SECT images and has been shown to have better stability in CT numbers across varying patient sizes [33]. This is especially the case close to metals, and thus MonoCT is often considered a metal artefact reduction algorithm.

A downside of DECT is that the acquired image set should not use a higher radiation dose than a regular SECT scan, according to the ALARA principle. Because of this DECT scans, and resulting MonoCTs or other reconstructions, also have this issue. This could impact the accuracy of RSP prediction based on DECT images.

2.5 Proton Radiotherapy

As previously mentioned, the appeal of proton therapy for cancer treatment is mainly the dose deposition characteristics of protons. In this section we take a closer look at characteristics of the proton beam.

2.5.1 Proton-matter interactions

In the energy ranges relevant for proton therapy, the main types interactions that take place between the protons and matter are; inelastic and elastic Coulomb scattering, as well as non-elastic nuclear reactions, see figure 2.10.

Inelastic Coulomb Scattering
The proton kicks out an atomic electron, but is not deflected from its trajectory due to its high mass relative to the electron. The proton loses a small fraction of its energy in the interaction. Through enough collisions like this the proton will finally lose all its energy and come to a stop. This electronic stopping of the proton beam when traveling through a material, is referred to as the materials stopping power, S, described by the Bethe equation [34][35].

$$S = \rho_e \frac{4\pi e^4}{m_e c^2 \beta^2} L(I, \beta) \tag{2.6}$$

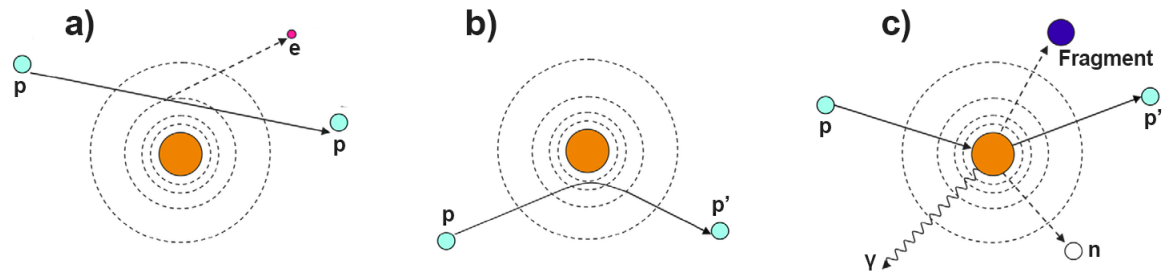

Figure 2.10: Proton interactions with matter. a) inelastic Coulomb scattering - energy loss b) elastic Coulomb scattering - deflection c) non-elastic nuclear reactions - removal of primary proton and creation of secondary particles. p: proton, e: electron, n: neutron, γ: gamma ray. (Adapted from [8] fig. 1)

where ρ_e is the electron density of the material, e is the electron charge, $m_e c^2$ is the electron rest energy, β is the velocity of the proton in units of speed of light, and $L(I, \beta)$ is the stopping number. The stopping number, or Fano's term, contains several corrections to the Bethe formulation. When shell and density corrections are neglected $L(I, \beta)$ can be written:

$$L = \ln \frac{2m_e c^2 \beta^2}{1 - \beta^2} - \beta^2 - \ln I \ , \tag{2.7}$$

where I is the mean ionization energy of the material. The protons energy loss is at its highest right before it comes to a complete stop, which gives rise to the characteristic Bragg peak, see figure 2.11.

Elastic Coulomb Scattering

If passing close enough to nuclei in the target material to interact with their positive charge, protons in the beam can also be elastically scattered [8]. The proton loses a small amount of energy in this interaction. A single account of these events may not deflect the proton much, but through multiple coulomb scatterings the effect is a significant lateral broadening of the beam. This is one of the reasons why heavier ion therapy is considered, where the increased mass leads to a reduction in lateral broadening [36].

Non-elastic Nuclear Reactions

If the incoming protons have enough energy to overcome the Coulomb barrier, they can interact with nuclei in the target material [8]. This interaction removes primary protons from the beam, reducing the proton flux, and creating secondary particles like protons, electron, neutrons, gamma-photons and radioactive fragments. These secondaries also contribute to the treatment dose and might lead to increased risk of late effects [37].

2.5.2 The Proton Beam and Dose Deposition

As the protons interact with the material they are travelling through, they start slowing down as they transfer their energy to the material. As the velocity of the protons decrease their energy loss increase, which can be seen from the inverse dependency on the particle velocity (equation 2.6). This is explained by the proton having more time to interact with the material when it is traveling at a lower velocity. As a consequence, the protons transfer most of their energy to the material right before they come to a stop. This peak of energy deposition is usually referred to as the Bragg Peak (see figure 2.11).

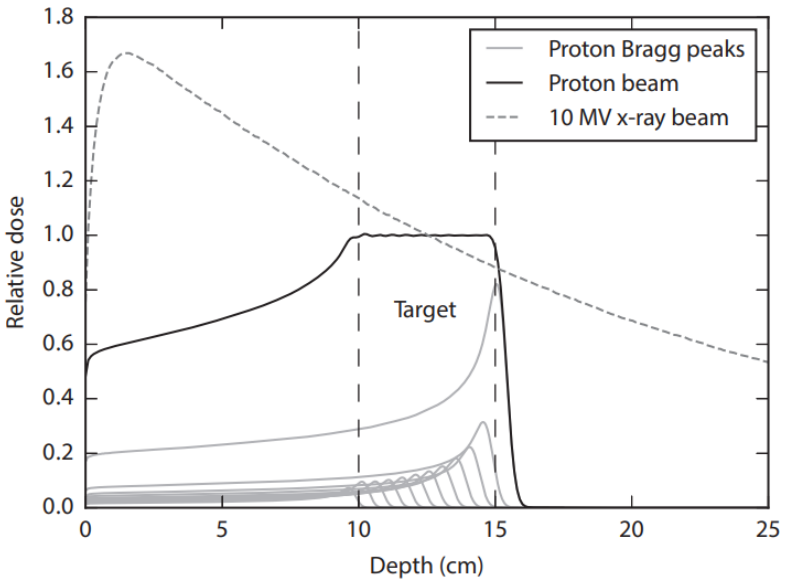

Figure 2.11: Comparison of depth-dose profiles for an x-ray beam and a proton beam in water. The target for dose delivery is located at 10 cm to 15 cm depth. The resulting broadened flat part of the proton beam peak is usually referred to as the spread out Bragg peak (SOBP).

To cover the entire tumor area, the proton beam energy needs to be modified in steps, as illustrated in figure 2.11. The spread out Bragg peak (SOBP) is the sum of the dose profiles of several proton beams at different energies. Figure 2.11 also illustrates the clear advantage of protons in comparison to photons in RT: significantly reduced dose to healthy tissue in front of the tumor, and negligible dose behind the tumor, see figure 2.12.

The characteristic depth-dose profile of the proton beam therefore carries great potential of reducing dose to healthy tissue, if the range of the beam is predicted accurately.

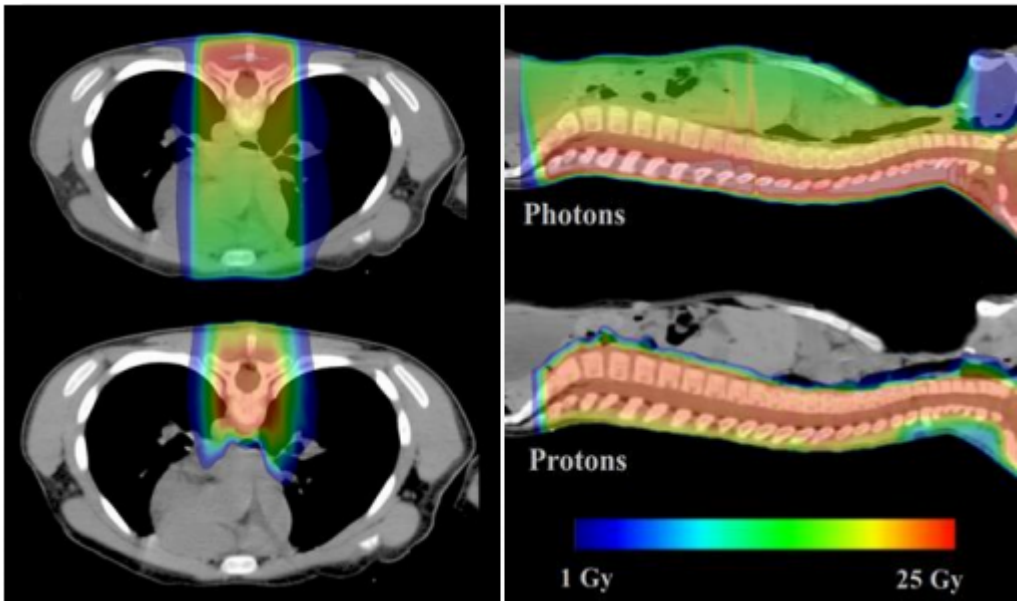

Figure 2.12: Healthy tissue sparing. Reduced dose to heart and lungs using proton RT for craniospinal irradiation, most usually prescribed for medulloblastoma where there is a substantial risk of cancer spread in the central nervous system. Top row: photon RT plan. Bottom row: proton RT plan. Used with permission [9]

2.5.3 The Range of the Proton Beam

One of the most important challenges in proton therapy is the uncertainty in the range of the proton beam. The high number of small individual energy losses with atomic electrons in the target material can be approximated to be a continuous slowing down process. According to the continuous slowing down approximation (CSDA), the range of a proton can be approximated by integrating the inverse stopping power[8]:

$$R = \int_0^{E_0} \frac{1}{S(E)}\, dE \tag{2.8}$$

Because of statistical variations in interactions, not all protons in the beam with the same initial energy will have the same exact range. The range in equation 2.8 is therefore considered the mean range, or the depth at which 50% of the protons have stopped. Clinically another definition of the beam range is usually used, defining the range as the depth at which the dose has decreased to 80% of maximum distally to the Bragg peak. These two definitions both result in approximately the same range [8, 31]. The variation in range between infividual protons broadens the Bragg peak and is called range straggling. This effect is even more prominent when the protons travel through heterogeneous material, e.g. inside a patient.

To calculate the range of the proton beam inside a patient, the tissue's relative stopping

powers (RSP) to the one of water is used:

$$RSP \;=\; \hat{\rho}_e \, \frac{\ln\left(\frac{2mc^2\beta^2}{1-\beta^2}\right) - \beta^2 - \ln(I)}{\ln\left(\frac{2mc^2\beta^2}{1-\beta^2}\right) - \beta^2 - \ln(I_w)}, \tag{2.9}$$

where I and I_w is the ionization energies, or I-values, of the tissue and of water respectively[2]. The RSP can therefore also be expressed simply as $\hat{\rho}_e$ times the relative stopping number $\hat{L} = L/L_w$ (see eq. 2.7):

$$RSP \;=\; \hat{\rho}_e \hat{L}\,, \tag{2.10}$$

meaning that the RSP is directly proportional to the relative electron density $\hat{\rho}_e$. This shows that to calculate a tissue's RSP we first need to determine its electron density and mean ionization energy. The main contributor to the RSP of these two parameters is the relative electron density, $\hat{\rho}_e$, however the mean ionization energy also impacts the RSP calculation.

The accuracy of determining I-values, even of materials of known elemental composition such as water, is a disputed topic. For compound materials the I-value is usually calculated using tabulated ionisation values for the elements in a solid state, estimated by Berger and Zeltzer in 1982 [38] (see table A.2), and the Bragg additivity rule:

$$\ln(I) = \frac{\sum_i w_i \frac{Z_i}{A_i} \ln(I_i)}{\sum_i w_i \frac{Z_i}{A_i}}, \tag{2.11}$$

where Z_i, A_i, and w_i is the atomic number, atomic weight, and weight fraction in the material composition. The problem with this is that the I-value of an atom or molecule depends on what it is bound to and the types of bonds [39]. Thus the I-values of the elements in real biological tissues will not be the same as the I-values of the elements in a solid state. The resulting uncertainty contribution from the I-value was estimated by Paganetti [17] to be approximately 1.5% of the total 3.5% uncertainty in the RSP calculation. For materials of unknown elemental composition, e.g. tissues inside a living patient, the problem is even more complex. The problem with the I-value determination falls beyond the scope of this thesis and the current clinical derivation will therefore be used.

Beam range margin

As previously stated margins are added to the calculated beam range. This margin is typically in the order of 3.5% + 1 mm. This number does not only account for theoretical range calculations uncertainties, but also includes other clinical uncertainties, such as patient positioning, organ motion, etc. The contribution from the CT based range prediction alone is therefore expected to be lower.

To obtain the needed tissue-information, medical imaging is used, which brings us back to CT.

[2]The subscript w will generally refer to water in this thesis

2.6 Tissue Characterization using CT imaging

The most effective way of obtaining accurate information on the proton stopping powers of tissues inside patients, would be to use proton CT (pCT) for the planning CT images. By increasing the energy of the proton beam so that the Bragg peak is located in a detector outside the patient, it is possible to reconstruct a SP-map of the patient. The accuracy of this RSP calculation drastically exceeds the calculations based on x-ray CT [40]. Unfortunately this modality is not yet clinically available, so a calibration from CT numbers to RSP is used instead. There exists a somewhat standardised method of obtaining RSP from CT images, although, as pointed out by Wohlfahrt and Richter in their recent review article on the topic [18]:

> *"The definition of CT scan protocols and image reconstruction as well as in particular the conversion from CT number to SPR are challenging and error-prone processes demanding special attention and quality assurance. This is especially true for PT centers in the preclinical preparation phase, when it is often difficult to allocate resources for such, at first glance, sideline tasks."*

And indeed, although the general physical principles are agreed upon, there exist no guidelines on how to apply the calibration in practise. The current clinically applied method of estimating RSP, is using patient data obtained from an x-ray CT scan. The energy dependent x-ray attenuation coefficient, obtained from CT scans, of a compound material is given by [26, 28, 41]:

$$\mu(E) = \rho N_A \sum_i \frac{w_i}{A_i} \left(k_C(E) Z_i \, + \, k_R(E) Z_i^{2.86} \, + \, k_p(E) Z_i^{4.62} \right), \tag{2.12}$$

where E is the energy of the x-rays, ρ is the mass density, N_A is Avogadro's number and the three terms in the sum are the energy and atomic number parametrisation of the cross sections for Compton scattering, Rayleigh scattering, and the photoelectric effect, respectively. Since the CT number, $\mathscr{H}$, is directly related to $\mu(E)$ (see equation 2.4), density and compositional information can be acquired from CT scans. However due to the polychromatic photon energies used in CT, the measured μ represents the average attenuation coefficient over the x-ray spectrum, μ_s, calculated as:

$$\mu_s = \int_E \mu(E)\, s(E)\, d(E)\, dE \, , \tag{2.13}$$

where $s(E)$ is the x-ray spectrum as produced by the CT scanner and $d(E)$ describes the number of photons registered by the detector. These two factors will vary from scanner to scanner, even between scanners of the same type/model. This means that any calibration from CT-numbers to RSP should be specific for every scanner used for PT treatment planning. Because the measured μ_s, and therefore the CT number, is an average value, two similar tissues might be assigned the same CT number, and will therefore not be differentiable, if only one x-ray spectrum is used. Using two x-ray spectra, as in DECT, this will rarely be the case for both spectra and thus the tissues can be differentiated.

