Photon-Counting CT for proton therapy

SPR Prediction Comparison with SECT and DECT and Validation with Proton Beam Measurement

BM51035: BME MSc Thesis

Ries van Walsum

Delft University of Technology

TU Delft

Photon-Counting CT for proton therapy

SPR Prediction Comparison with SECT and DECT and Validation with Proton Beam Measurement

by

Ries van Walsum

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday July 9, 2024 at 14:30 PM.

Student number:	4888898	
Project duration:	December, 2023 – July, 2024	
Thesis committee:	Dr. M.C. Goorden	TU Delft, supervisor
	S.C. Huijskens	HollandPTC, Erasmus MC, daily supervisor
	A.G. Denkova	TU Delft
Faculty:	Faculty of Mechanical Engineering, Delft	

This thesis is confidential and cannot be made public until July 1, 2025.

TU Delft

Voorwoord

Voor u ligt het eindproduct van mijn afstudeeronderzoek van afgelopen half jaar. Ik heb onderzoek gedaan naar het gebruik van photon-counting CT voor protonentherapie bij het HollandPTC, in samenwerking met het Erasmus MC. Dit onderzoek heb ik uitgevoerd als student Biomedical Engineering - Medical Physics aan de TU Delft.

Ik wil mijn dagelijkse begeleider Sophie Huijskens bedanken voor het sparren over het onderzoek en de goede feedback die ik heb ontvangen op mijn werk. Dit is het academisch gehalte van mijn thesis absoluut ten goede gekomen. Daarnaast wil ik mijn begeleider vanuit de TU Delft, Marlies Goorden, bedanken voor haar feedback en het meedenken tijdens dit gehele traject. Tot slot wil ik ook Antonia Denkova bedanken, die het derde lid is van mijn afstudeercommissie.

Verder wil ik Marcel van Straten bedanken voor het meedenken en meehelpen met het maken van de photon-counting CT scans in het Erasmus MC. Ik ben ook geholpen door Patricia Cambraia Lopes bij het maken van de single- en dual-energy CT scans bij het HollandPTC. Ook wil ik Ernst van de Wal en Thomas Toet bedanken voor hun hulp met de metingen met de protonenbundel bij het HollandPTC en Ernst voor de hulp met het 3D ontwerpen en printen van de bakjes.

Naast deze academische steun, ben ik mijn studie- en huisgenoten dankbaar voor hun adviezen afgelopen jaar, maar bovenal voor de vele gezamenlijke koffie- en lunchsessies. Specifiek Thijs, wiens printtegoed ik langzaam heb opgemaakt, die mij tot het einde heeft bijgestaan in C4 en nog een paar maandjes doormoet. Mijn ouders wil ik ook bedanken voor hun inhoudelijke en financiële steun tijdens mijn studie. Tot slot wil ik Marit bedanken voor haar onvoorwaardelijke steun tijdens mijn studie, ook als ik soms weekenden en avonden moest doorstuderen.

Ries van Walsum
Delft, July 2024

Samenvatting

Beeldvorming met Computertomografie (CT) is een belangrijke stap in de behandelplanning voor protontherapie. Het omrekenen van Hounsfield Units (HU) naar de 'stopping power ratio' (SPR) veroorzaakt meer dan de helft van de 3,5% foutmarge waar in de behandelplanning rekening mee wordt gehouden. Het doel van deze studie is om de effectiviteit van 'photon counting CT' (PCCT) te onderzoeken voor het voorspellen van de SPR bij protontherapie. De prestaties van PCCT worden vergeleken met die van 'single-energy CT' (SECT) en 'dual-energy CT' (DECT).

Vers dierlijk weefsel – biefstuk, gehakt en beenmerg – is gefixeerd in 3D-geprinte bakjes om luchtinsluiting te minimaliseren en de stabiliteit van het weefsel tijdens de metingen te waarborgen. Deze bakjes zijn geplaatst in een hoofd- en een lichaamsconfiguratie van een fantoom, waarvan CT-scans zijn gemaakt met standaard klinische protocollen. De SPR-waarden zijn afgeleid van de CT-gegevens met behulp van geautomatiseerde DirectSPR-software voor de DECT- en PCCT-scans, en omgezet van HU's naar SPR met behulp van een Hounsfield look-up tabel (HLUT) voor de SECT-scans. De resultaten zijn gevalideerd tegen gouden standaard SPR-metingen, gemeten in een opstelling met een protonenbundel.

Er zijn twee meetreeksen uitgevoerd. Voor de validatie van de eerste meetreeks resulteerden energieën van 150 en 175 MeV voor de protonenbundel beide in een SPR-waarde van 1,00 voor het gehakt. Voor de biefstuk werden SPR-waarden van 1,02 en 1,03 gevonden. De op CT gebaseerde SPR-voorspellingen resulteerden in fouten tot 1,5% voor gehakt en 3,5% voor biefstuk. Voor de validatie van de tweede meetreeks werden SPR-waarden voor gehakt (0.89-0.92), biefstuk (0.96-0.98) en beenmerg (0.87-0.91) verkregen. Het procentuele verschil tussen de op CT gebaseerde SPR-voorspellingen en de validatiemetingen toonde geen substantiële verschillen tussen de CT-modaliteiten, behalve dat SECT lagere SPR's voor de beenmergmonsters opleverde dan DECT en PCCT. De validatiemetingen van de tweede serie resulteerden in lagere SPR-waarden dan die in de literatuur en de eerste serie, wat wijst op mogelijke methodologische inconsistenties. Factoren zoals heterogeniteit in het vlees en meetfouten kunnen hebben bijgedragen aan deze inconsistenties, maar verklaren ze niet volledig. Deze resultaten benadrukken de noodzaak voor verder onderzoek met gestandaardiseerde fantomen en dierlijk weefsel om het werkelijke potentieel van PCCT in combinatie met DirectSPR-voorspelling te evalueren.

Uit deze studie kan niet worden geconcludeerd dat PCCT een voordeel biedt ten opzichte van SECT en DECT als CT-modaliteit voor de HU-SPR omrekening in de behandelplanning van protontherapie.

Contents

1

Paper

Is latest always the greatest? The potential of Photon-Counting CT for better SPR predictions

Van Walsum, M.J.
TU Delft, Delft, Netherlands
Dept. of Radiology and Nuclear Medicine, Dept. of Radiotherapy of Erasmus MC, Rotterdam, Netherlands
HollandPTC, Delft, Netherlands
m.j.vanwalsum@student.tudelft.nl

***Abstract*—Computed Tomography (CT) imaging is an important step in treatment planning for proton therapy. The conversion from Hounsfield Unit (HU) to stopping power ratio (SPR) accounts for more than half of the 3.5% error that is considered in treatment planning. The aim of the study is to investigate the efficacy of photon-counting CT (PCCT) for predicting the SPR in proton therapy. Its performance will be compared to single-energy CT (SECT) and dual-energy CT (DECT).**

Fresh tissue samples—steak, ground beef, and bone marrow—were secured in 3D-printed holders to minimize air inclusion and maintain tissue stability during measurements. These samples were secured in a head and a body configuration of a phantom and CT scans were conducted using standard clinical protocols. SPR values were derived from the CT data using automated DirectSPR software for the DECT and PCCT scans, and converted from HUs to SPR using a Hounsfield look up table (HLUT) for the SECT scans. Results were validated against actual SPR measurements obtained via proton beam measurements.

Two measurement series have been performed. For the validation of the first measurement series, proton beam energies of 150 and 175 MeV both resulted in an SPR value of 1.00 for the ground beef sample. For the steak sample, SPR values of 1.02 and 1.03 were found. The CT-based SPR predictions resulted in errors up to 1.5% for ground beef and 3.5% for steak. For the validation of the second measurement series, SPR values for ground beef (0.89-0.92), steak (0.96-0.98), and bone marrow (0.87-0.91) were obtained. The percentage error between the CT-based SPR predictions and the validation measurements showed no substantial differences between the CT-modalities, except for SECT providing lower SPRs for the bone marrow samples than DECT and PCCT. The validation measurements of the second series resulted in lower SPR values than found in literature and the first series, indicating possible methodological inconsistencies. Factors such as sample heterogeneities and measurement errors might have contributed to these inconsistencies, but do not account for these inconsistencies completely. These results highlight the need for further research with standardized phantoms and animal tissue samples to evaluate PCCT's true potential in combination with DirectSPR prediction.

From this study it cannot be concluded that PCCT demonstrates an advantage over SECT and DECT as CT-modality for the HU-SPR conversion in proton therapy treatment planning.

***Index Terms*—Proton therapy, stopping power ratio, CT, photon-counting CT**

I. Introduction

A. Background

Radiotherapy is a common treatment method for cancer [1]. It aims to kill the tumor with ionizing radiation, while sparing healthy tissues as much as possible. Photon therapy, the most common form of radiotherapy, exhibits a dose distribution where the highest dose deposition occurs at the beam's entry point. The dose deposition decreases as the beam penetrates deeper into the tissue. In contrast, proton therapy has a distinct dose distribution characterized by a 'Bragg peak' at a specific depth, where dose deposition is maximal before rapidly diminishing to near zero (Fig. 1). This allows for precise targeting of tumors, while minimizing the dose deposition in surrounding organs at risk (OARs) and reducing long-term side effects [2]. To exploit this benefit, the stopping power ratios (SPRs) of all tissues in the beam path need to be known. This is crucial because it quantifies how much energy protons will deposit as they pass through different tissues. This enables precise positioning of the Bragg peak, which in turn allows for targeting tumors effectively while sparing healthy tissues. These SPRs are determined from Computed Tomography (CT) scans that are made before treatment.

B. Current issues with SPR prediction

1) Single-energy CT: CT is a medical imaging technique that uses X-ray beams to create cross-sectional images of the body. Each voxel is assigned a Hounsfield Unit (HU), which quantifies the relative density of tissues by comparing their photon attenuation to that of water and air. This photon attenuation in CT imaging is described by the equation

$$I = I_0 e^{-\mu x}\,, \tag{1}$$

where I_0 is the initial intensity of the beam, I is the intensity of the beam after passing through the body, μ is the linear attenuation coefficient and x is the path length of the beam through the material. The beam traverses several materials in the body that all have a certain μ and x, which add up to the total attenuation in the body. The intensity of the beam is defined by the set peak tube voltage (kVp), current and exposure time. The peak tube voltage is the highest energy that the photons can have, however the distribution of the energy of the photons can be considered a spectrum with the set kVp

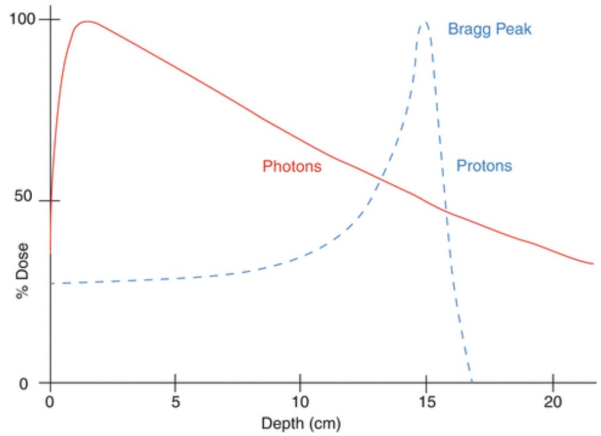


Fig. 1: Depth dose distribution of photons and protons. From [5]

as highest energy. The μ is dependent on the energy that the photons have and will change when irradiated with photons of different energies.

It is important to highlight that there is a direct relation between HUs and the attenuation coefficients that are used in photon therapy, since both are a measure for how much photons passing through tissue will attenuate. This is a fundamental difference with proton therapy, where HUs relate to the SPR in a more complex way. Proton interactions in the body are expressed by the Bethe-Bloch equation, as reported in Schneider et al.[3], which describes the energy loss of protons when traveling through a medium. This equation is used to calculate the stopping power of a certain medium for protons. The SPR is the stopping power of a certain material, relative to that of water. It is showed that the SPR for a certain material is mainly dependent on the relative electron density (RED) and the excitation energy (I-value) [4]. The I-value is the energy that is required to ionize the atoms in the medium. However, photon interactions are also dependent on RED, since there is an almost linear relation between the attenuation coefficient in CT imaging and the RED. This mutual dependency on the RED is defined in a Hounsfield look-up table (HLUT), which is the current clinical standard for predicting the SPR with single energy CT (SECT). This HLUT, used to relate HUs to SPR, consists of several line segments for different tissue types. There are two main methods to define and fit the HLUT: the empirical and the stoichiometric calibration methods [4].

- Empirical method: This involves measuring HUs and SPRs of tissue surrogates, which are materials designed to simulate human tissue properties. However, discrepancies arise because these surrogates do not perfectly mimic the actual human tissues, leading to potential inaccuracies in SPR predictions.
- Stoichiometric method: As proposed by Schneider et al. [3], this method involves direct measurements on human tissues, primarily from cadavers, to calculate the SPR. Though more accurate, this approach is limited by the availability of human tissue datasets and the need for CT scanner-specific calibration, which is often impractical.

As a consequence of the impracticalities related to the stoichiometric method, tissue surrogates are still widely used, necessitating careful selection of calibration phantoms as emphasized by Wohlfahrt et al. [6].

The HLUT is fundamentally based on the RED, while it was observed that the I-value also affects the SPR. This leads to inherent imperfections in HU-SPR correlations. These limitations contribute significantly to uncertainties in range predictions for proton therapy treatment planning. Currently, uncertainty margins of 3-3.5% are used in the clinic, some adding an absolute term of 1-2 mm [7]. Despite advances in the technique, these uncertainty margins have remained largely unchanged since 1985 [8]. One of the long-term goals of proton therapy research, defined by Paganetti et al. [2], is to improve the range prediction accuracy to $\leq 1\%$. The fact that the HU-SPR conversion using a HLUT contributes for the larger part of this error [2], emphasizes the need for improved methods to reduce the error in this conversion. This would lead to reduced dose deposition in the healthy tissues surrounding the tumor and thus may lower the side effects of proton therapy.

2) Inter-center variability: Taasti et al. [7] conducted a study on the variability in proton therapy workflows across different proton therapy centers. Their research involved sending questionnaires to twelve clinics, aiming to explore all critical aspects of the process from CT planning to SPR prediction. The findings revealed significant diversity in scanning parameters and methods for HLUT calibrations. To address these discrepancies and improve standardization, Peters et al. [9] conducted a follow-up study where a 'ground truth' phantom was scanned at seventeen proton therapy centers, each using its own scanning protocols and SPR prediction methods. This study highlighted the inter-center variation in range predictions, reporting standard deviations (2σ) of 2.9% for prostate-, 2.6% for lung-, and 1.3% for primary brain-tumor treatments. These findings showed that both accounting for beam hardening (increased beam energy as beam penetrates denser mediums) and conducting independent HLUT validation reduced the absolute error on average. In response to the observed variability, Peters et al. [10] proposed a step-by-step guide for specifying a HLUT to standardize SPR predictions. Their guide outlines six key steps:

1) Phantom Setup: Properly configuring the phantom to simulate patient tissues.
2) CT Acquisition: Collecting CT images with consistent parameters.
3) CT Number Extraction: Extracting HUs from the CT data.
4) SPR Determination: Calculating the SPR from the CT data.
5) HLUT Specification: Defining the HLUT based on the extracted HUs and SPR values.
6) HLUT Validation: Validating the HLUT with experimental data.

The guide was validated with three different CT scanners, demonstrating its potential to reduce inter-center variability in SPR predictions. The authors also suggested an optional verification method using biological animal tissues to better

simulate patient scenarios, which they proposed as a future focus for research.

C. Spectral CT

1) Dual-energy CT: Dual-Energy CT (DECT) operates on the principle of capturing CT scans at two different tube voltages, providing spectral information that can be used to reveal the composition of scanned tissues. Introduced in 1973, DECT's potential for enhancing charged particle therapy was recognized early, with Goitein [11] in 1977 proposing that scans at 90 and 140 kVp could improve treatment planning. Clinical adoption was not realized until 2015, when Wohlfahrt et al. [12] demonstrated its implementation in practice.

The spectral data from DECT can be processed to create virtual monoenergetic CT images (VMI), where the HUs from low and high voltage scans are combined using a weighting factor (α):

$$HU_{VMI}(\alpha) = \alpha * HU_{low} + (1 - \alpha) * HU_{high}\,. \qquad (2)$$

Optimizing this α can help to minimize beam hardening effects, leading to more consistent HUs, which is crucial for accurate proton therapy planning. This VMI can then be used in combination with a HLUT, partially compensating for the inherent issues of the SECT-based HU-SPR conversion [13].

Since two scans are made, the attenuation for two different kVps is known. This spectral information can also be used to directly determine both photon attenuation based tissue parameters, namely the RED and the effective atomic number [4]. The I-value cannot be directly determined using aforementioned tissue parameters. This means that, just as with the HLUT, an imperfect approximation is used to determine a relevant parameter for the HU-SPR conversion. However, the resulting SPR accuracy of this method is substantially higher than for the SECT-based approach with the HLUT [4]. This spectral information thus allows using the Bethe-Bloch equation to determine the SPR of the different tissues that the proton beam will traverse during treatment. This advancement addresses the inherent issues in the HLUT-SPR conversion process and contributes to more accurate range predictions in proton therapy planning.

The benefit of DECT over SECT for SPR prediction has been proven in literature in manifold (Appendix A) [14–19]. It should be noted that Taasti et al. [7] found that, despite recognizing the potential of DECT to enhance SPR predictions, in 2018 only one center had integrated DECT into their workflow.

An example of this benefit of DECT-based SPR predictions is commercially available DirectSPR (Siemens Healthineers, Erlangen, Germany) software that has been introduced, which can directly predict the SPR, using the spectral information from the two DECT scans. The novelty of this software, and also a great benefit, is that the SPR prediction is already integrated in the CT image processing instead of the treatment planning system. In clinical practice, HLUTs need to be calibrated for different patient sizes due to variations in HUs caused by beam hardening. However, due to the use of DECT scans and processing them with this new software, this extensive calibration is no longer required. The software facilitates patient-specific SPR predictions based on the DECT scans. Additionally, it incorporates advanced image processing algorithms that suppress noise, enhancing the accuracy and quality of the predictions. Besides its potential to simplify the workflow of proton therapy planning, literature [19, 20] has shown significant improvements in SPR prediction accuracy compared with SECT.

2) Photon-counting CT: Conventional CT scanners (SECT or DECT) use energy-integrating detectors (EID), which convert incoming photons into light via a scintillator, and subsequently into an electrical signal using photodiodes. These signals are integrated over a set time interval to form a signal that is converted into an image. This means that all signals are added up, leading to a higher weight for photons with higher energies, which also is the reason that beam hardening is a problem. A photon-counting CT scanner (PCCT), on the other hand, employs a photon-counting detector (PCD), which directly converts incoming photons into electrical pulses using a semiconductor. Pulses exceeding a threshold are recorded as incoming photons, with multiple thresholds allowing for energy binning to acquire spectral information [21]. This spectral information facilitates VMI reconstruction.

However, PCD's also come with already existent and new signal-degrading effects. Photoelectric absorption of incoming photons with K-fluorescence can emit characteristic X-rays that land in a different pixel. Compton scatter can also lead to photons landing in a different pixel than the one in which the primary X-ray interacted. Lastly, charge sharing distributes the energy of a single photon across multiple pixels, leading to photons being registered with incorrect energies and at the wrong pixels. These three effects are regarded as crosstalk, since they lead to photons being misassigned to adjacent pixels [22]. In addition, pulse pile-up occurs when a detector cannot register a second photon arriving shortly after the first due to its recovery time. Although pulse pile-up was a significant challenge during PCD development and can not be completely eliminated, it can be partially mitigated through image processing [21].

The first experimental research into the potential of PCCT for proton therapy was performed by Taasti et al. [23]. It was demonstrated that PCCT yields significantly lower root mean square (RMS) errors in SPR prediction compared to SECT and DECT, with RMS errors of 0.8%-1.0% using two and four energy bins on the organic tissue samples from [15]. Simard et al. [24] showed with an experimental setup with human substitute phantoms that PCCT outperforms DECT in SPR predictions. Hu et al. [25] confirmed comparable SPR accuracy to DECT in phantom studies using PCCT. Larsson [26] utilized neural networks for converting PCCT images to SPR maps, achieving RMS errors of 0.54-1.25%, which is comparable to DECT-based methods.

D. Research question

It has been observed that PCCT has shown to reduce the errors of the SPR predictions compared with SECT and DECT,

and that DirectSPR has done the same when applied to DECT scans. Combining these insights, it can be expected that the combination of PCCT with DirectSPR could reduce the errors of SPR predictions even further. Besides that, it is acknowledged that the recently published consensus guide suggests that additional verification for a HLUT can be performed using biological tissues, mimicking a more clinical setting than experiments with phantoms [10].