2.7 CT-based RSP calculation

In order to calculate the RSP of a material with the Bethe-equation (eq. 2.9) one needs to know the materials' relative electron density, $\hat{\rho}_e$, and the mean ionisation energy, the I-value. The relative electron density of a compound material can be calculated similarly to the Bragg additivity rule with:

$$\hat{\rho}_e = \frac{\rho \sum_i w_i \frac{Z_i}{A_i}}{\rho_w (w_H \frac{Z_H}{A_H} + w_O \frac{Z_O}{A_O})}, \tag{2.14}$$

where the subscripts H and O represent the elemental constituents of water; Hydrogen and Oxygen. From this and equations 2.12 and 2.4, one can see that the CT number depends linearly on $\hat{\rho}_e$, meaning that this information can be obtained from CT imaging. The I-value on the other hand, can only be obtained if imaging with directly ionizing radiation, which x-rays are not. The I-value thus can not be determined from CT imaging and an exact relation between CT numbers and RSP is not possible to obtain using this modality. Recall also that the RSP also depends linearly on $\hat{\rho}_e$ (see equation 2.10). The CT number and the RSP both depend linearly on the electron density, but not with the same proportionality constant, see figure 2.13 [31] This proportionality constant is σ for the CT number and L (containing the I-value) for the RSP. Because of this, and because the electron density is the main contributor to the RSP (compared to the I-value) it is assumed that the relation between CT number and RSP can be approximated with a piecewise linear function. This idea is the basis for today's state-of-the-art RSP calculations.

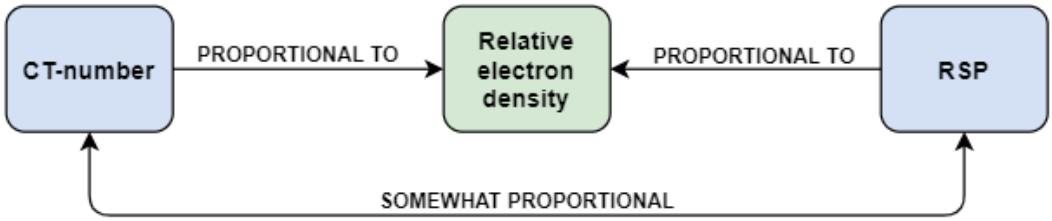

Figure 2.13: The CT number and the RSP are both proportional to $\hat{\rho}_e$, but not with the same proportionality constant. Clinically, the assumption is that the CT number and RSP are somewhat proportional to each other.

2.7.1 Tissue Substitution Calibration

The easiest way of establishing a relation between HU and RSP is to use a set of tissue like materials where the elemental compositions are known, usually referred to as tissue substitutes (the workflow is illustrated in figure 2.14). The RSP values of each tissue substitute can then be calculated theoretically, and plotted against their HU obtained from a SECT scan. A piecewice linear curve is then fitted to the plot, so that any given HU corresponds to a RSP. This semi-linear relation is called a Hounsfield Lookup Table (HLUT).

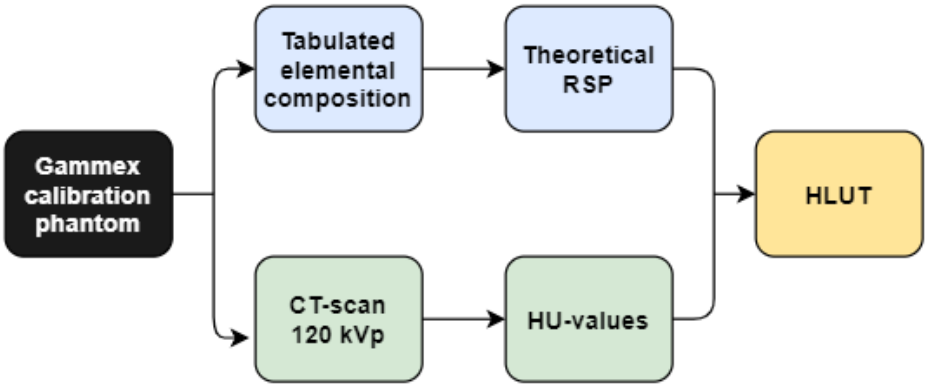

Figure 2.14: Tissue substitution calibration workflow. Tissue substitutes of known elemental composition are scanned. The RSP is calculated theoretically and plotted against the measured HU values from the CT scan. A piecewise linear curve is fitted to the plot to make the HLUT.

2.7.2 Stoichiometric Calibration

One problem with the tissue substitute calibration is that there are inherent differences between tissue substitutes and real human tissues. Ideally the calibration should be done with real tissues, however this is not possible for practical reasons. In 1996 Schneider *et al.* proposed a method for including real tissue data in the HLUT calibration [42] (the workflow is illustrated in figure 2.15). In this method two CT scanner specific parameters k_1 and k_1 are determined (see section 3.3.3) using tissue substitutes. These k-values characterize the scanners' x-ray spectrum and detector response. Once the scanner is characterized it is possible to calculate synthetic HU values for a number of real tissues as if they were scanned at the characterized scanner. This human tissue data is usually taken from the ICRU recommended human tissue database [43], originally made by White *et al.* in the 1980s. Theoretical RSP values are also calculated for the real tissues. This set of synthetic HU values and theoretical RSP values for the real human tissues are then used to make the HLUT. This stoichiometric calibration is considered the gold standard today, although one third of the European PT centers included in the 2019 review from Wohlfahrt and Richter, still use the tissue substitution method [18].

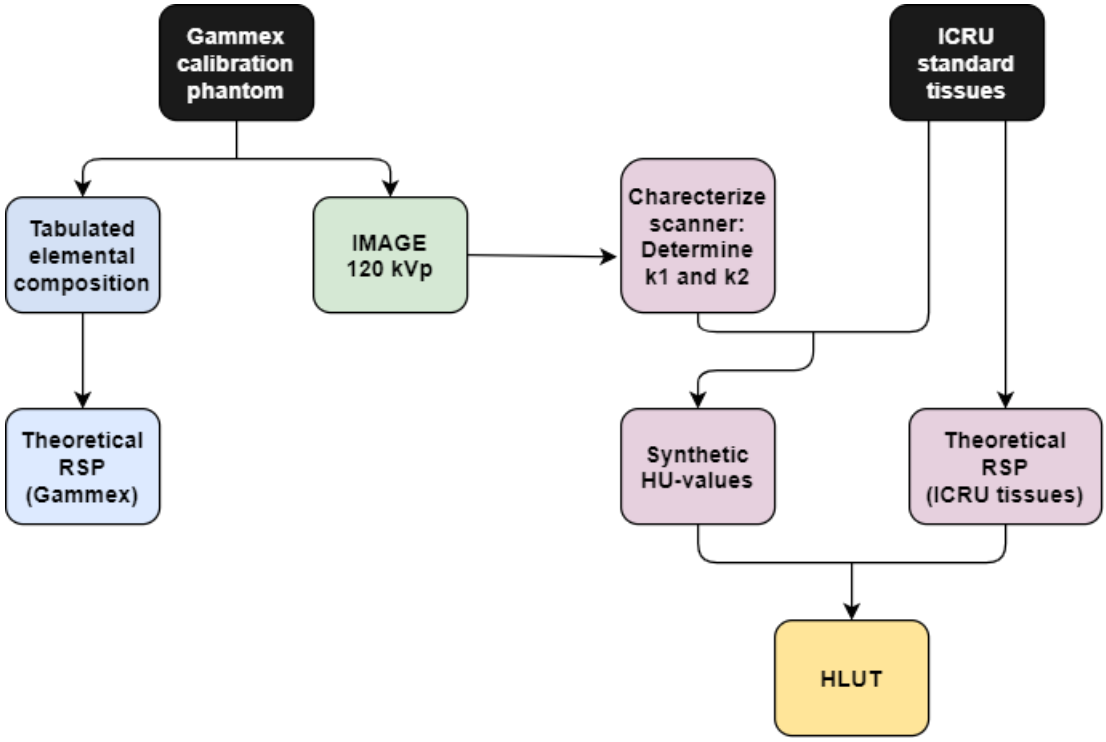

Figure 2.15: Stoichiometric calibration workflow. A tissue substitute calibration phantom (here Gammex 467) is used to characterize the scanner via two parameters k_1 and k_2. Synthetic HU values can then be calculated for tabulated real human tissues. The synthetic HU values and theoretical RSP values of the human tissues are then used to make the HLUT.

2.7.3 Problems with the HLUT approach

Even though the stoichiometric calibration is considered the gold standard, there exist no general guidelines on how the HLUT should be made, other than that it should be a piecewise fit. The most common approach seems to be using three line segments for lung, soft, and bone tissues. How to choose the HU intervals for the segments, how many segments should be used, and how to connect them is not standardised. Neither is there any recommendations on using different HLUTs for different patient groups or tumor sites. As explained in section 2.3.1, the patient geometry will affect the measured HU due to beam hardening. Further, the HLUT approach does not account for the relative abundance of the various tissues in the treatment area. The HLUT could be weighted for the ratio of lung/soft/bone tissue that is present in different body locations or patient specific variations in body composition. A "one size fits all" HLUT is simply not going to provide the best possible HU to RSP conversion for all patients.

2.7.4 DECT-based methods

DECT as a tool to reduce range uncertainty in PT treatment planning has been suggested by by several research groups, and has been shown to be superior to SECT based calibration on organic tissue samples [19, 39, 44, 45]. The stopping power of a material depends on two parameters; $\hat{\rho_e}$ and the I-value. With SECT calibration, both of these parameters are estimated through the HLUT. The additional information obtained from imaging with two x-ray spectra in DECT enables us to distinguish these [18] by directly extracting $\hat{\rho_e}$ through a weighted subtraction of the attenuation coefficients from each spectrum [11]:

$$\hat{\rho_e} = \alpha\hat{\mu_H} + (1-\alpha)\hat{\mu_L} \tag{2.15}$$

It is also possible to estimate the so called effective atomic number Z_{eff} of different tissues. As this is not a real physical parameter it needs not be an integer, and can be interpreted as the mean atomic number of a compound material [31]. Z_{eff} can then be used to estimate the I-value with a HLUT-like approach. Although there is still no exact way of extracting the I-value from CT images, this greatly reduces the reliance of the entire calibration on the HLUT approach, as only the less important of the two relevant parameters has to be estimated in this way.

In this thesis the focus will be on two specific DECT calibration methods. The first, MonoCT based calibration, can be considered a mix of DECT and SECT approaches. It offers some range uncertainty improvement while requiring very small changes in the calibration regime. The second is a direct use of DECT data to estimate the RSP, following the RhoSigma approach described by Möhler *et al.* [11]. The reasoning behind these choices are similar. The methods are both easy to implement and already clinically applied at PT centers in Europe. Additionally the RhoSigma approach forms the basis for the upcoming first ever TPS integrated implementation of DECT calibration for PT planning [18, 31]. This makes these two approaches the most relevant candidates for implementation at the new facilities in Norway.

MonoCT based calibration

The ability of MonoCT images to reduce noise, beam hardening effects and thereby metal artifacts, is one of the reasons for implementing these images in PT treatment planning. The method mainly provides more stability in the measured HU values independent placement inside the patient. Another reason is simply that the MonoCT images behave exactly like the regular SECT images in the calibration, making the method easily adaptable as no other changes to the calibration regime is needed.

In 2015, MonoCT calibration made its way into clinical application for the first time at the University Proton Therapy Dresden. The MonoCT based HLUT approach can be seen as a stepping stone towards full implementation of a DECT based approach. Aarhus University Hospital in Denmark also applied a MonoCT based calibration regime recently. In 2018 Möhler *et al.* presented good results from the clinical trials and implemented a pure DECT based calibration method. Several other direct DECT

based RSP calibration regimes have been proposed, many of which are very complex and challenging to implement [11, 31]. Although these are theoretically robust, their complexity might make them ill suited for clinical application.

RhoSigma Calibration

The RhoSigma calibration estimates the RSP directly instead of using a HLUT (the workflow is illustrated in figure 2.16). In this calibration a DECT scan of a calibration phantom, with tissue substitute inserts of known elemental composition, is performed. The x-ray spectra used is typically 80 kVp and 140 kVp with additional tin (Sn) filtration. MonoCT images are made by a weighted sum of the high and low energy CT images, using the parameter α. For a specific value of α, the produced image corresponds to the $\hat{\rho}_e$-image [11], see equation 2.15. Firstly, the α is calibrated based on a DECT scan of bone-tissue substitutes of known elemental composition (explained in detail in section 3.5). The calibration of α characterizes the scanner parameters. A $\hat{\rho}_e$-image and a MonoCT image is then calculated. This can be done directly in the Siemens software Syngo.via.

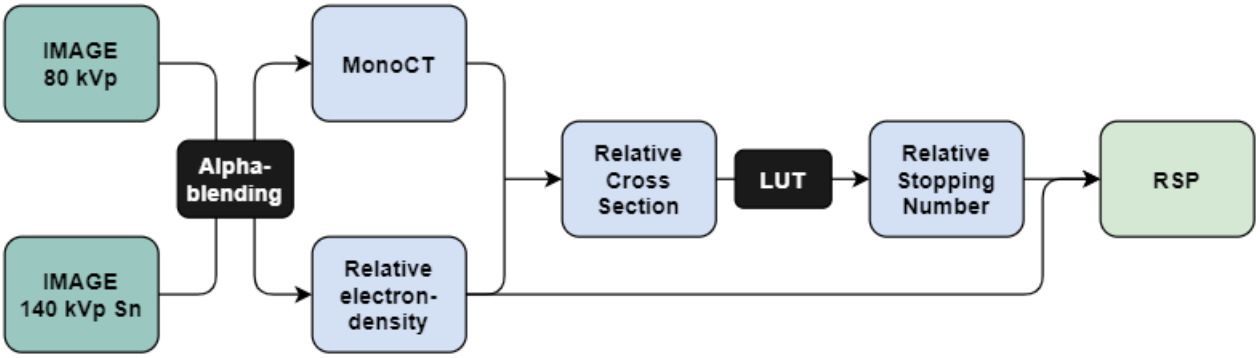

Figure 2.16: RhoSigma calibration workflow. A MonoCT and a relative electron density image (RED) is directly derived by a weighed subtraction of the two image sets (alpha blending). A relative cross section image (RCS) is derived by dividing the MonoCT by the RED. A relative stopping number image (RSN) is derived via a LUT. Finally the RSP image is obtained by multiplying the RED by the RSN.

Dividing the MonoCT by the RED image then yields a RCS image, since $\sigma = \mu/\rho_e$. From this a relative stopping number (RSN) image, can be made, using a RCS to RSN lookup table. This lookup table (LUT) has similar drawbacks as the usual HLUT, but in this case the LUT uncertainties only affect one of the two parameters needed for the calibration. Because of this, the uncertainties from this process are far lower than from the HLUT. The RSN image can then be multiplied with the RED image to directly produce a RSP image, according to equation 2.10.

Based on negligible differences in calibration results obtained when testing this calibration regime on different CT setups and protocols, Möhler *et al.* [11] propose that only one general calibration per scanner type and voltage pair is sufficient. Meaning

that in-house calibrations, at every PT facility, is not necessary. This is a rather bold statement that is likely to be met with some scepticism. Nevertheless, the method has demonstrated excellent results on reducing range uncertainties [18]. Due to this and to its claimed universal applicability, the method will be integrated into Siemens software, likely within the coming year, making its clinical application feasible in the very near future.

2.8 Water Equivalent Path Length

The RSP errors in and of themselves are not translatable in clinical use. To assess the clinical relevancy of the RSP error, a complete dose plan must be produced with a robustness analysis that adjusts the RSP values to within the accepted uncertainties (usually 3.5%) and translates the RSP errors into tumor coverage.. This is a time consuming process. Water equivalent path length (WEPL) is therefore usually used as a surrogate for the RSP error. The WEPL, usually given in millimeters, is used in clinical settings to describe the stopping abilities of a heterogeneous material with a single parameter. Considering the protons traveling through tissues of different stopping abilities, the WEPL scales all these tissues to the equivalent depth of water needed to produce the same integrating stopping power. When evaluating different RSP calculations this is done by calculating the difference in WEPL to the depth of the tumor. This method of assessing target coverage with the given RSP calculation is more time effective although not as precise as it does not account for patient specific treatment parameters such as beam spots, margins, etc. A concrete clinical application of the WEPL is to utilize it in connection to the cone beam CT (taken prior to every treatment fraction) to e.g. receive a warning if the WEPL to a reference point in the patient has changed by a set distance [46].

The WEPL is calculated by integrating all the RSP values along a straight lines in the RSP image. This sum is then multiplied by the pixel size to obtain the WEPL in mm:

$$\mathrm{WEPL} = x \cdot \sum_i \mathrm{RSP}_i \quad , \tag{2.16}$$

where x is the pixel size in mm and the subscript i represents all RPS values along a track through a phantom.

MATERIALS AND METHODS 3

In this chapter the materials and methods used in this study are presented. The materials section show the CT acquisition details (section 3.1.1), phantoms used (section 3.1.2), and software used (section 3.1.3). The calibration procedures described are the SECT calibration (section 3.3), the MonoCT calibration (section 3.4), and the RhoSigma calibration (section 3.5). Lastly the WEPL calculation is described in section 3.6.

3.1 Materials

3.1.1 CT Acquisition

A dual source CT scanner (Siemens SOMATOM Definition Flash, Siemens Healtineers, Forchheim, Germany) belonging to the radiological department of Haukeland University Hospital, was used for all CT scans. The SECT scans were acquired at 120 kVp and the DECT scans at 80 kVp and 140 kVp, with settings as shown in table 3.1.

Table 3.1: CT scan protocols used for SECT and DECT scans. The slice thicknesses used were 5 mm for the calibration phantom and 3 mm (shown in parentheses) for the anthropomorphic phantom

	120 kVp	**80 kVp**	**140 kVp**
Convolution Kernel	B40f	Q40f/2	Q40f/2
Slice Thickness [mm]	5 (3)	5 (3)	5 (3)
Exposure Time [ms]	500	500	500
Tube Current [mA]	480	636	246
Exposure [mAs]	300	530	205
Additional filtration	n/a	n/a	Sn

The following doses were used for the acquisitions: Gammex 20.25 mGy (CTDI 32). Abdomen 20.04 mGy (CTDI 32, for both SECT and DECT).

The SECT and DECT acquisitions were performed in direct succession, without re positioning the phantoms. Nevertheless, the two image sets were not perfectly overlapping. This would cause problems in comparing the modalities, and as such the image sets had to be registered to match each other (see section 3.1.3).

3.1.2 Phantoms

The RMI 467 tissue substitute phantom (Gammex, Middleton, WI, USA) was scanned using both SECT and DECT. This phantom is currently used for clinical photon radiotherapy. This phantom is 32 cm in diameter and contains 16 cylindrical inserts arranged as shown in figures 3.1 and 3.2.

Figure 3.1: Gammex 467 calibration phantom setup

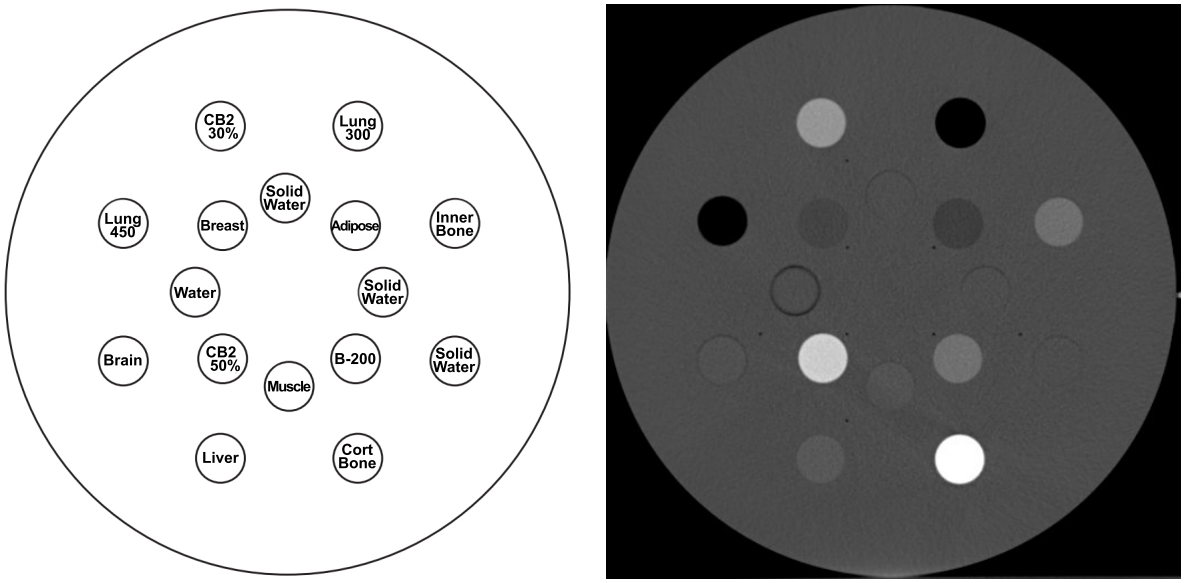

Figure 3.2: Left: arrangement of tissue substitute inserts in the Gammex 467 phantom. Right: CT-image of the same arrangement.

The elemental composition of each insert for this specific phantom, as provided by the manufacturer, is tabulated in table A.1. It is important to use the tabulated elemental

composition of the exact calibration phantom being used for the calibration, as the composition of the inserts may vary slightly between phantoms. This is especially important for the lung inserts where the variations are larger due to air bubbles in the inserts.

For WEPL calculations the anthropomorphic phantom (Kyoto Kagaku, CT Abdomen Phantom PH-5), see figure 3.3, was scanned using both SECT and DECT. For the Gammex and Abdomen phantoms, slice thickness was set to 5 mm and 3 mm respectively. According to previous studies, the slice thickness does not affect the calibrations significantly.

Figure 3.3: Kyoto Kagaku antropomorphic abdomen phantom. Image from [10]

3.1.3 Computer Software

The Siemens software Syngo.via version VA48A was used to make pseudo monoenergetic, and electron density images. Newer versions of syngo will likely offer an integrated method similar to the RhoSigma method used in this thesis.

All programming was done using Python 3.7. The Python code will be made publically available on `github.com/BergenParticleTherapy/DualEnergyCTForProtons`. The SECT image set was registered to the DECT image set using MICE 1.1.3.

3.2 Calibration Procedures

In this section the three HU to RSP calibration methods implemented in this work are described. The first is a stoichiometric calibration on SECT data. The second is a stoichiometric calibration on MonoCT data. The third is based on direct use of DECT data, the RhoSigma calibration. For all calibrations the phantom used was the Gammex 467 calibration phantom. Lastly the anthropomorphic phantom was used for WEPL calculations and comparisons.

3.3 Stoichiometric Calibration: SECT

In order to calculate dose distribution from proton therapy in a patient, the treatment planning software (TPS) needs a mapping that for any given voxel, gives the relationship between its HU value and its RSP. Such a HU to RSP calibration curve is called a Hounsfield lookup table, or HLUT. As the response of the detectors in the CT scanner, the produced energy spectrum, filtration of the beam, beam hardening, anode angle, etc. vary between machines, the HLUT is machine and protocol specific and needs to be calibrated for every CT scanner and protocol used for treatment planning. Today's gold-standard calibration involves performing a so called stoichiometric fit using SECT data of a calibration phantom containing tissue-substitute inserts, as well as tabulated information on the elemental constituents of real human tissues, as proposed by Schneider *et al.* [42]. See figure 2.15 for a visual overview of the calibration.

3.3.1 Calculating Ground Truth RSP

The Gammex calibration phantom was scanned at 120 kV_p as specified in table 3.1, as this is the spectrum that is normally used clinically for treatment planning. The theoretical RSPs for all tissue substitute inserts were calculated using equation 2.9, with the tabulated elemental compositions and electron densities, $\hat{\rho}_e$, see table A.1. The I-values was calculated using the Bragg additivity rule (eq. 2.11) and the tabulated I-values for the elements in a solid state, see table A.2.

3.3.2 Measuring HU's of Inserts

The average HU value for every tissue insert was measured with a python package for reading DICOM files[1], using a circular region of interest (ROI) with radius 50% that of the insert radius. A program was written that automatically measures all 16 ROI's. The measurements were made for the three middle slices and averaged (see table 3.2).

[1]DICOM files are the format of CT images, containing the image itself and all metadata

Table 3.2: Measured HU-value for all 16 inserts in the Gammex 467 calibration phantom, from the SECT images. The inserts were measured in three middle slices and averaged. The average HU was used for the calibration.

	Insert	**Slice 1**	**Slice 2**	**Slice 3**	**Avg HU**
1	CB2-30%	437.6	428.1	416.2	427.3
2	Lung 300	-714.3	-707.2	-699.2	-706.9
3	Inner Bone	190.7	198.2	193.5	194.1
4	Solid Water	-3.6	-0.1	4.6	0.3
5	Cort Bone	1096.3	1134.1	1132.6	1121.0
6	Liver	55.7	64.2	69.9	63.2
7	Brain	20.3	19.4	18.6	19.5
8	Lung 450	-529.7	-530.3	-523.2	-527.8
9	Solid Water	-1.1	-1.2	-6.8	-3.0
10	Adipose	-88.6	-84.6	-86.5	-86.5
11	Solid Water	-10.0	-0.1	-0.1	-3.4
12	B-200	193.1	209.1	206.4	202.8
13	Muscle	20.6	28.2	26.3	25.0
14	CB2-50%	737.2	752.5	741.1	743.6
15	Water	-1.9	-2.8	-12.1	-5.6
16	Breast	-50.3	-50.3	-55.0	-51.9

3.3.3 Characterizing the Scanner

Predicting, or calculating, synthetic HU values without performing a CT scan, requires determination of two constants k_1 and k_2 that account for the polychromatic spectrum of the x-ray source and the energy dependency of the detector in the CT scanner. The scanner specific parameters were determined using the reduced HU, defined as $\hat{\mu}$, in the following form [26]:

$$HU_{red} \;=\; \hat{\mu} \;=\; \hat{\rho}\,\frac{\sum_i \frac{w_i}{A_i}(Z_i + k_1 Z_i^{2.86} + k_2 Z_i^{4.62})}{\frac{w_H}{A_H}(1 + k_1 + k_2) + \frac{w_O}{A_O}(8 + k_1 8^{2.86} + k_2 8^{4.62})}, \tag{3.1}$$

for derivation of this expression see appendix A in Goma *et al.* 2018 [26]. The k_1 and k_2 parameters may be determined by minimizing the expression [42]:

$$\sum_j \left[\left(HU_{red}(k_1, k_2)\right)_j - \left(\frac{HU}{1000} + 1\right)_j \right]^2 , \tag{3.2}$$

where j runs over all tissue-substitutes used for the calibration, i.e. minimizing the squared sum of the differences between the measured and synthetic HU values. This was done using a gradient descent minimization in Python with packages `minimize` from `scipy.optimize`, and `linregress` from `scipy.stats` (see appendix B). A gradient decent minimization is an iterative optimization algorithm for finding a local minimum. The k-values bounds was set to those with physical meaning at imaging energies [41]: $2 \cdot 10^{-4} < k_1 < 6 \cdot 10^{-3}$ and $3 \cdot 10^{-6} < k_2 < 6 \cdot 10^{-4}$.

3.3.4 Synthetic HU and RSP for Human Tissues

With the scanner characterized (k_1 and k_2 determined), synthetic HU values for any material of known elemental composition can be calculated. This was done for the 61 ICRU adult human tissues listed in table A.3 using equation 3.1 and the definition of the CT number equation 2.4. The theoretical RSPs for the human tissues were also calculated, using the same method as for the tissue substitute inserts.

3.3.5 Making and Using the HLUT

To make the HLUT, the human tissue RSPs and synthetic HUs were plotted against each other. Three HU intervals were chosen for the line segments based on typical HU values for lung, soft, and bone tissues. A piecewise linear regression was made, using the chosen intervals and the python package PWLF [47]. This produces three connected line segments, each in the form $RSP = a \cdot HU + b$. This set of functions is the HLUT. The measured insert HU values was input to the piecewise function to obtain experimental RSP values. These were then compared to the theoretically calculated RSP values. The HU intervals of the piecewise fit were adjusted manually a few times to achieve the smallest possible[2] difference between ground truth RSP and experimental RSP. The breakpoints chosen after adjustments were 20HU and 40HU, which is roughly the HU interval for soft tissue and muscle. The resulting HLUT can be seen in figure 3.4.

3.4 Stoichiometric Calibration: MonoCT

The Gammex calibration phantom was scanned using DECT at 80 kVp and 140 kVp/Sn as specified in table 3.1. A 70 keV pseudo monoenergetic CT scan was made from the DECT scan, using the Siemens software syngo.via Monoenergetic Plus. The other steps in the MonoCT calibration is identical to the SECT calibration steps. The measured and averaged HU values used can be seen in table 3.3.

[2]these subjective visual estimates can not in any way guarantee to find the best possible intervals

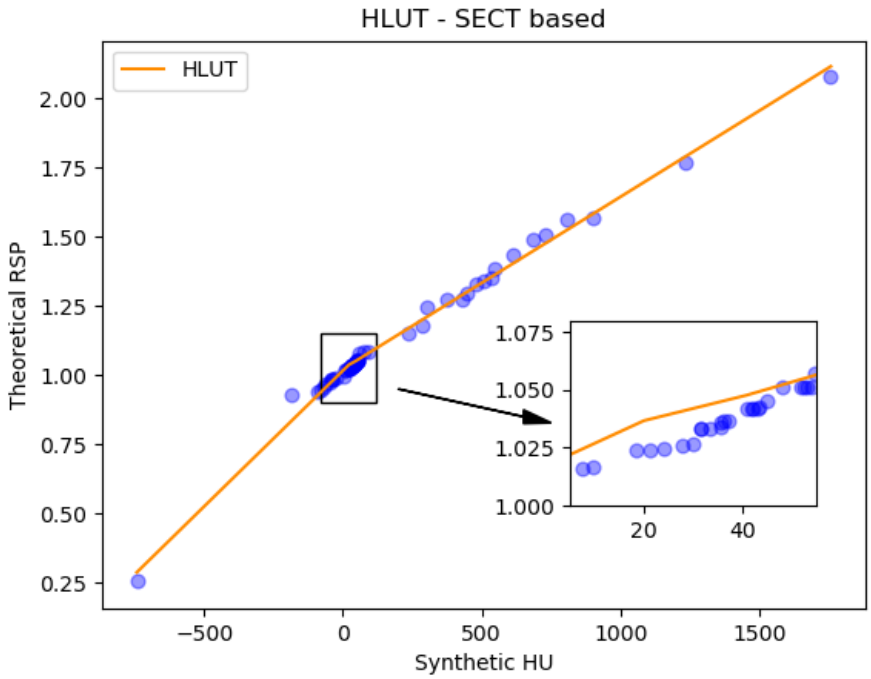

Figure 3.4: HLUT made with SECT stoichiometric calibration based on 61 recommended tissues

Table 3.3: Measured HU-value for all 16 inserts in the Gammex 467 calibration phantom, from the MonoCT scan. The inserts were measured in three middle slices and averaged. The average HU was used for the calibration.