Consequently, the research gap lies in the fact that no experimental validation with animal tissue has been performed on the use of PCCT for SPR predictions with the DirectSPR software tool. Therefore the following research question has been posed:

> *To what extend can the implementation of Photon-Counting CT with DirectSPR prediction enhance proton therapy treatment planning by improving the SPR predictions?*

To answer this research question, the PCCT-based SPR predictions will be compared with SECT and DECT-based SPR predictions of samples with animal tissue. In order to make a valid comparison, validation measurements of the sample will be performed with a multi-layer ion chamber (MLIC).

II. Methods

A. Sample preparation

Sample holders (cylinders with an outer diameter of $10\,\mathrm{cm}$) were designed in Onshape (PTC, Boston, MA, USA) and 3D-printed (Fig. 2a, such that the lid of the sample holder fitted inside the sample holder. This allowed pushing the lid into the tissue, making sure that no air was present in the sample. Having air within the sample compromises the assumption that the inner length of the sample is made up of animal tissue, which would lead to incorrect SPR predictions. The lid was glued to the inside of the sample holder to fixate the sample, ensuring that the sample did not change between the proton beam measurement and the CT scans, which would in turn would also lead to wrong assumptions and thus incorrect SPR predictions.

The sample holders were designed in such a manner that two sample holders would fit exactly inside the used phantoms, as can be seen from the front in Fig. 2b and from the top in 2d. Sample holders were filled with animal tissue that was collected on the day of the measurements at the butcher. The following tissue samples were used: beef steak, ground beef and beef bone marrow.

Two measurement series have been performed. Each series was performed on a single day, to prevent changes in the composition and position of the animal tissue within the sample holder. For the first measurements series, ground beef and steak were used. The steak had to be rolled up to fit in the sample holder. For the second measurement series, ground beef, steak and bone marrow were used. A piece of steak was cut from a larger solid chunk of steak, so it fitted exactly in the sample holder. This prevents that the steak had to be rolled or cut to fit, which can lead to air being present in the samples.

TABLE I: CT acquisition parameters

	Modality	Voltage	CTDIvol	Exposure
First series	SECT	120 kVp	Body: 19.9 mGy	300 mAs
			Head: 40.3 mGy	280 mAs
	DECT	80 kVp	Body: 8.4 mGy	450 mAs
			Head: 18.5 mGy	350 mAs
		140 kVp	Body: 10.9 mGy	106 mAs
			Head: 21.6 mGy	83 mAs
	PCCT	120 kVp	Body: 19.9 mGy	mAs
			Head: 40.3 mGy	253 mAs
Second series	SECT	120 kVp	Body: 11.5 mGy	300 mAs
			Head: 36.8 mGy	280 mAs
	DECT	80 kVp	Body: 5.5 mGy	450 mAs
			Head: 12.6 mGy	350 mAs
		140 kVp	Body: 7.1 mGy	106 mAs
			Head: 15.1 mGy	83 mAs
	PCCT	120 kVp	Body: 11.4 mGy	145 mAs
			Head: 36.7 mGy	218 mAs

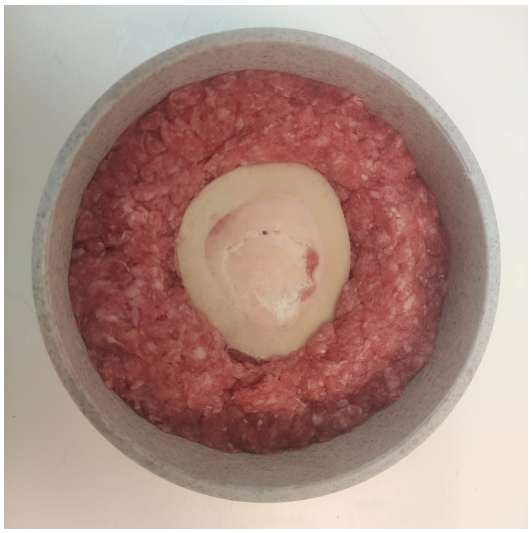
B. CT scanning

For the CT acquisition, the samples were placed in a head and a body configuration of a CTDI phantom (Sun Nuclear, Melbourne, Australia) with diameters of respectively 16 and $32\,\mathrm{cm}$ (Body configuration in Fig. 2b). To mimic clinical practice, scans were made with a standard clinical abdomen protocol for the body phantom and a neuro protocol for the head phantom. The SECT and DECT scans were made at the HollandPTC with the SOMATOM Definition Edge 128-slice dual-energy CT scanner (Siemens Healthineers, Erlangen, Germany). The NAEOTOM Alpha photon-counting CT scanner (Siemens Healthineers, Erlangen, Germany) was used to make the PCCT scans at the Erasmus MC. The PCCT scans were saved as Spectral Post Processing files, which include all spectral data that the software will use. CT acquisition parameters can be seen in Table I. Within the same series, the $\mathrm{CTDI_{vol}}$, which is a standardized measure of radiation dose in CT imaging, of the PCCT scan is matched to that of the SECT scan. In addition, all scans were reconstructed with a Qr40 reconstruction kernel.

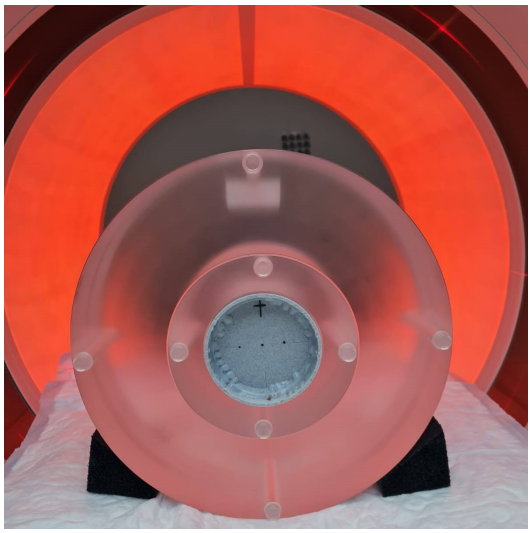
C. SPR prediction

For the conversion of the DECT and the PCCT scans to SPR values, DirectSPR software has been used. This fully automated software (eXamine (Siemens Healthineers, Forchheim, Germany)) makes use of the spectral information from the DECT and the PCCT scans to differentiate between different tissues. The output of the software is a volume with SPR values for every voxel instead of the original HU value. These are subsequently imported in Raystation (Raysearch Laboratories, Stockholm, Sweden) software. Volume of interest (VOI) sizes were based on Gafchromic film that provided a spot size of the proton beam of $2\,\mathrm{cm}$ and the tissue inside the sample holder that was $\sim 4\mathrm{cm}$ long. After characterization of different VOI sizes, in order to test for homogeneity influences, VOIs with a cylinder shape with a length of $3\,\mathrm{cm}$ and a diameter of $2\,\mathrm{cm}$ were created to include a sufficient amount of tissue sample, while excluding the edges (Fig. 2c and 2d). The average SPR in the VOIs has now been extracted from the DECT and PCCT-based SPR volumes.

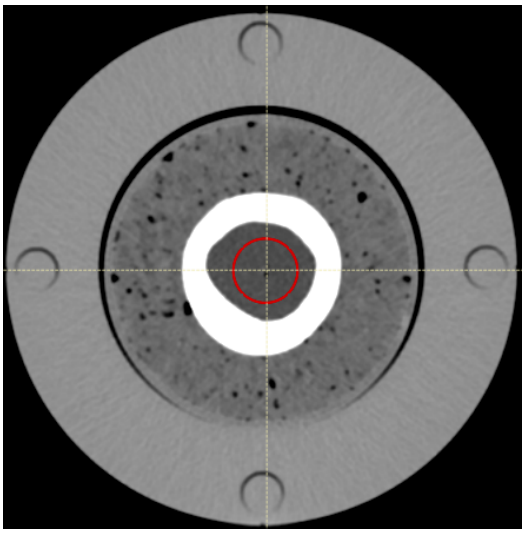
To go from a SECT scan to a SPR prediction, the CT scans are directly imported in Raystation and identical VOIs



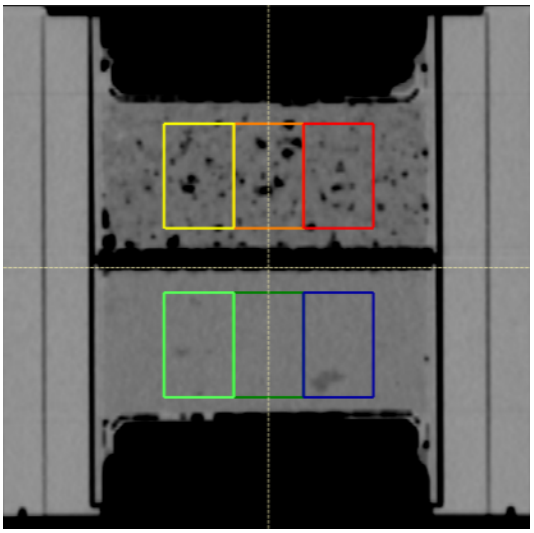
(a) 3D-printed sample holder with bone marrow, filled up with ground beef.



(b) Body phantom with sample in PCCT scanner. Three dots indicate the left, center and right mark where the proton beam was centered.



(c) Cranial view of VOI selection in bone marrow sample. Head phantom in PCCT from second measurement series.



(d) Ventral view of VOI selection in ground beef (cranial) and steak (caudal). Head phantom in DECT from second measurement series.

Fig. 2: Overview of the workflow

are made. This provides the average HU value in the VOI, which still need to be converted to SPR values. As explained in Section I-B, HLUTs are used to do this. This table was calibrated according to the consensus guide from [10]. Two different HLUTs are calibrated, one for the head phantom and once for the body phantom.

D. Validation

Validation was performed by measuring the SPR of each sample. After irradiating the sample with a proton beam and measuring the depth-dose profile using a MLIC, the SPR was calculated with two decimal places accurate. The results of these measurements were used as ground truth to compare the CT-based SPR predictions with.