	Insert	**Slice 1**	**Slice 2**	**Slice 3**	**Avg HU**
1	CB2-30%	508.4	503.1	500.6	504.1
2	Lung 300	-720.5	-716.9	-720.4	-719.3
3	Inner Bone	236.8	233.9	233.7	234.8
4	Solid Water	3.5	1.4	2.2	2.4
5	Cort Bone	1358.8	1348.5	1345.9	1351.1
6	Liver	74.7	71.2	72.8	72.9
7	Brain	18.5	16.9	15.6	17.0
8	Lung 450	-533.4	-535.0	-540.2	-536.2
9	Solid Water	2.0	0.5	-0.6	0.7
10	Adipose	-93.6	-95.4	-95.3	-94.8
11	Solid Water	0.2	-1.2	-1.5	-0.8
12	B-200	249.2	248.1	246.1	247.8
13	Muscle	42.5	37.9	36.4	39.0
14	CB2-50%	892.8	889.2	887.2	889.7
15	Water	4.9	3.8	2.3	3.7
16	Breast	-57.4	-59.3	-58.7	-58.4

3.5 RhoSigma Calibration: DECT

Recall figure 2.16 for a visual overview of the calibration. The Gammex calibration phantom was scanned with dual energy at 80 kVp and 140 kVp/Sn as specified in table 3.1. In this approach the RSP is estimated for each voxel in the CT scan. Only one step in this calibration involves a lookup table as opposed to using a HLUT for the entire conversion.

Alpha Blending

When two CT numbers are available for the same voxel, i.e. in DECT, $\hat{\rho_e}$ can be obtained directly by a linear superposition of the two spectrally weighted μs, μ_H and μ_L, with a single blending parameter α [11].

$$\hat{\rho_e} = \alpha\, \hat{\mu_H} + (1 - \alpha)\, \hat{\mu_L} \tag{3.3}$$

The subscripts H and L correspond to the high and low voltage on the tubes in the DECT scan respectively. These will be referred to as S = {H,L} as a collective term.

Eliminating $\hat{\rho_e}$ using the relation $\mu = \hat{\rho_e}\sigma$ for both spectra; $\hat{\mu_H} = \hat{\rho_e}\hat{\sigma_H}$ and $\hat{\mu_L} = \hat{\rho_e}\hat{\sigma_L}$, equation 3.3 becomes

$$\hat{\sigma_H} = \frac{1}{\alpha} + \left(1 - \frac{1}{\alpha}\right)\, \hat{\sigma_L} \tag{3.4}$$

By identifying the best fit between the two spectra, we get a simple parametrisation of the scanner specific effects such as energy spectrum, detector response, and acquisition protocol. The α is calibrated using the Gammex 467 tissue substitutes (see appendix B.2). Only the higher Z materials, i.e. bones, are used due to them giving the most reliable calibration [11]. The calibration can become unstable if materials with $\hat{\rho_e}$ close to 1 is used as this leads to $\hat{\mu_H} \approx \hat{\mu_L}$. The issue is clear when rewriting 3.3 as

$$\alpha = \frac{\hat{\rho_e} - \hat{\mu_L}}{\hat{\mu_H} - \hat{\mu_L}} \tag{3.5}$$

The average HU values for the bone-tissue inserts was measured with a python program, using a circular region of interest (ROI) with radius 50% that of the insert radius. The measurements were made for the three middle slices and averaged. The inserts used were: CB2-30%, Inner Bone, Cortical Bone, B-200, and CB2-50%. Following the steps of Möhler *et al.* [27], by combining eq. 2.2b and 2.4, the relative cross sections $\hat{\sigma_S}$ can be found from the DECT scan via

$$\hat{\sigma_S} = \left(\frac{\mathscr{H}}{1000HU} + 1\right) / \hat{\rho_e}, \tag{3.6}$$

The relative cross sections, $\hat{\sigma_S}$, of the high and low energies were plotted against each other. The α parameter was determined by linear regression to the expression in equation 3.4. With the blending factor α one can use the two original DECT images (80/140Sn kVp) to calculate pseudo monoenergetic images (MonoCT) and relative

electron density images (RED) [11].

The attenuation coefficients were calculated using:

$$\mu_S = \frac{\mathscr{H}_S}{1000} + 1\ , \tag{3.7}$$

then the $\hat{\rho}_e$ images were calculated, using [11]:

$$\hat{\rho}_e\ =\ (1-\alpha)\mu_L\ +\ \alpha\mu_H. \tag{3.8}$$

From the MonoCT and RED images a relative cross section image (RCS) can be made by the simple relation

$$\text{RelativeCrossSection} = \frac{\text{MonoCT}}{\text{RED}} \tag{3.9}$$

The relation

$$\hat{\sigma}\ \ =\ \ \sum_i \frac{\hat{\rho_{e_i}}}{\sum_j \hat{\rho_{e_j}}}\ \hat{\sigma_i}\ \ =\ \ \sum_i \nu_i \hat{\sigma_i} \tag{3.10}$$

is obtained from equation 3.9 and the attenuation sum for compound materials

$$\hat{\mu} = \sum_i \hat{\mu_i}. \tag{3.11}$$

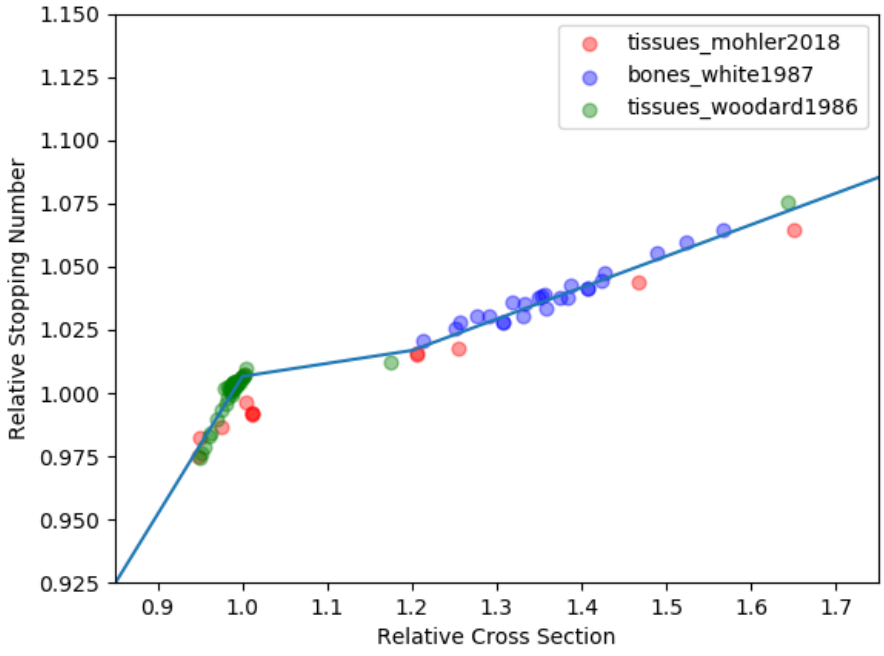

Figure 3.5: Lookup table for relative cross section to relative stopping number conversion. The calibration procedure is the same as described in Möhler *et al.* [11]

To make the RCS to RSN lookup table, relative cross sections and stopping numbers were calculated for a number of materials. The data used to calculate the relative

cross sections was taken from the NIST XCOM database [48]. The cross sections and stopping numbers were then plotted against each other and the lookup table was made using the same principles as the regular HLUT approach (see figure 4.3). With this approach, $\hat{\rho_e}$ is obtained directly, and only the relative stopping number $\hat{L}$ relies on a lookup table approach.

Based on the Bragg's additivity rule for stopping powers [49], eq. 2.9, and eq. 3.10, the stopping number $\hat{L}$ can be found through

$$\hat{L} = \sum_i \nu_i \, \hat{L}_i. \tag{3.12}$$

With the LUT defined, the RSN image can be made from the RCS image. Finally the RSP image can be calculated using eq. 2.10, namely by multiplying the RSN image by the RED image.

3.6 WEPL calculation

Calculating a theoretical ground truth, or reference, RSP for the Gammex inserts, is possible because their compositions are known. Thus the RSP predictions of the SECT, MonoCT, and DECT calibrations can be compared to this reference RSP. When calculating the WEPL projections of a patient, or in the case of this study an antropomorphic abdomen phantom, no ground truth can be calculated unless the ground truth WEPL values are measured with a proton beam, or the phantom-specific materials are fully known. The three calibration methods are therfore compared in relation to each other. Specifically the MonoCT and DECT calibrations were both compared to the SECT calibration as this is the state-of-the-art calibration. The comparison is done by subtracting the SECT based WEPL from the MonoCT based WEPL, and the DECT based WEPL in turn (see section 4.4.2). The root mean square error (RMSE) of the difference is then obtained. To assign meaning to the obtained RMSE, one needs to know which of the calibrations being compared, is theoretically more accurate. This is decided based on the methods reported MAE of the experimental RSP to the reference RSP. The value of the RMSE then describes how much more accurate, on average, the method with the lowest MAE is, compared to the other.

For the WEPL calculations, SECT and DECT based images of the abdomen phantom were used. RSP maps for all the image slices was calculated so as to make WEPL projections of the entire phantom, not just one axial slice. The RSPs were integrated along the beam path through the image slices and multiplied with the pixel spacing, to get the WEPL in mm. By subtracting the SECT-based WEPL from the MonoCT and DECT based WEPL projections, two WEPL difference maps were generated. Any difference in the calculated WEPLs was assumed to be in favour of the most accurate RSP calibration method of the two being compared. The difference of the WEPLs were calculated as root mean square of the entire projections. Only data from inside the phantom was used in the WEPL comparison. This was done by setting a lower threshold for RSP values included in the WEPL subtraction, so that the air outside the phantom was not included.

In the process of comparing the SECT-based and DECT-based WEPLs, it was discovered that the SECT and DECT image sets did not overlap perfectly. This occured even though the two image acquisitions were made in direct succession, without moving the phantom. The SECT image set thus had to be registered to match the DECT-based image set before comparing the WEPL projections.

RESULTS 4

In this chapter the results of the studied methods are presented. Sections 4.1, 4.2, and 4.3, show the results from the three calibration methods that were implemented in this thesis. Section 4.4 contain comparisons of the three methods via error-plots and WEPL projection comparisons.

The accuracy of the RSP prediction of each method was estimated in comparison to the calculated theoretical RSP values for the Gammex calibration phantom. These were calculated as described in section 3.3 and are considered the ground truth, RSP_{ref}, for all methods, see table 4.1.

Table 4.1: Theoretical RSP values for the Gammex 467 phantom inserts used in this study

Insert:	**CB2-30**	**Lung300**	**Inner B**	**Solid W**	**Cort Bone**	**Liver**	**Brain**	**Lung450**
RSP ref:	1.276	0.281	1.084	0.99	1.679	1.065	1.052	0.458
(cont.)	**Solid W**	**Adipose**	**Solid W**	**B-200**	**Muscle**	**CB2-50**	**Water**	**Breast**
	0.99	0.931	0.99	1.10	1.02	1.462	1.00	0.957

4.1 Stoichiometric Calibration on SECT data

The scanner was characterized, i.e. the two k-parameters were determined, following the method described in section 3.3. For the SECT-based calibration this resulted in the k-values: $k_1 = 2.000 \cdot 10^{-4}$ and $k_2 = 2.005 \cdot 10^{-5}$. The HLUT based on 61 recommended ICRU tissues theoretical RSPs and synthetic HUs can be seen in figure 3.4. This HLUT has three linear segments with the two breakpoints at HU = 20 and HU = 40.

The HLUT was used on the measured HU values of the Gammex inserts to obtain the experimental RSP, $\mathrm{RSP}_{\mathrm{exp}}$, see table 4.2. The error, $\mathrm{RSP}_{\mathrm{exp}} - \mathrm{RSP}_{\mathrm{ref}}$, relative error $(\mathrm{RSP}_{\mathrm{exp}} - \mathrm{RSP}_{\mathrm{ref}})/\mathrm{RSP}_{\mathrm{ref}}$, and absolute error, $|\mathrm{RSP}_{\mathrm{exp}} - \mathrm{RSP}_{\mathrm{ref}}|$, was calculated for all inserts. As can be seen in table 4.2, the relative error becomes very large for the lung tissues in this calibration, with 18.9% and 9.4% relative error for Lung300 and Lung450. The lung tissues are often excluded from the calibration for this reason. The mean relative error (MRE) was calculated to 3.3%. Bear in mind that in the MRE the positive and negative errors cancel. The mean absolute error (MAE) was calculated both including, and excluding, the lung tissues to be 2.6% and 2.3% respectively (the abbreviation MAE-L is used for the MAE excluding lung tissues).

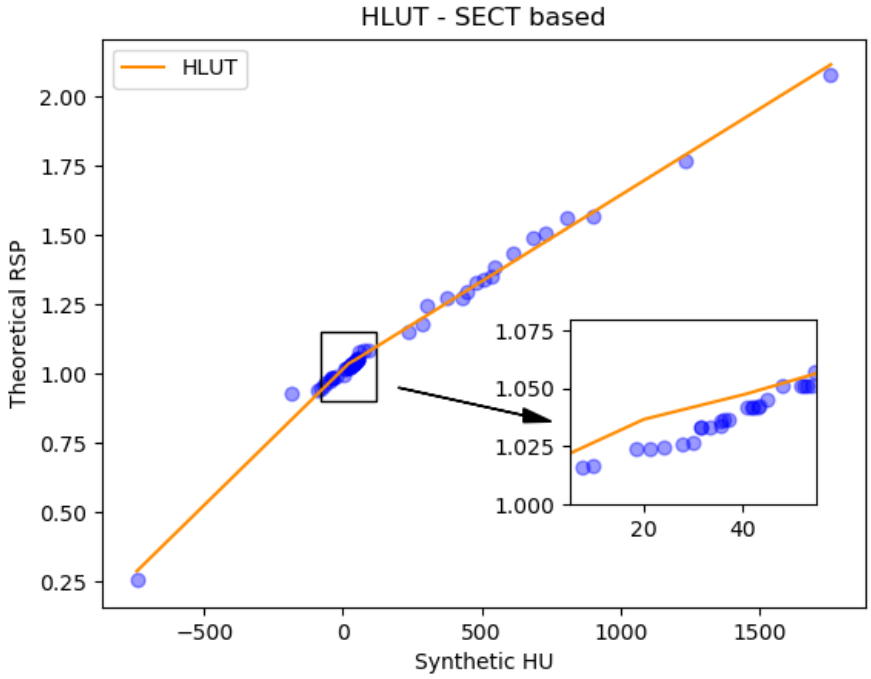

Figure 4.1: HLUT (orange line) made with stoichiometric calibration based on 61 recommended tissues (blue dots)

Table 4.2: SECT - RSP prediction errors. RSPref: calculated from tabulated compositional info. RSPexp: estimated with HLUT. Rel Err: RSPexp - RSPref / RSPref. Err: RSPexp - RSPref. Abs Err: abs(RSPexp - RSPref). MRE: Mean relative error, MAE. Mean absolute error, MAE-L: mean absolute error(no lung). STD: Standard deviation of Errors.

SECT (120kVp)					
Insert	**RSP ref**	**RSP exp**	**Rel Err**	**Err**	**Abs Err**
CB2-30	**1.276**	1.283	0.55%	0.70%	0.70%
Lung300	**0.281**	0.334	18.86%	5.30%	5.30%
Inner B	**1.084**	1.144	5.54%	6.00%	6.00%
Solid W	**0.990**	1.020	3.03%	3.00%	3.00%
Cort Bone	**1.679**	1.720	2.44%	4.10%	4.10%
Liver	**1.065**	1.064	-0.09%	-0.10%	0.10%
Brain	**1.052**	1.034	-1.71%	-1.80%	1.80%
Lung450	**0.458**	0.501	9.39%	4.30%	4.30%
Solid W	**0.990**	1.013	2.32%	2.30%	2.30%
Adipose	**0.931**	0.934	0.32%	0.30%	0.30%
Solid W	**0.990**	1.016	2.63%	2.60%	2.60%
B-200	**1.100**	1.152	4.73%	5.20%	5.20%
Muscle	**1.022**	1.041	1.86%	1.90%	1.90%
CB2-50	**1.462**	1.483	1.44%	2.10%	2.10%
Water	**1.000**	1.011	1.10%	1.10%	1.10%
Breast	**0.957**	0.963	0.63%	0.60%	0.60%
		MRE	**3.31%**	**MAE**	**2.59%**
		STD	**0.022**	**MAE-L**	**2.27%**

4.2 Stoichiometric Calibration on MonoCT data

The MonoCT-based scanner characterization resulted in the k-values: $k_1 = 2.000 \cdot 10^{-4}$ and $k_2 = 3.185 \cdot 10^{-5}$. Two HLUTs were made, one with three line segments and breakpoints at HU = 20 and HU = 40 (identical to the SECT based HLUT), and one with five line segments with breakpoints at HU = −30, HU = 0, HU = 35, and HU = 60. The resulting HLUTs can be seen in figure 4.2.