The sample was positioned in line with the proton beam with the help of lasers, ensuring the beam traverses the center of the sample. The output of the MLIC is a depth-dose profile of the proton beam. The distance to the 80% distal fall-off of the Bragg-peak is used as range. To be able to calculate the SPR of the animal tissue in the sample holder, these measurements are performed with an empty sample holder (R_{empty}) and with the filled sample holders (R_{filled}). The difference between the two resulting ranges is the range difference due to the animal tissue, called the water equivalent thickness (WET):

$$WET_{\text{sample}} = R_{\text{filled}} - R_{\text{empty}} \,. \tag{3}$$

The water equivalent ratio (WER) is then calculated by dividing the WET over the physical thickness (h_{sample}) of

TABLE II: SPR predictions and true SPR (175 MeV) from first measurement series

Phantom Configuration	Sample	SECT	DECT	PCCT	True
Body	Ground beef middle	1.00	1.01	1.01	1.00
	Steak middle	1.04	1.04	1.05	1.02
	Steak right	1.05	1.05	1.06	1.03
Head	Ground beef middle	1.00	1.00	1.00	1.00
	Steak middle	1.05	1.05	1.05	1.02
	Steak right	1.05	1.06	1.06	1.03

the animal tissue (measured with an electric caliper):

$$WER_{\text{sample}} = \frac{WET_{\text{sample}}}{h_{\text{sample}}}\,. \tag{4}$$

This allows calculating the SPR of the sample using the following formula

$$SPR_{\text{sample}} = \frac{WER_{\text{sample}}}{WER_{\text{water}}}\,, \tag{5}$$

where $WER_{water} = 1$, which means that

$$SPR_{\text{sample}} = WER_{\text{sample}}\,. \tag{6}$$

For the first measurement series, the ground beef sample was scanned at 150 MeV and 175 MeV to verify that beam energy had no significant effect (< 0.01) on the resulting SPR of the sample, this would ensure that a single measurement is sufficient to determine the SPR of the sample. The steak sample was scanned at 175 MeV. The measurements of the second series were performed at 150 MeV.

In addition, for the steak and ground beef samples in the second series, measurements were also performed $2\,\text{cm}$ right and left of the center of the sample to verify the homogeneity of the animal tissue.

E. Analysis

The results from the SPR predictions will be compared with the SPR values from the validation measurements. The relative error of the predicted SPR to the ground truth SPR from the validation measurements will be calculated as follows:

$$Error = (\frac{SPR_{\text{predicted}} - SPR_{\text{validation}}}{SPR_{\text{validation}}} - 1) * 100 \tag{7}$$

The resulting error will be compared with values from literature on SECT and DECT-based SPR predictions of animal tissues (Appendix A).

III. Results

A. Proton beam measurement

Using the results of the first measurement series (Tab. II), it was checked whether the energy of the proton beam affected the outcome of the measurement. A measurement of the ground beef sample at 150 MeV and 175 MeV yielded WETs of respectively 42.19 and 42.13 mm for the animal tissue itself, both resulting in an SPR of 1.00. Since this led to differences in SPR < 0.01, the measurements of the second series were performed at 150 MeV. The steak measurements of the first series were still irradiated at 175 MeV. In addition, the WET of steak sample was measured through the center of the sample and $2\,\text{cm}$ to the right. This yielded respectively WETs of 46.13 and 46.47 mm and SPR values of 1.02 and 1.03. Since this difference of ≥ 0.01 is present, one can not assume entirely homogeneous samples. Therefore, the WET measurements of the ground beef and steak samples were performed at three spots on the sample for the second measurement series, being in the center and $2\,\text{cm}$ left and $2\,\text{cm}$ right of the center.

TABLE III: SPR predictions and true SPR (150 MeV) from second measurement series

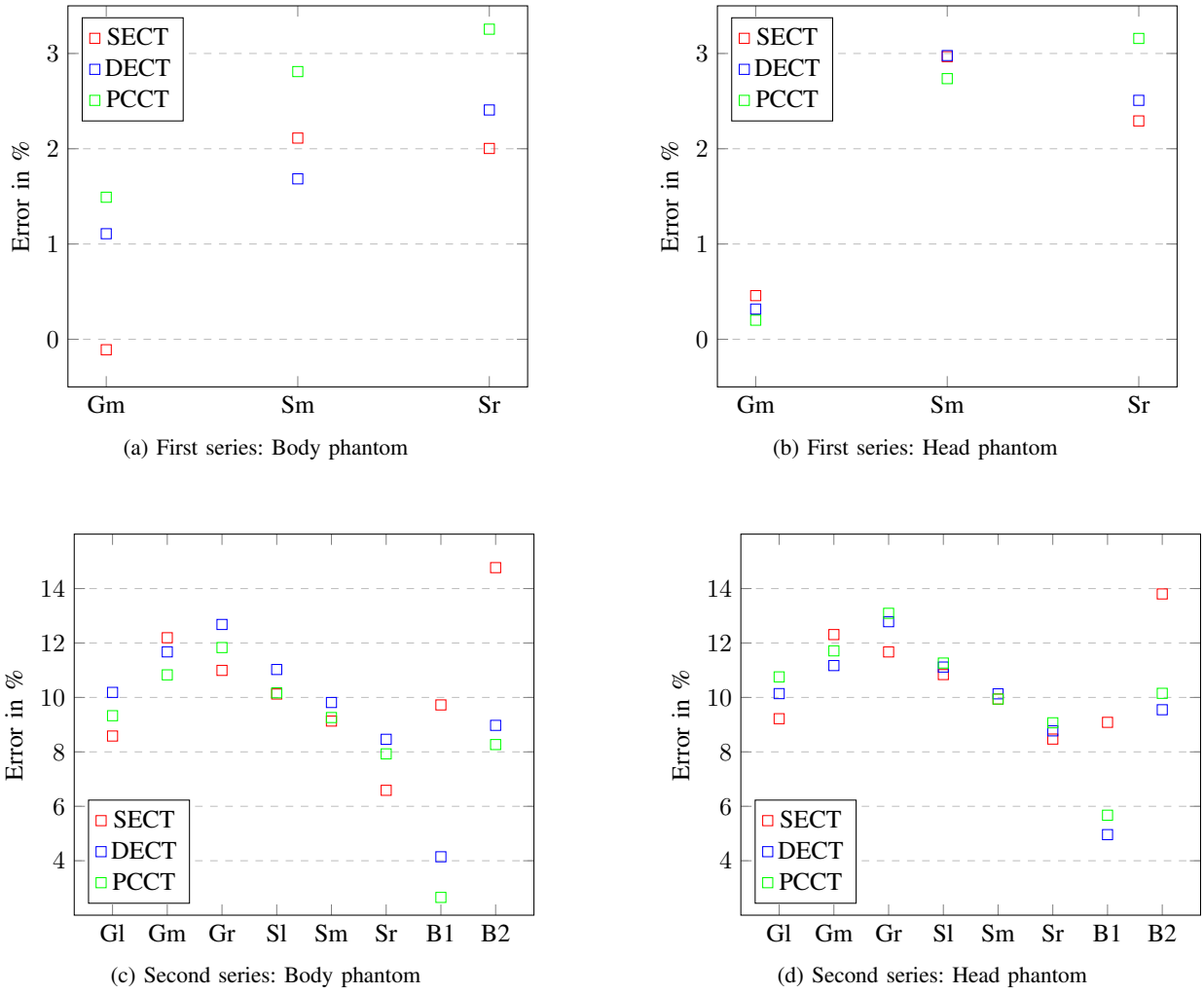
Phantom Configuration	Sample	SECT	DECT	PCCT	True
Body	Ground beef left	1.00	1.01	1.01	0.92
	Ground beef middle	1.00	0.99	0.99	0.89
	Ground beef right	1.00	1.01	1.01	0.9
	Steak left	1.06	1.07	1.06	0.96
	Steak middle	1.06	1.07	1.06	0.97
	Steak right	1.04	1.06	1.06	0.98
	Bone marrow 1	1.00	0.95	0.93	0.91
	Bone marrow 2	1.00	0.95	0.94	0.87
Head	Ground beef left	1.00	1.01	1.02	0.92
	Ground beef middle	1.00	0.99	0.99	0.89
	Ground beef right	1.01	1.02	1.02	0.9
	Steak left	1.06	1.07	1.07	0.96
	Steak middle	1.07	1.07	1.07	0.97
	Steak right	1.06	1.07	1.07	0.98
	Bone marrow 1	0.99	0.96	0.96	0.91
	Bone marrow 2	0.99	0.95	0.96	0.87

When comparing results from the second series (Tab. III) with results from the first series, it stands out the resulting SPRs for ground beef and steak came out $\sim 10\%$ lower than the SPR for ground beef and steak in the first series.

B. SPR prediction

The results of the first measurement series (Table II) show SPR values of 1.00 and 1.01 for ground beef and between from 1.04 to 1.06 for steak. Within the same measurement, the difference between all CT-modalities is $\leq 1\%$. In addition, the difference in SPR values between the body and the head phantom is on average 0.59%. The error (Fig. 3a and 3b) is low for ground beef, especially for the SECT-based SPR prediction, which shows $< 0.5\%$ error. The error of the SPR predictions for steak however are above 2.5% on average. Within the same measurement, the difference between all CT-modalities is $\leq 1\%$.

The results of the second measurement series (Table III) show comparable performance of all CT-modalities for the ground beef and steak samples, with the difference between all CT-modalities within the same measurement being $\leq 1\%$. However SECT gives SPR values of 0.99 and 1.00 for the bone marrow samples, whereas DECT and PCCT give SPR values from 0.93 to 0.96 for the bone marrow samples. In addition, there is a difference present between the body and head phantom for the bone marrow sample. The difference between SECT and spectral CT is approximately 6% for the body phantom and 4% for the head phantom. The SPR predictions for ground beef (0.99-1.01) and steak (1.04-1.07) are in the same range as the SPR predictions from the first



(a) First series: Body phantom

(b) First series: Head phantom

(c) Second series: Body phantom

(d) Second series: Head phantom

Fig. 3: Percentage error of SPR predictions for all tissue samples: Ground beef left (Gl), ground beef middle (Gm), ground beef right (Gr), steak left (Sl), steak middle (Sm), steak right (Sr), bone marrow 1 (B1) and bone marrow 2 (B2).

measurement series for those tissues. In addition, difference in SPR values between the body and the head phantom is on average 0.78%. The error (Fig. 3c and 3d) shows differences over 10% between the SPR predictions and the true SPR.

IV. Discussion

Animal tissue samples have been used to compare SPR prediction performance of three CT modalities. These were compared with ground truth measurements from each sample, measured using a MLIC in a proton beam setup. Results from the first measurement series show errors ranging from close to 0% up to 3%, while results from the second series show errors over 10%.

When comparing the ground truth SPR values to literature, it can be seen that the results of the first series are comparable with literature, whereas the results of the second series are not. For the second series, an SPR value of 0.9 on average was found for ground beef, which is lower than 1.017, found by [15]. An SPR value of 0.97 on average for steak was found, which is lower than 1.06, found by [17], 1.042, found by [14] and 1.054 for beef shoulder, found by [18]. An SPR values of 0.89 on average for adipose tissue was found, which is lower than 0.96, found by [17], 0.947, found by [14] and 0.975 for beef suet, found by [15]. However, the results of the first series match the aforementioned values from literature.