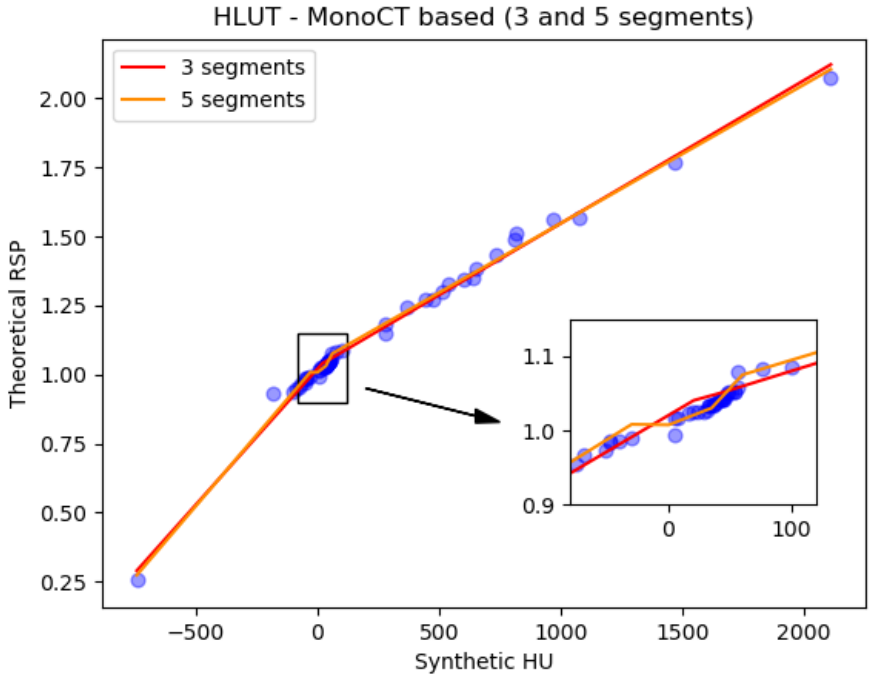

Figure 4.2: HLUTs (red and orange lines) made with stoichiometric calibration MonoCT data and 61 recommended tissues (blue dots)

Both HLUTs were used on the measured HU values of the Gammex inserts to obtain the experimental RSPs, see table 4.3. The error, relative error, and absolute error, was calculated for all inserts (see table 4.3). From the HLUT with three segments: MRE = 3.1%, MAE = 2.8%, and MAE-L = 2.8%. From the HLUT with five segments: MRE = 3.6%, MAE = 2.8%, and MAE-L = 2.9%.

Both MonoCT HLUTs thus performed poorer than the SECT HLUT in predicting the RSP of the tissue inserts. The MonoCT based predictions were however more stable for the lung tissues with relative errors of 11.4% and 7.9% for Lung300 and Lung450.

Table 4.3: MonoCT - RSP prediction errors. RSPref: calculated from tabulated compositional info. RSPexp: estimated with HLUT. Rel Err: RSPexp - RSPref / RSPref. Err: RSPexp - RSPref. Abs Err: abs(RSPexp - RSPref). MRE: Mean relative error, MAE. Mean absolute error, MAE-L: mean absolute error(no lung). STD: Standard deviation of Errors.

		Mono (70keV)				Mono (70keV) (5 SEGMENTS)			
Insert	**RSP ref**	**RSP exp**	**Rel Err**	**Err**	**Abs Err**	**RSP exp**	**Rel Err**	**Err**	**Abs Err**
CB2-30	**1.276**	1.291	1.18%	1.50%	1.50%	1.299	1.80%	2.30%	2.30%
Lung300	**0.281**	0.313	11.39%	3.20%	3.20%	0.298	6.05%	1.70%	1.70%
Inner B	**1.084**	1.150	6.09%	6.60%	6.60%	1.163	7.29%	7.90%	7.90%
Solid W	**0.990**	1.023	3.33%	3.30%	3.30%	1.009	1.92%	1.90%	1.90%
Cort Bone	**1.679**	1.730	3.04%	5.10%	5.10%	1.725	2.74%	4.60%	4.60%
Liver	**1.065**	1.066	0.09%	0.10%	0.10%	1.082	1.60%	1.70%	1.70%
Brain	**1.052**	1.037	-1.43%	-1.50%	1.50%	1.018	-3.23%	-3.40%	3.40%
Lung450	**0.458**	0.494	7.86%	3.60%	3.60%	0.488	6.55%	3.00%	3.00%
Solid W	**0.990**	1.021	3.13%	3.10%	3.10%	1.008	1.82%	1.80%	1.80%
Adipose	**0.931**	0.927	-0.43%	-0.40%	0.40%	0.942	1.18%	1.10%	1.10%
Solid W	**0.990**	1.020	3.03%	3.00%	3.00%	1.007	1.72%	1.70%	1.70%
B-200	**1.100**	1.157	5.18%	5.70%	5.70%	1.170	6.36%	7.00%	7.00%
Muscle	**1.022**	1.047	2.45%	2.50%	2.50%	1.030	0.78%	0.80%	0.80%
CB2-50	**1.462**	1.491	1.98%	2.90%	2.90%	1.494	2.19%	3.20%	3.20%
Water	**1.000**	1.024	2.40%	2.40%	2.40%	1.010	1.00%	1.00%	1.00%
Breast	**0.957**	0.963	0.63%	0.60%	0.60%	0.979	2.30%	2.20%	2.20%
		MRE	**3.12%**	**MAE**	**2.84%**	**MRE**	**2.63%**	**MAE**	**2.83%**
		STD	**0.022**	**MAE-L**	**2.76%**	**STD**	**0.026**	**MAE-L**	**2.90%**

4.3 RhoSigma Calibration on DECT data

The two DECT spectra were calibrated according to section 3.5 to obtain the best match for the alpha parameter. This calibration yielded $\alpha = 1.494$, which is close to what other studies have found. The software syngo.via was used to produce the MonoCT (70keV) and the relative electron density image (RED). This was also done by Möhler *et al.* The relative cross section image (RCS) was made by dividing the MonoCT image by the RED image. The RCS to relative stopping number (RSN) LUT was made following the method described in section 3.5. Our LUT behaves almost as expected, except that it is inverted compared to the LUT produced by Möhler *et al.*[50], see figure 4.3.
Using the LUT a $\hat{\sigma}$-image was made. This image was then multiplied with the $\hat{\rho}_e$-image, to produce the final RSP-image mapping the RSP of every pixel in the original CT images, see figure 4.4.

The RSP-values of the insert was then measured by averaging a ROI inside each insert in the RSP image, see table 4.4. From the RSP measurements the following : MRE = 1.0%, MAE = 1.3%, and MAE − L = 1.4%. Note that this method performs better when including the lung tissues than if they were not included.

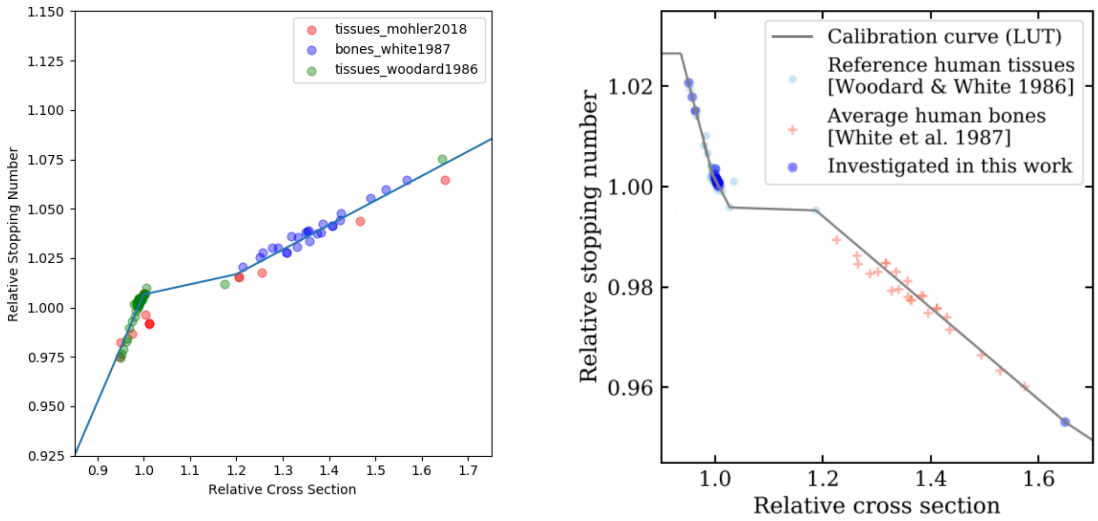

Figure 4.3: Lookup table for relative cross section to relative stopping number conversion. The calibration procedure is described in section 3.5 and is the same as in Möhler *et al.* [11]. Left: LUT produced in this thesis. Right: LUT from Möhler et al.

Table 4.4: DECT - RSP prediction errors. RSPref: calculated from tabulated compositional info. RSPexp: estimated with RhoSigma calibration. Rel Err: RSPexp - RSPref / RSPref. Err: RSPexp - RSPref. Abs Err: abs(RSPexp - RSPref). MRE: Mean relative error, MAE. Mean absolute error, MAE-L: mean absolute error(no lung). STD: Standard deviation of Errors.

RhoSigma (DECT 80/140Sn)

Insert	RSP ref	RSP exp	Rel Err	Err	Abs Err
CB2-30	**1.276**	1.276	0.00%	0.00%	0.00%
Lung300	**0.281**	0.285	1.42%	0.40%	0.40%
Inner B	**1.084**	1.089	0.46%	0.50%	0.50%
Solid W	**0.990**	0.998	0.81%	0.80%	0.80%
Cort Bone	**1.679**	1.652	-1.61%	-2.70%	2.70%
Liver	**1.065**	1.088	2.16%	2.30%	2.30%
Brain	**1.052**	1.081	2.76%	2.90%	2.90%
Lung450	**0.458**	0.464	1.31%	0.60%	0.60%
Solid W	**0.990**	1.000	1.01%	1.00%	1.00%
Adipose	**0.931**	0.954	2.47%	2.30%	2.30%
Solid W	**0.990**	0.998	0.81%	0.80%	0.80%
B-200	**1.100**	1.100	0.00%	0.00%	0.00%
Muscle	**1.022**	1.036	1.37%	1.40%	1.40%
CB2-50	**1.462**	1.446	-1.09%	-1.60%	1.60%
Water	**1.000**	1.010	1.00%	1.00%	1.00%
Breast	**0.957**	0.985	2.93%	2.80%	2.80%
		MRE	**0.99%**	**MAE**	**1.32%**
		STD	**0.015**	**MAE-L**	**1.44%**

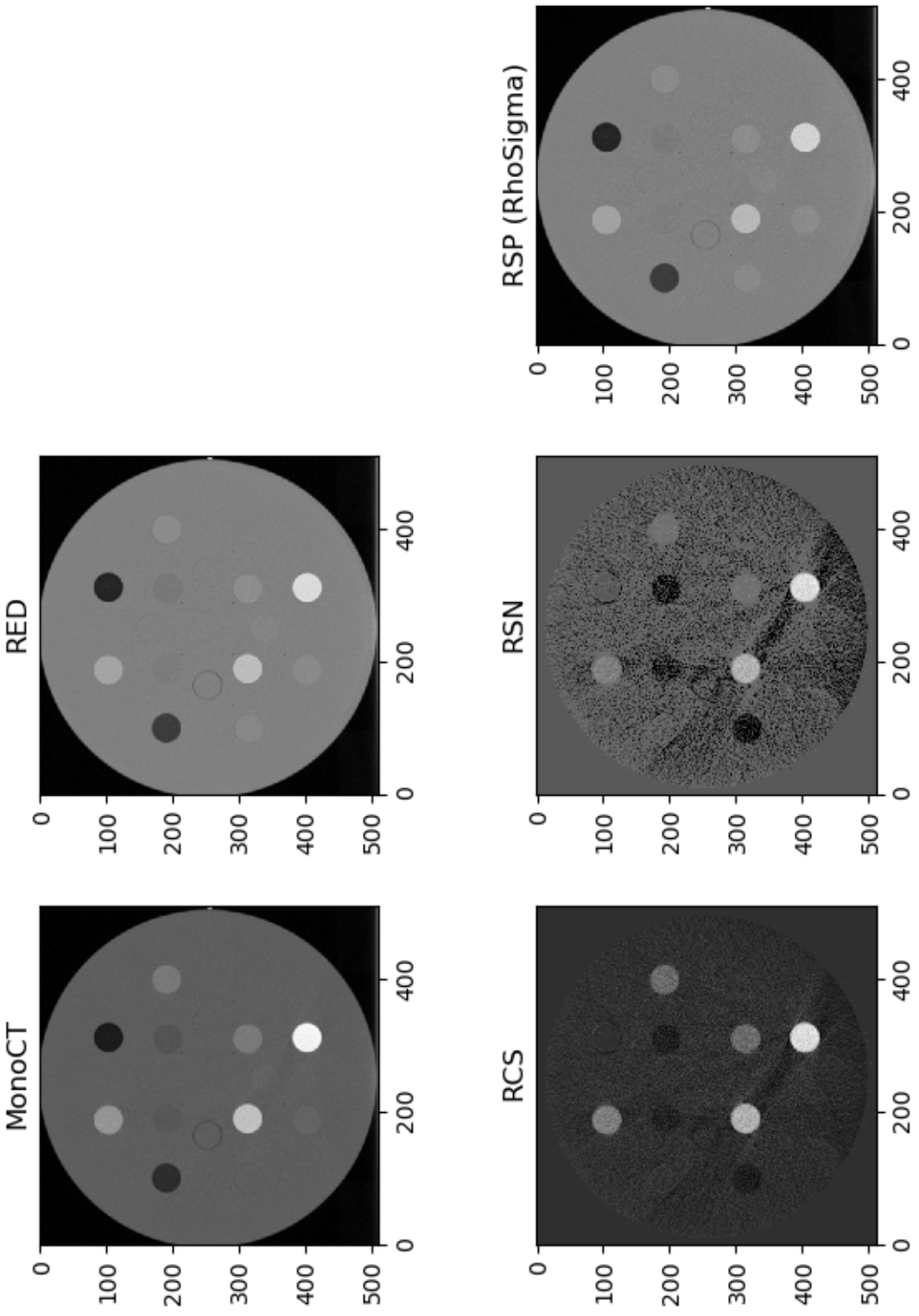

Figure 4.4: Steps of the RhoSigma approach. MonoCT: pseudo monoenergetic image, RED: relative electron density image, RCS: relative cross section image, RSN: relative stopping number image, RSP: relative stopping power image.

4.4 Comparisons

4.4.1 RSP prediction accuracy

Table 4.5 shows accuracy of the RSP predictions of all the applied calibration regimes. Surprisingly, the SECT-based calibration performs better than both MonoCT-based calibrations, including and excluding lung tissues. The barplots in figure 4.5 shows a visual representation of the actual difference and relative difference in RSP prediction for all the methods. Only one MonoCT-dataset, from the HLUT with 3 segments, is included in this plot.

Table 4.5: Uncertainty analysis for all calibration methods used in this study. MAE: Mean absolute error, MAE-L: Mean absolute error excluding the lung tissues Lung300 and Lung450, Err STD: The standard deviation in the actual error between ground truth and experimental RSP. SECT: Stoichiometric calibration on SECT data, MonoCT-3: Stoichiometric calibration on MonoCT data using 3 segments in the HLUT, MonoCT-5: Stoichiometric calibration on MonoCT data using 5 segments in the HLUT, DECT: RhoSigma calibration using DECT data.