The error of the SECT-based SPR predictions of the first measurements series for ground beef are better than values from literature (Appendix A). However, the DECT and PCCT based SPR predictions perform worse than stated in literature, which makes it plausible that this SECT-based result was incidental. In addition, the error of the SPR predictions of all CT-modalities for the steak sample is higher than found in literature.

The comparison with the true SPR for the second measurements series is not valid, because of the unreliable results from the validation measurements. However, these SPR predictions can be compared with each other. As observed, it stands out that no specific trend can be observed for any of the CT-modalities. A more accurate performance of DECT and PCCT compared with SECT that is found in literature (Appendix A) is not present in these results.

Differences between the body and the head configuration

of the phantom were small, being 0.59% and 0.78% for respectively the first and the second measurement series. The only difference was observed for the bone marrow samples, where the predicted SPR values were lower for the body phantom than for the head phantom. This could be caused by beam hardening in the surrounding bone.

A few steps in this experimental process can induce some uncertainties in the results. The physical thickness of the sample is defined as the inner length of the sample holder. This value is used to calculate the WET and thus the SPR. The way this value was measured might induce some error in the true SPR calculation, since it was measured by hand with an electronic caliper. A recommendation would be to design sample holders that have a fixed inner dimension (instead of pushing the lid into the tissue), excluding this source of uncertainty.

Another butcher was found to get the meat for the second measurements series, since the first butcher could not provide large cuts of steak to fit inside the sample holders. A difference in meat-fat ratio in the ground beef can induce an error up to 0.01 in the SPR, based on a maximum difference of 10% in the ratio.

In addition, during the process of making the samples, the lid was pushed into the meat to ensure that no air was present in the sample. This can lead to a higher density of the meat, leading to a higher SPR. This can explain a difference in measured ground truth SPR compared with literature. However, the CT-based SPR predictions on that sample should provide SPR values in the same range as the measured ground truth SPR values, which is not the case for the results of the second series, seeing that the error is approximately 10% on average.

These sources of error do not explain the lower ground truth SPR values for all samples from the second measurement series. These values were approximately 10% lower compared with the first series (only ground beef and steak) and values from literature. No particular reason or combination of reasons has been found for these discrepancies. The fact that the difference was this large, rules out the possibility of a mere measurement error in the inner diameter of the sample for example. It could be that the combination of all sources of uncertainties led to a larger error, this is however unlikely, because this would lead to large variations between samples as well, which does not seem to be the case.

This would however be in line with what Peters et al. [10] stated in the consensus guide on additional verification measurements with animal tissue, namely that they 'need to be weighed against the potentially large experimental uncertainties in these complex validation scenarios'.

In order to verify whether the discrepancies of the second measurements series were incidental, repetition of this measurement series is necessary. Standardized samples holders with a fixed inner volume and more homogeneous meat could reduce the risk of having large uncertainties in the results. In addition, proton beam measurements with a water absorber of variable thickness [17] are known to be more accurate than MLICs. However, measurements with such a system take a few hours, which would compromise the goal to perform all proton beam measurements and CT scans on the same day to prevent change in the sample itself.

More research should be conducted into a comparison of these results with results of the same scans and SPR predictions using standardized phantoms instead of animal tissue samples. In case that these would show similar results, this would confirm that PCCT does not show potential for better SPR predictions.

V. Conclusion

The SPR predictions based on the PCCT scans were not more accurate than those based on the SECT and DECT scans. Therefore, it cannot be concluded that that PCCT would enhance the treatment quality by improving the SPR predictions. Further research with animal tissues as well as standardized phantoms is necessary to evaluate PCCT's true potential in improving proton therapy with DirectSPR predictions.

Acknowledgments

I would like to thank my HollandPTC supervisor Sophie Huijskens and my TU Delft supervisor Marlies Goorden. In addition, I would like to thank Marcel van Straten for the help with the PCCT scans at the Erasmus MC and Patricia Cambraia Lopes for the help with the SECT and DECT scans at the HollandPTC. Lastly, I would like to thank Ernst van der Wal and Thomas Toet for the help with the proton beam measurements.

References

[1] Delaney, G., Jacob, S., Featherstone, C., & Barton, M. (2005). The role of radiotherapy in cancer treatment. *Cancer*, *104*(6), 1129–1137.

[2] Paganetti, H., Beltran, C., Both, S., Dong, L., Flanz, J., Furutani, K., Grassberger, C., Grosshans, D. R., Knopf, A.-c., Langendijk, J. A., Nystrom, H., Parodi, K., Raaymakers, B. W., Richter, C., Sawakuchi, G. O., Schippers, M., Shaitelman, S. F., Teo, B. K. K., Unkelbach, J., ... Lomax, T. (2021). Roadmap : proton therapy physics and biology. *Physics in Medicine & Biology*, *66*, 1–61.

[3] Schneider, U., Pedroni, E., & Lomax, A. (1996). The calibration of CT Hounsfield units for radiotherapy treatment planning. *Physics in Medicine & Biology*, *41*, 111–124.

[4] Wohlfahrt, P., & Richter, C. (2020). Proton therapy special feature: Review article status and innovations in pre-treatment CT imaging for proton therapy. *British Journal of Radiology*, *93*(July), 1–14.

[5] Radiology Key. (2017). Particle Therapy in the Third Millennium: Current Status and Future Outlook. https://radiologykey.com/particle-therapy-in-the-third-millennium-current-status-and-future-outlook/

[6] Wohlfahrt, P., Möhler, C., Greilich, S., & Richter, C. (2017). Comment on: Dosimetric comparison of stopping-power calibration with dual-energy CT and single-energy CT in proton therapy treatment planning [Med. Phys. 43(6), 2845-2854 (2016)]. *Medical Physics*, *44*(10), 5533–5536.

[7] Taasti, V. T., Bäumer, C., Dahlgren, C. V., Deisher, A. J., Ellerbrock, M., Free, J., Gora, J., Kozera, A., Lomax, A. J., De Marzi, L., Molinelli, S., Kevin Teo, B. K., Wohlfahrt, P., Petersen, J. B., Muren, L. P., Hansen, D. C., & Richter, C. (2018). Inter-centre variability of CT-based stopping-power prediction in particle therapy: Survey-based evaluation. *Physics and Imaging in Radiation Oncology*, *6*, 25–30.

[8] Goitein, M. (1985). Calculation of the uncertainty in the dose delivered during radiation therapy. *Medical Physics*, *12*(5), 608–612.

[9] Peters, N., Wohlfahrt, P., Dahlgren, C. V., de Marzi, L., Ellerbrock, M., Fracchiolla, F., Free, J., Gomà, C., Góra, J., Jensen, M. F., Kajdrowicz, T., Mackay, R., Molinelli, S., Rinaldi, I., Rompokos, V., Siewert, D., van der Tol, P., Vermeren, X., Nyström, H., ... Richter, C. (2021). Experimental assessment of inter-centre variation in stopping-power and range prediction in particle therapy. *Radiotherapy and Oncology*, *163*, 7–13.

[10] Peters, N., Trier Taasti, V., Ackermann, B., Bolsi, A., Vallhagen Dahlgren, C., Ellerbrock, M., Fracchiolla, F., Gomà, C., Góra, J., Cambraia Lopes, P., Rinaldi, I., Salvo, K., Sojat Tarp, I., Vai, A., Bortfeld, T., Lomax, A., Richter, C., & Wohlfahrt, P. (2023). Consensus guide on CT-based prediction of stopping-power ratio using a Hounsfield look-up table for proton therapy. *Radiotherapy and Oncology*, *184*, 109675.

[11] Goitein, M. (1977). The measurement of tissue heterodensity to guide charged particle radiotherapy. *International Journal of Radiation Oncology, Biology, Physics*, *3*, 27–33.

[12] Wohlfahrt, P., Möhler, C., Hietschold, V., Menkel, S., Greilich, S., Krause, M., Baumann, M., Enghardt, W., & Richter, C. (2017). Clinical Implementation of Dual-energy CT for Proton Treatment Planning on Pseudo-monoenergetic CT scans. *International Journal of Radiation Oncology*Biology*Physics*, *97*(2), 427–434.

[13] Vilches-Freixas, G., Létang, J. M., Ducros, N., & Rit, S. (2017). Optimization of dual-energy CT acquisitions for proton therapy using projection-based decomposition. *Medical Physics*, *44*(9), 4548–4558.

[14] Xie, Y., Ainsley, C., Yin, L., Zou, W., McDonough, J., Solberg, T. D., Lin, A., & Teo, B. K. K. (2018). Ex vivo validation of a stoichiometric dual energy CT proton stopping power ratio calibration. *Physics in Medicine & Biology*, *63*(5), 055016.

[15] Taasti, V. T., Michalak, G. J., Hansen, D. C., Deisher, A. J., Kruse, J. J., Krauss, B., Muren, L. P., Petersen, J. B., & McCollough, C. H. (2018). Validation of proton stopping power ratio estimation based on dual energy CT using fresh tissue samples. *Physics in Medicine & Biology*, *63*(1), 1–12.

[16] Bär, E., Lalonde, A., Zhang, R., Jee, K. W., Yang, K., Sharp, G., Liu, B., Royle, G., Bouchard, H., & Lu, H. M. (2018). Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples. *Medical Physics*, *45*(1), 48–59.

[17] Möhler, C., Russ, T., Wohlfahrt, P., Elter, A., Runz, A., Richter, C., & Greilich, S. (2018). Experimental verification of stopping-power prediction from single- and dual-energy computed tomography in biological tissues. *Physics in Medicine & Biology*, *63*, 025001.

[18] Niepel, K. B., Stanislawski, M., Wuerl, M., Doerringer, F., Pinto, M., Dietrich, O., Ertl-Wagner, B., Lalonde, A., Bouchard, H., Pappas, E., Yohannes, I., Hillbrand, M., Landry, G., & Parodi, K. (2021). Animal tissue-based quantitative comparison of dual-energy CT to SPR conversion methods using high-resolution gel dosimetry. *Physics in Medicine & Biology Biology*, *66*(7), 075009.

[19] Sarkar, V., Paxton, A., Su, F., Price, R., Nelson, G., Szegedi, M., James, S. S., & Salter, B. J. (2023). An evaluation of the use of DirectSPR images for proton planning in the RayStation treatment planning software. *Journal of Applied Clinical Medical Physics*, *24*(5).