Methods	**MAE**	**MAE-L**	**Err STD**
SECT	2.59%	2.27%	0.022
MonoCT-3	2.84%	2.76%	0.022
MonoCT-5	2.83%	2.90%	0.026
DECT	1.32%	1.44%	0.015

These representations show how the relative difference becomes unstable for lung tissues in the HLUT-based calibrations. These are usually excluded from the calibration for this reason. Note, however that the DECT-based calibration performs very well for the lung tissues both in relative difference and in actual difference. There is a 42% reduction in MAE from the best performance of the HLUT calibrations: 2.27% to the best performance of the DECT calibration 1.32%. Additionally the DECT-based calibration has a lower standard deviation in absolute error, meaning that it is less vulnerable to which tissues are used in the calibration.

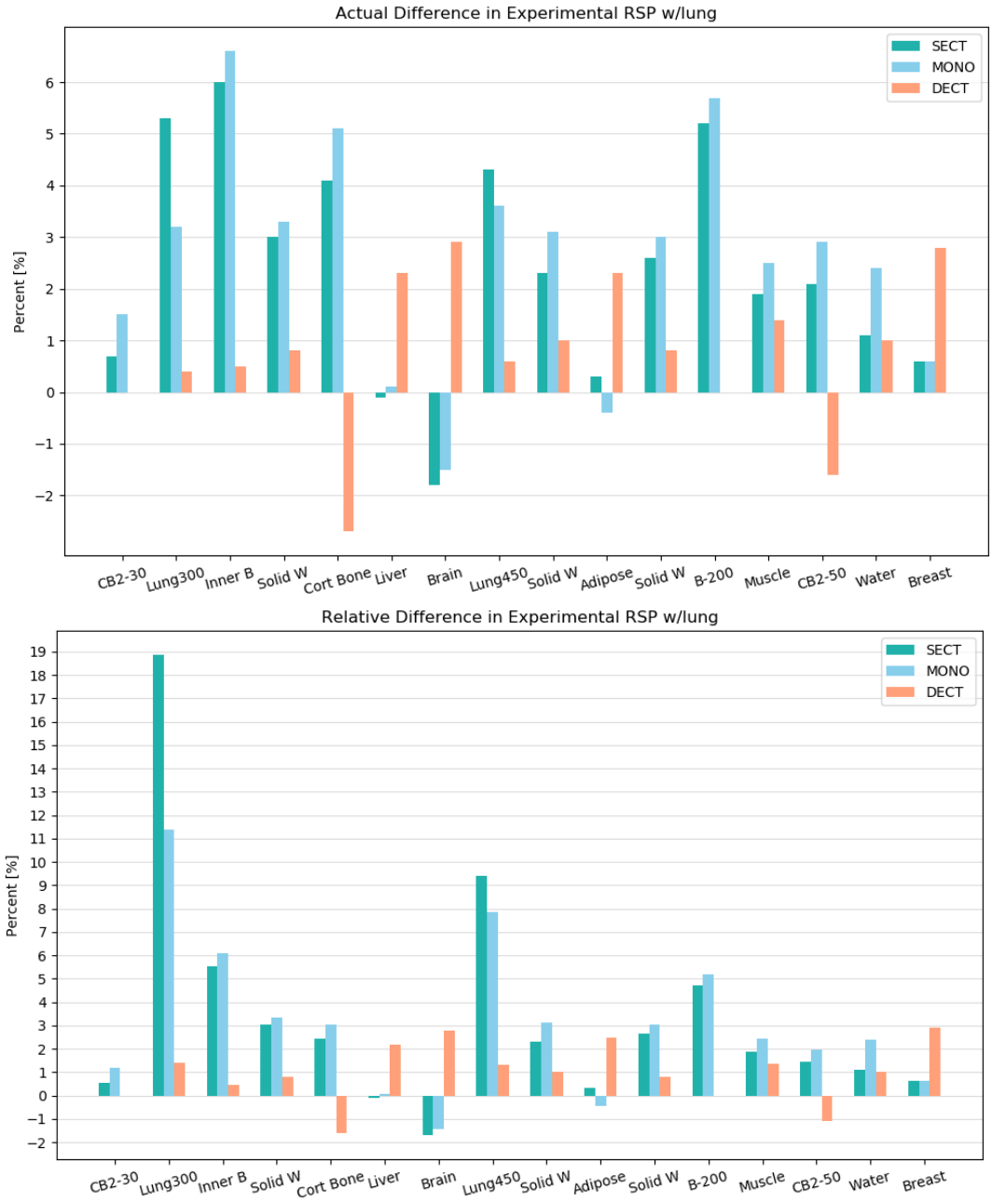

Figure 4.5: Top: Difference in percentage points between ground truth between ground truth and experimental RSP (RSPexp-RSPref). Bottom: Relative difference between ground truth and experimental RSP.

4.4.2 WEPL results

The WEPL projections for all methods were calculated, including only the MonoCT-based method using 3 segments in the HLUT. These WEPL projections can be seen in the upper row of figure 4.6, where the two first columns are the WEPL projections for the SECT and DECT based calibrations. The third column is the difference in the WEPL predictions between the two calibrations where the SECT-WEPL has been subtracted from the DECT-WEPL. The air outside of the phantom was removed from the comparison. White color indicates that the two WEPL projections are in agreement. Since the DECT based calibration has been shown, in other studies and in this thesis, to be more accurate than the SECT based calibration, any difference between the two WEPL calculations will be considered to favour the DECT based calibration.

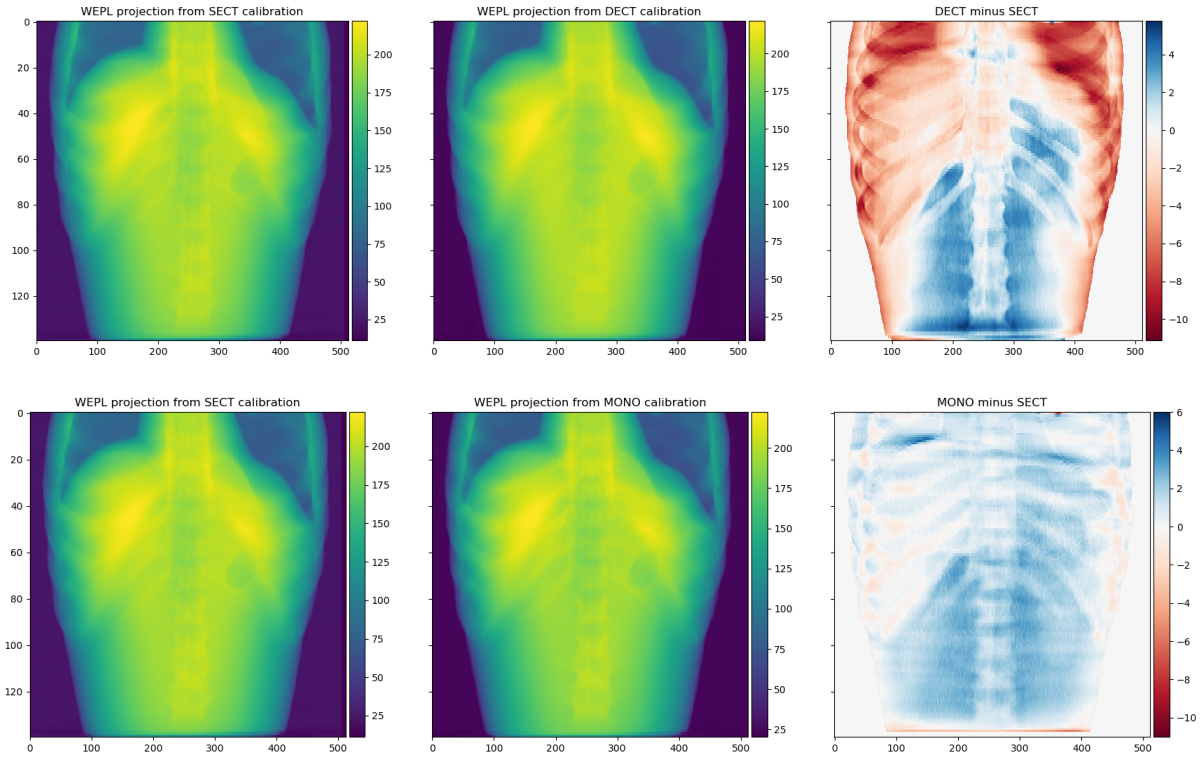

Figure 4.6: Top: Water equivalent path length (WEPL) projections from the SECT calibration and the DECT calibration. The RMSE is 3.16 mm in favour of the DECT calibration. Bottom: WEPL projections from the SECT calibration and the DECT calibration. The RMSE is 1.41 mm in favour of the SECT calibration

The RMSE, or the root mean square error, of the difference-image was calculated to be: 3.16 mm. This translates to a 3.16 mm mean error in range prediction using the SECT based calibration compared to the DECT based calibration, when the whole image is considered.

The WEPL comparison of the SECT and MonoCT based projections can be seen in

figure 4.6. Since the MonoCT based RSP prediction was shown to be less accurate than the SECT based calibration in this study, any difference between the two WEPL projections is considered to be in the favour of the SECT based calibration. The RMSE of the difference-image was calculated to be 1.41 mm. Thus, in this case the SECT based calibration is more accurate than the MonoCT based calibration.

5 DISCUSSION

In this thesis, three calibration methods for calculating RSP from CT-numbers were investigated and compared. A stoichiometric calibration was made using SECT. From the DECT images, two calibrations were made; a stoichiometric calibration using MonoCT and a RhoSigma calibration using the DECT data directly. In this section the MonoCT and direct DECT calibration will be discussed in comparison to the state-of-the-art SECT based calibration, and to each other. The SECT based HLUT calibration performed as expected from previous literature. A MAE of 2.6% in RSP prediction, seems reasonable as the typical 3.5% range margin accounts for all clinical uncertainties, not only the theoretical RSP prediction and that the CT related uncertainty often is reported as being in the order of 2% to 3% [51]. This makes this method a good comparative baseline method for discussing the other calibrations.

The clinical impact of reducing range calculation uncertainties in proton therapy, can be significant if it leads to the reduction of treatment margins. With the margins used today, a notable amount of healthy tissue recieves the same dose as prescribed for the tumor. Reducing this volume of radiated healthy tissue can reduce the late side effects seen in treated patients. This is especially the case for pediatric patients.

5.1 MonoCT

The MonoCT stoichiometric calibration method was expected to provide more stability and slightly better, or equal, RSP prediction compared to the SECT based calibration. Better stability in lung tissue was achieved, however the overall performance of the method was poorer than the state-of-the-art SECT calibration, with MonoCT calibration resulting in a MAE of 2.8%. Further, the MonoCT HLUT with five segments performed worse (MAE of 2.9%) than the MonoCT HLUT with three segments. We have not been able to find any obvious mistakes in the implementation of this method, or otherwise managed to identify why the RSP prediction from MonoCT is less accurate than expected. As shown in section 3.4, this method is the exact same as the SECT based method, except for the input data, which did perform as expected.

While the implementation of the MonoCT calibration at other facilities has been mostly due to other advantages, such as reduced beam hardening, than improved RSP prediction, the method has usually been reported to yield equal or slightly better RSP prediction than SECT calibration. This suggests that negative results might be

under-reported, or more likely, that there are more accurate ways of implementing the method than what has been demonstrated in this thesis. One step of this calibration that should be improved is the explorative approach to choosing the break points of the HLUT. The most precise directions for this step found in the literature, is that the intervals should be chosen to represent the different tissue types. If this can be done in a more sophisticated way, the calibration might yield a more optimal fit, and thus better results. Krah *et al.* [52] investigated methods for optimizing the HLUT (albeit with a radiograph as the input) and assessed the effects of differently placed breakpoints. They found an agreement with the used ground truth better than 0.5%.

Another possible point of investigation is the CT dose while acquiring the DECT scan from which the MonoCT is made. As DECT images, and followingly MonoCT images, are often acquired so that the patient dose is equal to that of a SECT scan, these images have more noise. If allowing a higher patient dose during the DECT acquisition leads to better accuracy in RSP prediction, it could be investigated whether the improved range calculation makes up for the added dose during the planning CT.

5.1.1 k-values

For both stoichiometric calibrations (SECT and MonoCT), k_1 was calculated to be equal to the lower bound value of the optimization algorithm: $2.000 \cdot 10^{-4}$. The behaviour of k_1 in the optimization was asymptotic towards (and beyond) its lower bound. It was therefore set to the value of the lower bound in accordance similar studies [26]. Nevertheless, varying the breakpoints of the HLUT had a much greater impact on the calibration than varying the k-values. As such it reasonable to assume that the uncertainty correlated to breakpoints outweigh any potential uncertainty correlated to the k-values.

5.2 DECT

The results of the DECT-based RSP prediction show that this method has great potential to reduce range uncertainties in proton therapy. The method showed superior stability in low density tissues as well as better overall stability, compared to today's gold standard HLUT approach. This is in accordance with existing literature. The results of other groups suggest that the MAE should be below one percent [20], to which our results (MAE of 1.32%) are close. Although this MAE is slightly larger than one percent, it is still more than one percent point lower than the best MAE of the other predictions which is 2.27% for the SECT calibration excluding lung tissue.

There occurred an unresolved problem while implementing the DECT calibration, namely that the relative cross section to relative stopping number LUT appeared inverted. This phenomenon leads to an error, of small yet unknown magnitude, in the estimated stopping number of the tissue substitutes. Fortunately the stopping number is not the main contributor to the final RSP, and as such the calibration performs well even with this flaw. The reason for this inversion should however be investigated and resolved. With this error corrected, the calibration will likely yield the expected below

one percent MAE in RSP prediction.

5.2.1 WEPL

The RMSE of the WEPL difference between the DECT and the SECT calibration was 3.16 mm. Since the DECT was proven to be more accurate from the RSP analysis, the improvement of the proton range accuracy in an anthropomorphic phantom is thus in the order of 3 mm. Clinically the WEPL difference will be somewhat lower since one does not integrate over the entire volume of the patient, but only to the depth of the tumor site. Investigation of this type falls beyond the scope of this work. Yet, this reported RMSE support that the DECT calibration is more accurate than the state-of-the-art SECT calibration.

5.3 Implementation

The process of implementing the different RSP calibrations has been error-prone and complex. While investigating the literature on this topic, several questions arose that were not answered. How should the HU intervals of the HLUT breakpoints be chosen and how many should there be? How should the segments of the HLUT be connected? Further, there are no recommendations on what optimization algorithms should be used to determine the k-values. Neither are there guidelines on CT protocols, reconstruction algorithms, or methods of measuring the HU values. These are only a few of the questions that were encountered.

There are thus, many considerations to keep in mind when implementing the required steps, and there are many opportunities for small errors to pass unnoticed. This process is truly more complex than it appears at first sight, which is the case for the state-of-the-art stoichiometric calibration and to some degree for the RhoSigma calibration, although the latter is less complex. After correspondence with some proton therapy facilities in Europe, the consensus regarding protocols for calibrations between proton facilities, seems to be that as long as uncertainties are under control, everything is up for interpretation. This is both reassuring and worrying as so many steps are left to the interpretation of the physicists implementing the calibration. This also makes quality assurance harder and inter-center variations bigger. There is a need for international guidelines, and ICRU are expected provide this in 2021.

Results aside, in working with both the stoichiometric calibration and with the RhoSigma calibration, the latter proved easier to implement. The RhoSigma calibration has fewer steps, which translates to less room for personal interpretation and therefore less misinterpretations or mistakes. There is a possibility that this statement is somewhat biased, due to lack of knowledge on the topic at hand in the beginning of this project and that the stoichiometric calibration was the first method to be implemented. Nevertheless, it must be stressed that the implementation of the RhoSigma calibration was quicker, performed better, and that it requires less heuristic assumptions, than the

stoichiometric calibration.

A way of implementing the benefits of DECT into existing protocols is by retrospective adaptation of a MonoCT HLUT based on patient data, as done by Wohlfahrt *et al.*, in 2020 [53]. This adaptation of the HLUT resulted impressive reduction of the proton range down to approximately 0.2%.

5.4 Implications and Future Work

The results in this thesis show that there are still some unresolved issues in both the MonoCT and the DECT calibrations, as implemented in this study. As routines for the new particle centers in Norway are being developed, these should be investigated further. Contrary to the current literature, the results in this thesis showed that MonoCT calibration does not improve the RSP prediction compared to the state-of-the-art SECT calibration. The method should be investigated further to uncover why the results disagree with literature. Unfortunately, we can not point to any specific part of this calibration that is responsible for the negative result, although the method of determining the break points should be looked into. If better results can be achieved, the method should be considered for potential clinical implementation as a stepping stone towards a full DECT based calibration.

The DECT based RhoSigma calibration, on the other hand, has been shown in this thesis, to improve the RSP prediction and thus the range calculation of protons. This is in agreement with existing literature. The recent review article by Wohlfahrt *et al.* [18] on the topic, concluded that there is substantial and broad evidence for the clinical benefits of DECT in proton treatment planning. This conclusion was based on more than 20 studies assessing the accuracy of RSP prediction with various DECT algorithms. For the RhoSigma calibration specifically, Möhler *et al.* [27] concluded that the maximum RSP uncertainty was below 1%. If results like these are clinically attainable, reduction of treatment margins should be considered and investigated furter.

If the RhoSigma calibration method is considered for clinical implementation, the inverted LUT needs to be resolved. The method can also be further improved upon by including patient specific weighted calibrations such as in Wohlfahrt *et al.* [53].

5.5 Conclusions

In this thesis, methods for proton range calculation using SECT and DECT were implemented and compared. Based on the results of this study, the DECT calibration method is concluded to improve range calculations in proton therapy treatment planning from 2.3% using SECT to 1.3% using DECT. This conclusion is in agreement with existing literature. If clinically implemented, such a range uncertainty reduction, might lead to the reduction of treatment margins which, in turn, can contribute to less irradiation of healthy tissue and its related side effects. Based on the results presented in this thesis and on the investigation of existing literature, it should therefore be considered to bypass the un-adapted state-of-the-art stoichiometric calibrations, and focus on establishing a sound DECT-based calibration regime.

The work done in this thesis has shown the clear potential of DECT to reduce uncertainties of range calculations for proton therapy, although further development of calibration methods is needed. The results presented can contribute to a foundation on high which quality RSP prediction methods can be developed and implemented in Norway.

Bibliography

[1] E. B. Podgoršak and International Atomic Energy Agency, editors. *Radiation oncology physics: a handbook for teachers and students*. International Atomic Energy Agency, Vienna, 2005. OCLC: ocm61877454.

[2] Dominik Maximilian Soliman. *Augmented microscopy: Development and application of high-resolution optoacoustic and multimodal imaging techniques for label-free biological observation*. PhD thesis, Technische Universität München, 2017.

[3] Perry Sprawls. *Physical principles of medical imaging*. Aspen Publishers Gaithersburg, 1993.

[4] Sarah Abdulla. Productrion of x-rays x-ray spectrum. `https://www.radiologycafe.com/radiology-trainees/frcr-physics-notes/production-of-x-rays`, 2020.

[5] Robert Pelberg. Basic principles in computed tomography (ct). In *Cardiac CT Angiography Manual*, pages 19–58. Springer, 2015.

[6] Thorsten RC Johnson. Dual-energy ct: general principles. *American Journal of Roentgenology*, 199(5_supplement):S3–S8, 2012.

[7] Isabel P Almeida, Lotte EJR Schyns, Michel C Öllers, Wouter van Elmpt, Katia Parodi, Guillaume Landry, and Frank Verhaegen. Dual-energy ct quantitative imaging: a comparison study between twin-beam and dual-source ct scanners. *Medical physics*, 44(1):171–179, 2017.

[8] Wayne D Newhauser and Rui Zhang. The physics of proton therapy. *Physics in Medicine & Biology*, 60(8):R155, 2015.

[9] Camilla H Stokkevåg, Grete-May Engeseth, Kristian S Ytre-Hauge, Dieter Röhrich, Odd Harald Odland, Ludvig P Muren, Marianne Brydøy, Liv B Hysing, Artur Szostak, Matthew B Palmer, et al. Estimated risk of radiation-induced cancer following paediatric cranio-spinal irradiation with electron, photon and proton therapy. *Acta Oncologica*, 53(8):1048–1057, 2014.

[10] Kyoto kagaku - crafting soultions in medical education. `https://www.kyotokagaku.com/en/products_data/ph-5/`, 2020.

[11] Christian Möhler, Patrick Wohlfahrt, Christian Richter, and Steffen Greilich. Methodological accuracy of image-based electron density assessment using dual-energy computed tomography. *Medical Physics*, 44(6):2429–2437, 2017.

[12] World health organization - causes of death. https://www.who.int/data/gho/data/themes/topics/causes-of-death, 2018.

[13] Cancer in norway 2018 – cancer incidence, mortality, survival and prevalence in norway. 2018.

[14] Kreftregisteret - nøkkeltall om kreft. https://www.kreftregisteret.no/Temasider/om-kreft/, 2020.

[15] Andrew Brown and Herman Suit. The centenary of the discovery of the bragg peak. *Radiotherapy and Oncology*, 73(3):265–268, 2004.

[16] Particle therapy co-operative group - facilities in operation. https://www.ptcog.ch/index.php/facilities-in-operation, 2020.

[17] Harald Paganetti. Range uncertainties in proton therapy and the role of monte carlo simulations. *Physics in Medicine & Biology*, 57(11):R99, 2012.

[18] Patrick Wohlfahrt and Christian Richter. Status and innovations in pre-treatment ct imaging for proton therapy. *The British Journal of Radiology*, 93(1107):20190590, 2020.

[19] Vicki T Taasti, Christian Bäumer, Christina V Dahlgren, Amanda J Deisher, Malte Ellerbrock, Jeffrey Free, Joanna Gora, Anna Kozera, Antony J Lomax, Ludovic De Marzi, et al. Inter-centre variability of ct-based stopping-power prediction in particle therapy: Survey-based evaluation. *Physics and imaging in radiation oncology*, 6:25–30, 2018.

[20] Wouter van Elmpt, Guillaume Landry, Marco Das, and Frank Verhaegen. Dual energy ct in radiotherapy: current applications and future outlook. *Radiotherapy and Oncology*, 119(1):137–144, 2016.

[21] Patrick Wohlfahrt, Christian Möhler, Christian Richter, and Steffen Greilich. Evaluation of stopping-power prediction by dual-and single-energy computed tomography in an anthropomorphic ground-truth phantom. *International Journal of Radiation Oncology* Biology* Physics*, 100(1):244–253, 2018.

[22] Measurements. Scientific Committee 46-3 on ALARA for Occupationally-Exposed Individuals in Clinical Radiology and Measurements. Scientific Committee 46 on Operational Radiation Safety. *Implementation of the principle of as low as reasonably achievable (ALARA) for medical and dental personnel: recommendations of the National Council on Radiation Protection and Measurements*. Number 107. Ncrp, 1990.

[23] Camilla H Stokkevåg, Daniel J Indelicato, Klaus Herfarth, Henriette Magelssen, Morten E Evensen, Maren Ugland, Terje Nordberg, Tove A Nystad, Camilla Hægeland, Mirjam D Alsaker, et al. Normal tissue complication probability models in plan evaluation of children with brain tumors referred to proton therapy. *Acta Oncologica*, 58(10):1416–1422, 2019.

[24] Joel S Bedford. Sublethal damage, potentially lethal damage, and chromosomal aberrations in mammalian cells exposed to ionizing radiations. *International Journal of Radiation Oncology* Biology* Physics*, 21(6):1457–1469, 1991.

[25] Faiz M. Khan and John P. Gibbons. *Khan's the physics of radiation therapy.* Lippincott Williams & Wilkins/Wolters Kluwer, Philadelphia, PA, fifth edition edition, 2014.

[26] Carles Gomà, Isabel P Almeida, and Frank Verhaegen. Revisiting the single-energy ct calibration for proton therapy treatment planning: a critical look at the stoichiometric method. *Physics in Medicine & Biology*, 63(23):235011, 2018.

[27] Christian Möhler, Patrick Wohlfahrt, Christian Richter, and Steffen Greilich. Range prediction for tissue mixtures based on dual-energy ct. *Physics in Medicine & Biology*, 61(11):N268, 2016.

[28] RA Rutherford, BR Pullan, and I Isherwood. Measurement of effective atomic number and electron density using an emi scanner. *Neuroradiology*, 11(1):15–21, 1976.

[29] Jiang Hsieh. *Computed tomography: principles, design, artifacts, and recent advances*, volume 114. SPIE press, 2003.

[30] Martin J Willemink, Mats Persson, Amir Pourmorteza, Norbert J Pelc, and Dominik Fleischmann. Photon-counting ct: technical principles and clinical prospects. *Radiology*, 289(2):293–312, 2018.

[31] Vicky Trier Taasti. *Improving proton range determination using new x-ray computed tomography principles.* PhD thesis, Aarhus University, 2017.

[32] Stefan Kuchenbecker, Sebastian Faby, Stefan Sawall, Michael Lell, and Marc Kachelrieß. Dual energy ct: How well can pseudo-monochromatic imaging reduce metal artifacts? *Medical physics*, 42(2):1023–1036, 2015.

[33] Gregory Michalak, Joshua Grimes, Joel Fletcher, Ahmed Halaweish, Lifeng Yu, Shuai Leng, and Cynthia McCollough. Improved ct number stability across patient size using dual-energy ct virtual monoenergetic imaging. *Medical physics*, 43(1):513–517, 2016.

[34] MJ Berger, M Inokuti, HH Andersen, H Bichsel, D Powers, SM Seltzer, D Thwaites, and DE Watt. Report 49. *Journal of the International Commission on Radiation Units and Measurements*, (2):NP–NP, 1993.

[35] J F Ziegler. The Stopping of Energetic Light Ions in Elemental Matter. *J. Appl. Phys*, page 51, 1999.

[36] Uli Weber and Gerhard Kraft. Comparison of carbon ions versus protons. *The Cancer Journal*, 15(4):325–332, 2009.

[37] Arabinda Kumar Rath and Narayan Sahoo. *Particle radiotherapy: Emerging technology for treatment of cancer.* Springer, 2016.

[38] Martin J Berger and Stephen M Seltzer. Stopping powers and ranges of electrons and positions. Technical report, NATIONAL STANDARD REFERENCE DATA SYSTEM, 1982.

[39] Esther Bär, Pedro Andreo, Arthur Lalonde, Gary Royle, and Hugo Bouchard. Optimized i-values for use with the bragg additivity rule and their impact on proton stopping power and range uncertainty. *Physics in Medicine & Biology*, 63(16):165007, 2018.

[40] Helge Egil Seime Pettersen. *A digital tracking calorimeter for proton computed tomography*. PhD thesis, The University of Bergen, 2018.

[41] Daphne F Jackson and David J Hawkes. X-ray attenuation coefficients of elements and mixtures. *Physics Reports*, 70(3):169–233, 1981.

[42] Uwe Schneider, Eros Pedroni, and Antony Lomax. The calibration of ct hounsfield units for radiotherapy treatment planning. *Physics in Medicine & Biology*, 41(1):111, 1996.

[43] M. J. Berger, M. Inokuti, H. H. Anderson, H. Bichsel, J. A. Dennis, D. Powers, S. M. Seltzer, and J. E. Turner. Report 37. *Journal of the International Commission on Radiation Units and Measurements*, os19(2):NP–NP, 04 2016.

[44] Christian Möhler. *Stopping-power prediction with dual-energy computed tomography*. PhD thesis, 2018.

[45] Yunhe Xie, Christopher Ainsley, Lingshu Yin, Wei Zou, James McDonough, Timothy D Solberg, Alexander Lin, and Boon-Keng Kevin Teo. Ex vivo validation of a stoichiometric dual energy ct proton stopping power ratio calibration. *Physics in Medicine & Biology*, 63(5):055016, 2018.

[46] El Hassane Bentefour, Stefan Both, Shikui Tang, and Hsiao-Ming Lu. Using cbct for pretreatment range check in proton therapy: a phantom study for prostate treatment by anterior–posterior beam. *Journal of applied clinical medical physics*, 16(6):472–483, 2015.

[47] Piecewise linear fit - python package contents. `https://jekel.me/piecewise_linear_fit_py/pwlf.html`, 2020.

[48] National institute of standards and technology - xcom: Photon cross sections database. `https://www.nist.gov/pml/xcom-photon-cross-sections-database`, 2019.

[49] William Henry Bragg and Richard Kleeman. Xxxix. on the α particles of radium, and their loss of range in passing through various atoms and molecules. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 10(57):318–340, 1905.

[50] Christian Möhler, Tom Russ, Patrick Wohlfahrt, Alina Elter, Armin Runz, Christian Richter, and Steffen Greilich. Experimental verification of stopping-power prediction

from single-and dual-energy computed tomography in biological tissues. *Physics in Medicine & Biology*, 63(2):025001, 2018.

[51] Ming Yang, X Ronald Zhu, Peter C Park, Uwe Titt, Radhe Mohan, Gary Virshup, James E Clayton, and Lei Dong. Comprehensive analysis of proton range uncertainties related to patient stopping-power-ratio estimation using the stoichiometric calibration. *Physics in Medicine & Biology*, 57(13):4095, 2012.

[52] Nils Krah, Vincenzo Patera, Simon Rit, Angelo Schiavi, and Ilaria Rinaldi. Regularised patient-specific stopping power calibration for proton therapy planning based on proton radiographic images. *Physics in Medicine & Biology*, 64(6):065008, 2019.

[53] Patrick Wohlfahrt, Christian Möhler, Wolfgang Enghardt, Mechthild Krause, Daniela Kunath, Stefan Menkel, Esther GC Troost, Steffen Greilich, and Christian Richter. Refinement of the hounsfield look-up table by retrospective application of patient-specific direct proton stopping-power prediction from dual-energy ct. *Medical Physics*, 47(4):1796–1806, 2020.

[54] Report 46 ICRU Report 46, editor. *Photon, electron, proton, and neutron interaction data for body tissues*. Number 46 in ICRU report. International Commission on Radiation Units and Measurements, Bethesda, Md., U.S.A, 1992.

Tabulated Material A

A.1 Gammex 467 insert constituents

The elemental constituents of the tissue like inserts used for scanner calibration with the Gammex 467 calibration phantom can be seen in Table A.1:

Table A.1: Tabulated elemental composition (in percent) of Gammex 467 inserts, relative electron density, $\hat{\rho}_e$, and relative mass density, $\hat{\rho}$, provided by manufacturer.