[20] Peters, N., Wohlfahrt, P., Hofmann, C., Möhler, C., Menkel, S., Tschiche, M., Krause, M., Troost, E. G., Enghardt, W., & Richter, C. (2022). Reduction of clinical safety margins in proton therapy enabled by the clinical implementation of dual-energy CT for direct stopping-power prediction. *Radiotherapy and Oncology*, *166*, 71–78.

[21] Willemink, M. J., Persson, M., Pourmorteza, A., Pelc, N. J., & Fleischmann, D. (2018). Photon-counting CT: Technical principles and clinical prospects. *Radiology*, *289*(2), 293–312.

[22] Flohr, T., Petersilka, M., Henning, A., Ulzheimer, S., Ferda, J., & Schmidt, B. (2020). Photon-counting CT review. *Physica Medica*, *79*, 126–136.

[23] Taasti, V. T., Hansen, D. C., Michalak, G. J., Deisher, A. J., Kruse, J. J., Muren, L. P., Petersen, J. B., & McCollough, C. H. (2018). Theoretical and experimental analysis of photon counting detector CT for proton stopping power prediction. *Medical Physics*, *45*(11), 5186–5196.

[24] Simard, M., Lapointe, A., Lalonde, A., Bahig, H., & Bouchard, H. (2019). The potential of photon-counting CT for quantitative contrast-enhanced imaging in radiotherapy. *Physics in Medicine & Biology*, *64*, 115020.

[25] Hu, G., Niepel, K., Risch, F., Kurz, C., Würl, M., Kröncke, T., Schwarz, F., Parodi, K., & Landry, G. (2022). Assessment of quantitative information for radiation therapy at a first-generation clinical photon-counting computed tomography scanner. *Frontiers in Oncology*, *12*(September), 1–10.

[26] Larsson, K. (2023). *Improving proton therapy planning with photon-counting spectral computed tomography* [Doctoral dissertation].

APPENDIX A
OVERVIEW OF EXPERIMENTAL RESEARCH ON DECT FOR SPR PREDICTIONS WITH ANIMAL TISSUES

TABLE IV: Overview of experimental research on DECT for SPR predictions with animal tissues

	Animal tissue: Beef	Animal tissue: Pork	Proton beam energy	Proton beam measurement	SPR prediction methods	Results
Bar et al.	Adipose, bone	Adipose, muscle, liver, kidney, brain, heart, blood and lung	195 MeV	Dose extinction method	SECT: HLUT DECT: 1. Bethe-Bloch (Rho and I-value) 2. Bethe-Bloch (elemental compositions)	RMS error for soft tissues and bones: SECT: 0.53% and 1. 37% DECT: 1. 0.19% and 1.06% 2. 0.38% and 1.06%
Mohler et al.	Bone marrow, cortical bone	Back fat (smoked), belly fat (fresh), loin (fresh), fillet (fresh), thick flank (fresh), liver (fresh), kidney (fresh), brain (fresh), heart (fresh), blood (fresh) and lung (fresh)	200.28 MeV	Water absorber of variable thickness	SECT: HLUT DECT: syngo Rho/Z software	Signed error: SECT: 0.84% +- 0.12% DECT: No significant error
Taasti et al.	Ground beef, kidney (chopped), suet, round (chopped and unchopped) and femur	Loin (chopped and unchopped), liver (chopped), fat (chopped and unchopped), brain (fresh), brain (frozen, thawed) and blood (mixed with heparin)	185.2 MeV	MLIC	SECT: HLUT DECT: syngo Rho/Z software	RMS error: SECT: 2.78% DECT: 0.47%
Xie et al.	Shank and muscle	Brain, kidney and liver	130, 140, 150, 160 MeV	MLIC	SECT: HLUT DECT: Bethe-Bloch (Rho and I-value)	Error: SECT: 0.55 +-1.94% DECT: 0.07 +- 0.58%
Niepel et al.	Shoulder, tongue, liver, filet and fat	Brain	200 MeV	Water absorber of variable thickness	SECT: HLUT DECT: 1. Bethe-Bloch (Rho and I-value) 2. Bethe-Bloch (elemental composition) 3. Direct conversion	Error for muscle, adipose, bone200 and cortical bone SECT: -2.51%, -0.58%, 3.72% and -0.29% DECT: 3. direct conversion: -1.21%, -0.81%, -0.37% and 1.05%
Sarkar et al.	Grounded fat, grounded heart and grounded sirloin	-	164.8 MeV	MLIC	SECT: HLUT DECT: DirectSPR	Signed and absolute difference with WER: SECT: 2.2 mm and 2.2 mm DECT: 0.1 mm and 0.3 mm

A

Literature study

Is latest always the greatest? The potential of Photon-Counting CT for better SPR predictions

Van Walsum, M.J.
TU Delft, Delft, Netherlands
Dept. of Radiology and Nuclear Medicine, Dept. of Radiotherapy of Erasmus MC, Rotterdam, Netherlands
HollandPTC, Delft, Netherlands
m.j.vanwalsum@student.tudelft.nl

***Abstract*—This literature study aims to provide a clear overview of research that has been performed stopping power ratio (SPR) predictions for proton therapy, using single-energy CT (SECT), dual-energy CT (DECT) and photon-counting CT (PCCT) scanners. Using the PRISMA method, it was found that studies on the comparison between conventional SECT and DECT-based SPR predictions conclude that DECT outperforms SECT as planning CT for proton therapy. In addition, PCCT was introduced and the working principle and differences with previously stated CT techniques were explained. Limited research has been found on the use of PCCT as base for the SPR prediction for proton therapy. Of which no studies used animal tissue for experiments and predicting the SPR using DirectSPR. It was concluded that this research gap exists and a research plan is proposed for a series of measurements, based on the experiments on SECT and DECT.**

***Index Terms*—Proton therapy, stopping power ratio, CT, photon-counting CT**

I. Introduction

A. Background

Among surgery and chemotherapy, radiotherapy is a common treatment method for cancer [1]. The aim of radiotherapy is to irradiate and kill the tumor, while sparing as much of the healthy surrounding tissues as possible. Therefore, healthy tissues should receive as little dose as possible, since ionizing radiation could damage these tissues. Even at low doses, long-term side effects can occur due to the probabilistic effects of ionizing radiation.

Photon therapy, the most common form of radiotherapy, has a dose distribution where the dose deposition is highest where the beam enters the body and then decreases as the beam penetrates deeper into the body. Proton therapy has a dose distribution that differs considerably from the dose distribution of photon therapy (Fig. 1) The protons exhibit a 'Bragg-peak' at a certain depth, where the dose deposition is at its highest. After this Bragg-peak, the dose deposition quickly becomes close to zero. This allows focusing proton beams in a manner that the dose is highest at the tumor, which is more difficult with photon therapy and its dose distribution. This is especially favorable when critical organs are close to or in pediatric cases where minimizing long-term side effects is crucial [2].

Just as photon therapy, CT imaging uses photons. However, these photons have an energy that is roughly one hundred times lower than that of photon therapy. Electrons are fired from a cathode to an anode, where they collide with the anode material. Some of these collisions lead to the release of X-rays (photons) towards the patient. A part of the photons are attenuated (scattered and/or absorbed) in the body. The other part of these photons traverse the body, reaching the detector, which gives the attenuation of X-rays in the body. By imaging from multiple angles, a 3D volume of the patient is reconstructed from multiple 2D images.

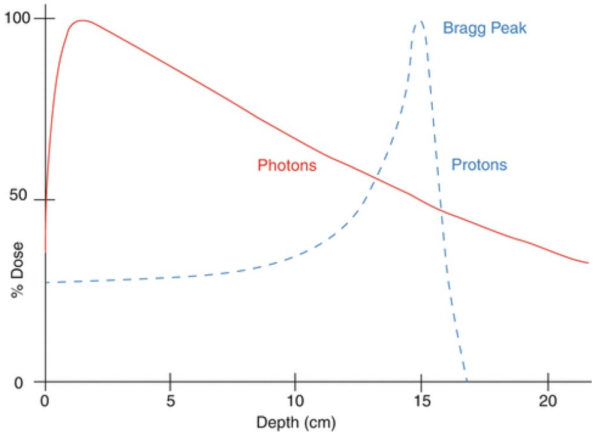


Fig. 1: Depth dose distribution of photons and protons. From [3]

B. Imaging for Proton Therapy

The importance of imaging for proton therapy comes from its crucial role within the proton therapy workflow (Fig. 2). There are two reasons for the use of imaging in the proton therapy workflow. The first reason, also necessary for photon therapy, is localizing and delineation of the tumor and the OAR's. In all cases, Computed Tomography (CT) is used for this purpose. In addition, Positron Emission Tomography (PET) and Magnetic Resonance Imaging (MRI) can also be used for this. Whereas CT can distinguish well between bone and soft tissue, MRI and PET are used to identify soft tissues. The second purpose of imaging however, which is only necessary for proton therapy, is deriving the Stopping Power Ratio (SPR). For this purpose, only the CT scan is used, which is then called a planning CT. The SPR is used in treatment planning to predict how the protons will interact with

Fig. 2: Schematic overview of the proton therapy workflow

the various tissues within the body. It accounts for the relative stopping power of the tissue compared to water along the path of the proton beam. Different body compositions affect the way and how much the protons interact with tissue and thus where the Bragg peak will appear. Having an accurate SPR is therefore of the utmost importance since the Bragg peak should be focused exactly on the tumor. The more accurate the SPR is, the smaller we can make the uncertainty margins in proton therapy. Currently, uncertainty margins of 3-3.5%, sometimes plus an absolute term of 1-2 mm, are used [4], which stems from research from 1985 that defined these margins [5]. One of the long-term goals of proton therapy research, defined by Paganetti et al. [2], is to improve the range prediction accuracy to $\leq$1%.

C. Research Question

As concluded above, a prerequisite of a good proton treatment plan is to have an accurate SPR map. Therefore, the research question for this literature study will be:

How do current SPR prediction methods perform and can photon-counting CT improve them?

This introduction has shown the essential role of CT imaging and how it can influence the treatment quality. This literature review will dig deeper into the use of CT imaging for SPR predictions. Firstly, the used methods for literature retrieval will be explained. The methods for predicting and measuring the SPR, will be discussed. In addition, an explanation of inter-center variability will be given. Thereafter, it will be discussed which improvements Dual-Energy CT (DECT) has brought in comparison with conventional Single Energy CT (SECT) for SPR accuracy. Lastly, a recently introduced Photon-Counting CT (PCCT) scanner and its potential as planning CT for proton therapy will be discussed.

II. Literature Retrieval

The PRISMA [6] flowchart model (Appendix A) is used to identify relevant papers regarding the research question. Since Section IV will answer the main part of the research question of this study, the PRISMA model will be focused on Section IV. The used search query in PubMed and Scopus was 'Dual-energy CT proton therapy'. Papers were included if a comparison was made between the performance of SECT and DECT for SPR prediction, including studies with actual experiments, but these could also be model-based or based on older patient cohorts. An exclusion criterion a methodology that lacks sufficient explanation.

For the other sections of this literature review some relevant literature for this study was provided by my supervisor. Where specific information was needed, academical search engines were used to find the proper literature. For formulas and basic physical principles, original sources were cited.

III. Stopping Power Ratio

This section will elaborate on the SPR. As the use of the SPR already has been explained in Section I-B, this will not be covered in this section. Firstly, common methods for predicting the SPR will be discussed and issues with these current methods will be addressed. In addition, the variability in used methods between different proton therapy centers will be addressed. Lastly, it will be explained how the SPR can be measured, which is an essential part of validating research into improving SPR predictions.

A. Predicting the SPR

As explained in Section I, a planning CT scan is made. This scan is the foundation of the SPR prediction. An elaborate review by Wohlfahrt and Richter [7] showed the different methods in literature to go from CT images to SPR prediction. The most used methods for SECT and DECT will be explained in this section.

1) Single-energy CT: Whereas a direct link is present between attenuation (Equation 1) in CT (HUs) and attenuation in photon therapy through the attenuation coefficient μ, this is not the case for the relation between HUs and SPR.

$$I = I_0 e^{-\mu x} \tag{1}$$

On one hand, attenuation of photons depends on photon interactions in the body, the photoelectric effect, Compton scatter and pair production. On the other hand, particle interactions depend on the Bethe-equation, of which the common form is the corrected Bethe-Bloch equation (Equation 2) [8].

$$-\frac{dE}{dx} = 2\pi N_a r_e^2 m_e c^2 \rho \frac{Z}{A}\frac{z^2}{\beta^2} \left[\ln\left(\frac{2m_e\gamma^2 v^2 W_{max}}{I^2}\right) - 2\beta^2 - \delta - 2\frac{C}{Z}\right] \tag{2}$$

This large formula contains mainly parameters that can not be changed, and shows that the SPR is mainly dependent on the electron density (ρ) and the mean excitation energy (I). However, both photon and particle interactions are linearly dependent on the electron density, which is defined in a generic Hounsfield look-up table (HLUT) [7] and used as standard for SPR predictions in the clinic. A HLUT (Figure 3) consists of multiple line segments for different types of tissues. Especially in the region around 0 HU (adipose and soft tissue), multiple line segments are connected to establish one curve.

The definition and fitting of the HLUT can be done in two ways, the empirical and the stoichiometric calibration method. The empirical method measures the HUs and the SPR of tissue surrogates. These surrogates should resemble the attenuation of photons and the particle interactions of human tissue. Therefore, one can calculate the SPR from these results. However, this introduces the largest issue with this technique, these tissue surrogates do not resemble human tissue as well as was assumed [10]. This was also concluded by Schneider et al. [11], who showed that human tissue leads to a better HLUT than tissue surrogates. Therefore, Schneider et al. [11] introduced the stoichiometric approach, where measurements

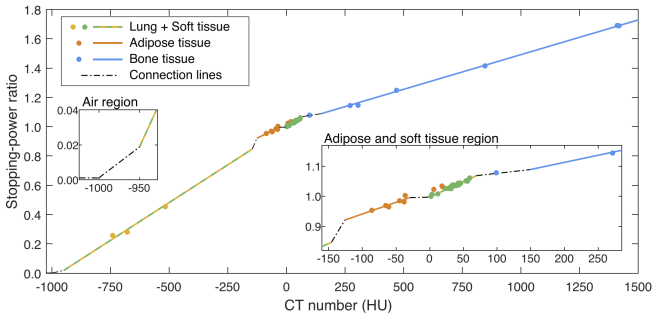


Fig. 3: Example of a Hounsfield look-up table by Peters et al. [9]

on real human tissue (mostly from corpses) are performed to calculate the SPR. Only a few of these datasets with human tissues exist. However, to make a correct HLUT, the measurements should be performed with the CT scanner and parameters of the specific clinic to retrieve the HUs as well. Since this is not possible, tissue surrogates are still used for calibration measurements in order to have reliable HUs calculations. This means that a proper selection of a calibration phantom is necessary, as explained in a comment by Wohlfahrt et al. [12]. The difference between both these methods is thus that the stoichiometric method only uses the tissue surrogates for mimicking the attenuation of photons (HUs) of human tissues and not the particle interactions (SPR) [7]. This is achieved by looking at similar parameters such as the electron density and the effective atomic number. The result of this is that certain soft tissues with different SPRs can not be distinguished since they have the same HUs.

This table is based on the mutual dependency on the electron density, however there are more factors affecting the SPR besides the electron density. The HU-SPR correlation is also dependent on the effective atomic number (photon interactions) and mean excitation energy (particle interactions) [7]. This means that HLUT-based SPR predictions inherently are not perfect.

These SPR prediction methods and their issues are the main cause of uncertainty in range predictions for treatment planning [7]. As explained in Section I, the uncertainty margins have not changed since 1985.

2) Dual-energy CT: The principle of DECT is that, when two CT scans are made with different tube voltages, two different material parameters become known. This spectral information can be used to determine the composition of the scanned tissues. Only four years after the introduction of DECT in 1973, the first ideas for the use of DECT for charged particle therapy were published [13]. Goitein stated that scanning at two tube potentials (90 and 140 kVp) showed potential to improve treatment planning for particle therapy. However, it was until 2015 that it was used in the clinic, when Wohlfahrt et al. [14] published a paper on the clinical implementation of DECT.

Multiple methods are available for using DECT as planning CT and they can categorized in methods for single-source and dual-source CT scanners. The first category consists of

consecutively (dual-spiral mode), fast voltage switching and simultaneously using either a dual-layer detector or a split-beam filter [7]. Of these methods, the first one is the preferred method, since the tube voltage and current can be adjusted independently. In addition, to overcome motion artifacts, deformable image registration is used, sometimes combined with time-resolved CT [7].

Dual-source CT scanners on the other hand, make use of two sources, which can be adjusted independently to increase spectral separation. On many subjects, this technique performs better than the single-source DECT scanner. Cross-scatter radiation in the detector signal can be removed in image processing. However, a main issue is the limited field of view (FOV), caused by the limited space of the gantry for two detectors. The resulting spectral information from a DECT scan can be used for many purposes, some needing their own image processing. A common method of image processing is to use weighted superposition of the HUs from the low (HU_{low}) and the high (HU_{high}) voltage. This results in a virtual monoenergetic CT image (VMI) dataset, which a CT image at one specific energy instead of a conventional energy spectrum, as function of weighting factor alpha α:

$$HU_{VMI}(\alpha) = \alpha * HU_{low} + (1 - \alpha) * HU_{high} \qquad (3)$$

Whereas scattering is more present for higher energies, the photoelectric effect is more present for lower energies. Therefore the contrast of the image can be changed. This means that from the same DECT scan, a high contrast image ($\alpha > 0$) or a minimized metal artifact image ($\alpha < 0$) can be reconstructed [7]. It is also possible to minimize beam hardening by optimizing α. This would lead to more consistent HUs, which is of great importance for proton treatment planning. This approach was also used for the first use of DECT in the clinic [14]. All other issues of the HLUT are still present, but more consistent HUs lead to better SPR predictions.

An overview of the available methods DECT-based SPR predictions methods can be seen in Fig. 4. These methods either use the previously explained VMI CT images, or they use direct SPR predictions. As explained with the Bethe-Bloch formula, the SPR is dependent on the I-value and the ρ. The intrinsic issues with the HLUT-SPR conversion are overcome by using the additional spectral information from the DECT scan [7], allowing direct determination of the tissues parameters.

B. Inter-center variability

Taasti et al. [4] performed research into the variability of the workflows in different proton therapy centers. Twelve clinics were sent a questionnaire, asking about all relevant aspects in the process from planning CT to SPR prediction. A wide variety of scanning parameters and HLUT definitions was found. Eight centers used the stoichiometric method, three centers based their HLUT on tissue-substitute calibrations and one center combined both methods. The number of HLUT line segments varied from two to eleven. Of all centers, only nine had investigated the influence of beam-hardening and ten centers validated their HLUT with experiments. These

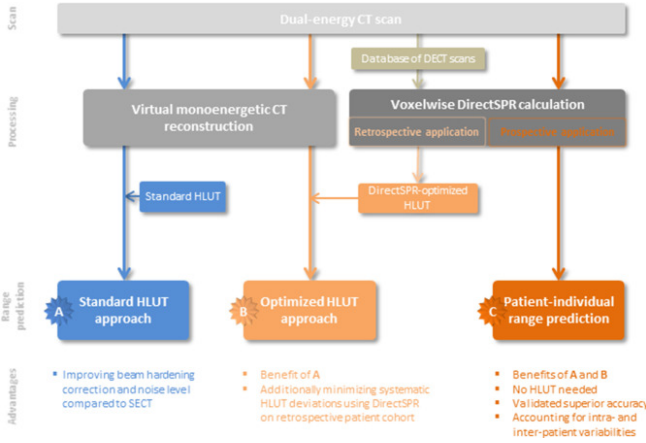


Fig. 4: Overview of DECT-based SPR prediction methods by Wohlfahrt et al. [7]

validation experiments were performed with tissue-equivalent phantoms (six centers) and/or animal tissues (seven centers). Lastly, whereas most centers acknowledged the potential of DECT in improving SPR predictions, only one center actually used DECT. It is recommended that CT-based SPR prediction is standardized for all proton therapy centers. These findings were confirmed by Peters et al. in the next year [15]. A 'ground truth' phantom was scanned at 17 proton therapy centers, all with their own scanning protocols and SPR prediction methods. Results showed inter-center variation in the range predictions (2σ) of 2.9%, 2.6% and 1.3% for typical beam paths of respectively prostate-, lung- and primary brain-tumor treatments. Both the consideration of beam hardening and performing an independent HLUT validation reduced the absolute error on average.

In order to standardize the process of SPR prediction, Peters et al. [9] have published a step-by-step guide on how to specify a HLUT for SPR predictions. The six following steps are defined in the process of setting up a HLUT: Phantom setup, CT acquisition, CT number extraction, SPR determination, HLUT specification, and HLUT validation. For all of these steps, clear explanation has been given and common challenges are assessed. The resulting guide was validated by testing it on three different CT scanners. It was concluded that this step-by-step guide can reduce the inter-center variation in SPR prediction. In addition, it was stated that an optional verification method is to use biological animal tissues to better mimic patient scenarios, which will be the focus of this literature study.