Element	**H**	**C**	**N**	**O**	**Mg**	**Si**	**P**	**Cl**	**Ca**	$\hat{\rho}_e$	$\hat{\rho}$
Z	**1**	**6**	**7**	**8**	**12**	**14**	**15**	**17**	**20**		
A	**1.008**	**12.01**	**14.01**	**16**	**24.31**	**28.09**	**30.97**	**35.45**	**40.08**		
CB2-30%	6.68	53.48	2.12	25.61	0	0	0	0.11	12.01	1.279	1.334
Lung 300	8.46	59.38	1.96	18.14	11.19	0.78	0	0.1	0	0.281	0.290
Inner Bone	6.67	55.64	1.96	23.52	0	0	3.23	0.11	8.87	1.086	1.133
Solid Water	8.02	67.23	2.41	19.91	0	0	0	0.14	2.31	0.988	1.017
Cort Bone	3.41	31.41	1.84	36.5	0	0	0	0.04	26.81	1.696	1.824
Liver	8.06	67.01	2.47	20.01	0	0	0	0.14	2.31	1.063	1.095
Brain	10.83	72.54	1.69	14.86	0	0	0	0.08	0	1.047	1.051
Lung 450	8.47	59.57	1.97	18.11	11.21	0.58	0	0.1	0	0.458	0.470
Solid Water	8.02	67.23	2.41	19.91	0	0	0	0.14	2.31	0.988	1.018
Adipose	9.06	72.3	2.25	16.27	0	0	0	0.13	0	0.928	0.945
Solid Water	8.02	67.23	2.41	19.91	0	0	0	0.14	2.31	0.988	1.017
B-200	6.65	55.52	1.98	23.64	0	0	3.24	0.11	8.87	1.102	1.150
Muscle	8.1	67.17	2.42	19.85	0	0	0	0.14	2.32	1.020	1.050
CB2-50%	4.77	41.63	1.52	32	0	0	0	0.08	20.02	1.471	1.561
Water	11.19	0	0	88.81	0	0	0	0	0	1.000	1.000
Breast	8.59	70.11	2.25	16.27	0	0	0	0.13	0.95	0.954	0.977

A.2 I-values

Table A.2: Mean ionization energies [eV] for the elements in a solid state from Berger *et al* 1984.

Element	**H**	**C**	**N**	**O**	**Mg**	**Si**	**P**	**Cl**	**Ca**
I (eV)	19.02	81	82	106	176.3	195.5	195.5	180	215.8

A.3 ICRU Standard Tissues

The following table (Table A.3) contains elemental composition data on 61 selected adult tissues obtained from [54]. These are used as guidance for creating human tissue-like substitutes used in radiotherapy. The same data is usually used to create the clinical HLUT, by applying the CT calibration to the standard tissues as well as scanned tissue substitutes used for calibration.

Table A.3: 61 adult standard tissues listed in ICRU Report 46.
Elemental compositions, mass density, ρ $[kgm^{-3}]$, electron density ρ_e $[m^{-3} \cdot 10^{26}]$, and electron density relative to water $\hat{\rho}_e$

Tissue–	**H**	**C**	**N**	**O**	**Na**	**Mg**	**P**	**S**	**Cl**	**K**	**Ca**	**Fe**	**I**	ρ	ρ_e	$\hat{\rho}_e$
Z	**1**	**6**	**7**	**8**	**11**	**12**	**15**	**16**	**17**	**19**	**20**	**26**	**53**			
A	**1.008**	**12.01**	**14.01**	**16**	**22.99**	**24.31**	**30.97**	**32.05**	**35.45**	**39.1**	**40.08**	**55.85**	**126.9**			
Adipose #1	11.2	51.7	1.3	35.5	0.1	0	0	0.1	0.1	0	0	0	0	970	3240	0.970
Adipose #2	11.4	59.8	0.7	27.8	0.1	0	0	0.1	0.1	0	0	0	0	950	3180	0.952
Adipose #3	11.6	68.1	0.2	19.8	0.1	0	0	0.1	0.1	0	0	0	0	930	3120	0.934
Lipoma	10.9	71.9	3	13.8	0.1	0	0	0.2	0.1	0	0	0	0	980	3270	0.979
ST ICRU-33	0.1	11.1	2.6	76.2	0	0	0	0	0	0	0	0	0	1000	3310	0.991
ST ICRU-44m	10.5	25.6	2.7	60.2	0.1	0	0.2	0.3	0.2	0.2	0	0	0	1030	3420	1.024
ST ICRU-44f	10.6	31.5	2.4	54.7	0.1	0	0.2	0.2	0.1	0.2	0	0	0	1020	3390	1.015
Blood—	10.2	11	3.3	74.5	0.1	0	0.1	0.2	0.3	0.2	0	0.1	0	1060	3510	1.051
Bone mineral	0.2	0	0	41.4	0	0	18.5	0	0	0	39.9	0	0	3225	9670	2.895
Brain—	10.7	14.5	2.2	71.2	0.2	0	0.4	0.2	0.3	0.3	0	0	0	1040	3460	1.036
Breast-calc	0.3	1.6	0.5	40.7	0	0	18.7	0	0	0	38.2	0	0	3060	9180	2.749
Breast-mam-1	10.9	50.6	2.3	35.8	0.1	0	0.1	0.1	0.1	0	0	0	0	990	3300	0.988
Breast-mam-2	10.6	33.2	3	52.7	0.1	0	0.1	0.2	0.1	0	0	0	0	1020	3390	1.015
Breast-mam-3	10.2	15.8	3.7	69.8	0.1	0	0.1	0.2	0.1	0	0	0	0	1060	3510	1.051
Breast 50/50	11.5	38.7	0	49.8	0	0	0	0	0	0	0	0	0	960	3220	0.964
Breast 33/67	11.6	51.9	0	36.5	0	0	0	0	0	0	0	0	0	940	3150	0.943
Carbohydrate	6.2	44.5	0	49.3	0	0	0	0	0	0	0	0	0	1560	4980	1.491
Cell nucleus	10.6	9	3.2	74.2	0	0	2.6	0.4	0	0	0	0	0	1000	3320	0.994
Eye lens	9.6	19.5	5.7	64.6	0.1	0	0.1	0.3	0.1	0	0	0	0	1070	3530	1.057
Gallstones	11.8	79.4	0.5	8.3	0	0	0	0	0	0	0	0	0	1075	3610	1.081
GI tract	10.6	11.5	2.2	75.1	0.1	0	0.1	0.1	0.2	0.1	0	0	0	1030	3420	1.024
Heart healty	10.4	13.9	2.9	71.8	0.1	0	0.2	0.2	0.2	0.3	0	0	0	1050	3480	1.042
Heart fatty	10.3	18.2	3.1	67.4	0.1	0	0.2	0.2	0.2	0.3	0	0	0	1040	3450	1.033
Heart filled	10.3	12.1	3.2	73.4	0.1	0	0.1	0.2	0.3	0.2	0	0.1	0	1060	3510	1.051
Kidney–	10.3	13.2	3	72.4	0.2	0	0.2	0.2	0.2	0.2	0.1	0	0	1050	3480	1.042
Lipid—	11.8	77.3	0	10.9	0	0	0	0	0	0	0	0	0	920	3090	0.925
Liver healthy	10.2	13.9	3	71.6	0.2	0	0.3	0.3	0.2	0.3	0	0	0	1060	3510	1.051
Liver cirrhotic	10.4	23	2.4	63	0.2	0	0.3	0.2	0.2	0.3	0	0	0	1040	3450	1.033
Liver fatty	10.3	18.6	2.8	67.1	0.2	0	0.2	0.3	0.2	0.3	0	0	0	1050	3480	1.042
Lung healty	10.3	10.5	3.1	74.9	0.2	0	0.2	0.3	0.3	0.2	0	0	0	260	862	0.258
Lung congested	10.5	8.3	2.3	77.9	0.2	0	0.1	0.2	0.3	0.2	0	0	0	1040	3450	1.033
Lymp—-	10.8	4.1	1.1	83.2	0.3	0	0	0.1	0.4	0	0	0	0	1030	3420	1.024

Table A.3 continued from previous page

Tissue–	**H**	**C**	**N**	**O**	**Na**	**Mg**	**P**	**S**	**Cl**	**K**	**Ca**	**Fe**	**I**	ρ	ρ_e	$\rho_{e,w}$
Muscle–	10.2	14.3	3.4	71	0.1	0	0.2	0.3	0.1	0.4	0	0	0	1050	3480	1.042
Ovary—	10.5	9.3	2.4	76.8	0.2	0	0.2	0.2	0.2	0.2	0	0	0	1050	3490	1.045
Pancreas	10.6	16.9	2.2	69.4	0.2	0	0.2	0.1	0.2	0.2	0	0	0	1040	3460	1.036
Protein-	6.6	53.4	17	22	0	0	0	1	0	0	0	0	0	1350	4330	1.296
Skeleton-cartil	9.6	9.9	2.2	74.4	0.5	0	2.2	0.9	0.3	0	0	0	0	1100	3620	1.084
Skeleton-cort.b	3.4	15.5	4.2	43.5	0.1	0.2	10.3	0.3	0	0	22.5	0	0	1920	5950	1.781
Skeleton crani	5	21.2	4	43.5	0.1	0.2	8.1	0.3	0	0	17.6	0	0	1610	5070	1.518
Femur - 30y	7	34.5	2.8	36.8	0.1	0.1	5.5	0.2	0.1	0	12.9	0	0	1330	4270	1.278
Femur - 90y	7.9	38.5	2.2	36.2	0.1	0.1	4.4	0.1	0.1	0	10.4	0	0	1220	3950	1.183
Humerus-	6	31.4	3.1	36.9	0.1	0.1	7	0.2	0	0	15.2	0	0	1460	4640	1.389
Mandible	4.6	19.9	4.1	43.5	0.1	0.2	8.6	0.3	0	0	18.7	0	0	1680	5270	1.578
Red Marrow	10.5	41.4	3.4	43.9	0	0	0.1	0.2	0.2	0.2	0	0.1	0	1030	3420	1.024
Ribs1—	6.4	26.3	3.9	43.6	0.1	0.1	6	0.3	0.1	0.1	13.1	0	0	1410	4500	1.347
Ribs2—	5.6	23.5	4	43.4	0.1	0.1	7.2	0.3	0.1	0.1	15.6	0	0	1520	4820	1.443
Sacrum M	7.4	30.2	3.7	43.8	0	0.1	4.5	0.2	0.1	0.1	9.8	0.1	0	1290	4160	1.246
Sacrum F	6.6	27.1	3.8	43.5	0.1	0.1	5.8	0.3	0.1	0.1	12.5	0	0	1390	4450	1.332
Spongiosa	8.5	40.4	2.8	36.7	0.1	0.1	3.4	0.2	0.2	0.1	7.4	0.1	0	1180	3840	1.150
Vert.col. C4	6.3	26.1	3.9	43.6	0.1	0.1	6.1	0.3	0.1	0.1	13.3	0	0	1420	4530	1.356
Vert.col. D6,L3	7	28.7	3.8	43.7	0	0.1	5.1	0.2	0.1	0.1	11.1	0.1	0	1330	4270	1.278
Yellow marrow	11.5	64.4	0.7	23.1	0.1	0.1	0	0	0.1	0	0	0	0	980	3280	0.982
Skin—-	10	20.4	4.2	64.5	0.2	0	0.1	0.2	0.3	0.1	0	0	0	1090	3600	1.078
Spleen–	10.3	11.3	3.2	74.1	0.1	0.3	0	0.2	0.2	0.3	0	0	0	1060	3510	1.051
Testis–	10.6	9.9	2	76.6	0.2	0	0.1	0.2	0.2	0.2	0	0	0	1040	3460	1.036
Thyroid-	10.4	11.9	2.4	74.5	0.2	0	0.1	0.1	0.2	0.1	0	0	0.1	1050	3480	1.042
Bladder empty	10.5	9.6	2.6	76.1	0.2	0	0.2	0.2	0.3	0.3	0	0	0	1040	3450	1.033
Bladder filled	10.8	3.5	1.5	83	0.3	0	0.1	0.1	0.5	0.2	0	0	0	1030	3430	1.027
Ur. stones cyst	5	30.7	11.9	26.5	0	0	0	25.98	0	0	0	0	0	1660	5240	1.569
Ur. stones oxa	1.5	11.7	0.5	50.7	0	0	6	0	0	0	29.6	0	0	2300	7010	2.099
Ur. stones acid	3.3	33.2	29.9	33.6	0	0	0	0	0	0	0	0	0	1745	5420	1.623
Water—	11.2	0	0	88.8	0	0	0	0	0	0	0	0	0	1000	3340	1.000

Python codes

B

B.1 Determine k-values

```
### Gradient decent minimization to determine k1 and k2

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy.optimize import minimize
from scipy.stats import linregress

data       = pd.read_csv("Data/datablad_gammex_NY.csv")
measured   = pd.read_csv("Data/Measured_Mean_HU.csv")


### HU_R = mu/mu_w = rho_rel * (f(k1,k2))/"water"
def objective(k):
    k1 = k[0]
    k2 = k[1]

    x_ = list()
    y_ = list()

    ## Denominator in eq. A.20 (Goma et al 2016) -water
    wH = float(data.iloc[16,1])
    AH = float(data.iloc[1,1])
    wO = float(data.iloc[16,4])
    AO = float(data.iloc[1,4])
    sum_water = wH/AH*(1+k1+k2) + wO/AO*(8+k1*8**2.86+k2*8**4.62)
    ##


    ## f(k1,k2)
    sum_tot=0
    ### Loop over Gammex tabulated material spreadsheet (Table A.1)
    for i in range(2,18):# itterate through rows (CB2-30% - Breast)
        sum_row = 0
```

```
        for j in range (1,10): # itterate through columns (H - Ca)
            wi  = float(data.iloc[i,j])
            Ai  = float(data.iloc[1,j])
            Zi  = float(data.iloc[0,j])
            rho = float(data.iloc[i,11]) #relative mass density

            sum_element = wi/Ai * (Zi + k1*Zi**2.86 + k2*Zi**4.62)
            sum_row += sum_element

        calc_HU = rho*(sum_row/sum_water)
        meas_HU = (float(measured.iloc[i-2,0]))/1000+1
    ##

        if kDRAW:
            x_.append(meas_HU)
            y_.append(calc_HU)

        diff_squared = (calc_HU - meas_HU)**2

        sum_tot += diff_squared

    if kDRAW:
        fig = plt.figure(figsize=(10,5))
        plt.subplot(121)
        plt.scatter(x_, y_, label="data")
        xlin = np.linspace(min(x_), max(x_), 100)
        plt.xlabel("Measured HU")
        plt.ylabel("Calculated HU")
        plt.ylim([0, 2.5])
        plt.title(f"k1={k1:.3e};k2={k2:.3e};error={sum_tot:.3e}")
        plt.plot(xlin, xlin, "r", label=f"x=y")
        plt.legend()

        plt.subplot(122)
        plt.scatter(x_, [a-b for a,b in zip(x_,y_)],
                    label="calculated - measured")
        xlin = np.linspace(min(x_), max(x_), 100)
        plt.xlabel("Measured HU")
        plt.ylabel("Calculated - measured HU")
        plt.ylim([-0.1, 0.1])
        plt.show()

    return sum_tot

### Determine k-values by minimizing the HU_meas - HU_synth difference
# Initial gues-values
```

```
k0 = [5.3*10**(-4),2.3*10**(-5)]
kDRAW = True
print(objective(k0))

kDRAW = False
# Bounds (Goma et al, p. 5)
b1 = (2*10**(-4),6*10**(-3))
b2 = (3*10**(-6),6*10**(-4))
bnds = (b1,b2)

# Solve:
sol = minimize(objective, k0, method='SLSQP', bounds=bnds)
print(sol)

kDRAW = True
print(objective(sol.x))
```

B.2 Determine alpha

```
### Calibration of alpha using bone inserts ###

import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import curve_fit

# Bone inserts -> ((HU/1000)+1)/rel_e_dens) -> relative cross section
x = np.array([1.239, 1.181, 1.534, 1.183, 1.392]) #rcs_low
y = np.array([1.079, 1.061, 1.177, 1.061, 1.128]) #rcs_high


def f(x,a):
    return 1/a + (1-1/a)*x

plt.plot(x, y, "x", label="Bone Inserts")


popt, pcov = curve_fit(f,x,y)

plt.plot(x, f(x, *popt), 'r-',
         label='fit: a=%5.4f' % tuple(popt))

plt.xlabel('Relative Cross Section (Low E)')
plt.ylabel('Relative Cross Section (High E)')
plt.legend()
plt.show()
```