C. Validating the SPR

Measuring the SPR of a sample is a necessary part of the validation. This allows comparing predicted SPR's with the measured (true) SPR, the closer a predicted SPR is to the measured SPR, the better the prediction method performed. Several methods exist to measure the SPR of a certain sample, these will be explained below.

A multi layer ion chamber (MLIC) can be used to measure the dose intensity of the proton beam over depth. It consists

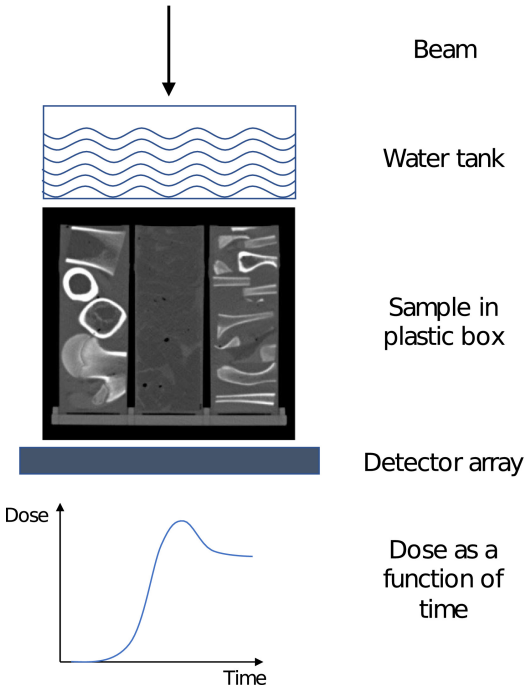


Fig. 5: Water equivalent range (WER) measurement setup by Bär et al. [16]

of many small ionization chambers stacked next to each other. Every individual ionization chamber can measure the dose intensity at that depth. Since the SPR of the whole MLIC is made to be approximately the same as water, for every individual chamber that the beam traverses, an increase of the thickness of a single chamber is added to the water equivalent range (WER). By irradiating a sample holder with and without its content, the difference in WER is retrieved, which can be converted to an SPR of the sample. A Zebra MLIC (IBA Dosimetry GmbH, Schwarzenbruck, Germany) was used by [17] and [18] and has an 0.5 mm accuracy [19]. Taasti et al. used a custom MLIC with 0.2 mm accuracy [20].

Another method was used by Mohler et al. [21]. They used a water absorber of variable thickness 'Peakfinder water column' (PTW, Montreal, US) to record the depth-dose curve. A dual-channel electrometer and two Bragg-peak ionization chambers are placed within a closed water column. This allows monitoring of the beam during the measurements. PTW states that increments of only 1 $0\mu m$ can be made. However, Wohlfahrt only reported sub-millimeter accuracy with a 150 MeV beam [22]. There also exist Bragg-peak ionization chambers that can move through a water phantom. These can also be used to measure dose intensity over depth, resulting in a depth-dose curve [22].

A comparable method (Fig. 5) was used by Bär et al. [16], where a water column was placed in front of the sample, while the detector was placed behind the sample. The water column was slowly drained, so the detector recorded the dose-depth curve in the same manner as for the previously explained method.

Radiochromic film can be used, which was done by Xie et al. [18]. The 'GAFchromic EBT3 film' (Ashland, Bridgewater, US) can be used to visualise the exit dose profile of the beam. After irradiation, the dose maps can then be reconstructed using specific software. Xie et al. state that the accuracy of the range measurements using the film with a solid water phantom and a 150 MeV beam was better than 1 mm [18]. Besides this paper, no other research has been found using radiochromic film for comparing CT techniques to improve SPR predictions.

Lastly, N-vinylpyrrolidone based polymer (VIP) gel can be used to retrieve the SPR of a sample, this was done by Niepel et al. [23]. A beam-shaped container is filled with the VIP gel and is placed behind the sample in the beam line. The irradiated areas of the gel turn white and cloudy. In addition, these areas also have a different spin-spin relaxation that can be detected with MRI. An advantage of this method is that it provides 3D information of the SPR in the sample.

IV. SECT vs DECT

This section will cover research performed to compare the accuracy of SPR predictions utilizing DECT scans versus those based on SECT scans. The studies have been categorized based on whether animal tissues were used in the experiments or not. The different methods of sample preparation, measurements, results and conclusions will be discussed.

A. No animal tissues

Wohlfart et al [24] recalculated treatment plans of 25 head and neck (HN) cancer and 25 prostate cancer patients. The plans were previously based on a HLUT and were recalculated using a DECT-based direct SPR prediction (RhoSigma). Range shifts between the HLUT based and the RhoSigma based SPR prediction were found to be averagely larger than 1%. In addition, large intra-patient range shift variations were found, which confirms that a generic HLUT can not be accurate for all patients.

Taasti et al. [25] included DECT scans of 14 HN cancer patients. For eight patients, a dual-source DECT scanner was used, while the other six patients were scanned by acquiring two consecutive SECT scans. A difference in WER of 1.9 mm was found between SECT and DECT. In addition, a difference of 1% was found between SPR predictions based on these two techniques. Overall, the SPR differences were lower for DECT than for SECT. Therefore, it was concluded that this could imply that accuracy of proton treatment planning would improve by using DECT.

Polf et al. [26] used 3D-printed inserts with septa and placed them in a tissue-equivalent head phantom. DECT scans were made and using the relative electron density (ρ_e) and the effective atomic number (Z_{eff}) the SPR was predicted. In addition, the SPR was calculated using the Bethe-Bloch equation. Then, for each plug the SPR was measured in a clinical proton beam. Results showed that for all inserts with septa thicker than 3 mm, DECT-based SPR predictions were within 3% of the calculated and the measured SPR. No comparison with SECT was made.

Peters et al. [27] investigated the use of a new software tool that can convert spectral CT data to an SPR directly. DirectSPR is a commercial software tool (DirectSPR™, Siemens Healthineers, Erlangen, Germany) that is able to take the raw CT data and convert it directly into an SPR, not needing any processing or other tools. The resulting uncertainty margins of this new method were compared with conventional margins of 3.5% +2 mm. For 150 brain-tumor and 100 prostate-cancer patients, the amount of healthy tissue that was spared in beam direction was respectively 2.6 mm and 4.4 mm. The use of DirectSPR allowed a reduction of the range uncertainty from the conventional margin to 1.7% + 2 mm for brain-tumor and to 2.0% + 2 mm for prostate-cancer patients.

B. Animal tissues

A detailed overview of the following studies can be seen in Table I in Appendix B.

Bär et al. [16] used plastic boxes to place the tissue samples in and filled the boxes up with saline water. The WER was measured using the dose-extinction method, the setup can be seen in Fig. 5. A beam energy of of 195 MeV was used. In order to estimate the SPR from the CT data, a calibration was performed using a tissue characterization phantom. Two different methods were used to predict the SPR from the DECT data, namely the one introduced by Bourque et al. [28] and the one by Lalonde et al. [29]. These were compared with the SPR prediction from the SECT data and then validated with the WER measurement.

For SECT, the uncertainties were 0.53% and 1.37% for soft tissues and bones, respectively. For the method by Bourque et al. [28] the uncertainties were 0.19% and 1.06% and for the method by Lalonde et al. [29] they were 0.38% and 1.06%. It was concluded that there are potential benefits to SPR predictions with DECT rather than SECT.

For the measurements of Möhler et al. [21] , a larger sample container was 3D-printed, consisting of smaller chambers where the different tissues were placed in. These chambers have an inner dimension of $1.5 \times 1.78 \times 1.78\ cm$. One hour before the measurements, the samples were taken out of the fridge to adapt to room temperature. This prevents temperature changes during the measurements. The tissue was cut into exactly fitting pieces to prevent air coming in between the samples. For each sample, a depth-dose curve was measured in a water absorber with variable thickness and a step size of 0.1 mm.

For the CT measurements, slabs were attached on the sides of the containers to make it a cylindrical tube, which was then fitted into a larger PMMA phantom with an inner and outer diameter of 2.8 and 16 cm. The SECT and DECT scans were made at 120 kVp and 80/140 kVp respectively. The tube current was set at respectively 390 mAs and 701/351 mAs, yielding the same $CTDI_{vol}$ of 59.7 mGy. The SPR prediction from the SECT was based on the method stoichiometric calibration by Schneider et al. [11]. For the DECT, the SPR prediction was based on the same method as for SECT, in combination with the use of the syngo.via Rho/Z module by Siemens that derived the ρ_e and Z_{eff}.

Results show that for all tissues except the brain and bone tissue, the SECT SPR predictions deviate significantly from the calculated SPR. With a mean signed error of (-0.84 $\pm$ 0.12)%, there is an overall negative bias. Largest outliers are found for adipose tissue with -2.1% and -2.7%. On the other hand, the DECT SPR predictions do not deviate significantly from the calculated SPR, except for one adipose sample, showing the potential of DECT to decrease range uncertainty for proton therapy.

Taasti et al. [20] also 3D-printed plastic boxes, which were sized $5 \times 5 \times 8\ cm$. CT scans were made before the WER measurements to verify that no air was present in the containers. In addition, some tissues were chopped up and mixed with water (SPR = 1) to prevent air (SPR $\approx$ 0) coming into the samples. Next to tissue samples, a blood sample was made by putting blood in a flask and a solid square bone sample was made by stacking flat cut pieces of bone on each other.

The WER measurements were performed with a 185.2 MeV proton pencil beam. The range was measured with a custom

made MLIC at the 80% distal fall-off. Measurements were performed with empty, water-filled and tissue-filled boxes. The proton beam was set at different points at the boxes to assess the homogeneity of the thickness of the walls of the boxes and also of the tissues. In addition, the boxes were shifted slightly between some measurements, this only led to very little changes in the proton range, which indicates that the samples were homogeneous. The samples were then vacuum sealed and refrigerated until the acquisition of the CT scans, which were performed within three days of the proton beam measurements.

The samples were placed within a torso-shaped water tank to simulate attenuation in a patient. A tube potential of 120 kVp was used with 180 mAs. For the DECT scan, corresponding $CTDI_{vol}$ was used to match the dose. On the DECT data, the syngo.via Rho.Z module was used again to retrieve the ρ_e and the Z_{eff}. Then, to calculate the SPR, The Bethe equation by Schneider et al. [11] was used. Since this equation relies on the mean excitation energy I instead of Z_{eff}, the images were converted using a piecewise linear fit by Yang et al [30].

Six different DECT acquisition methods were used and these all resulted in more accurate SPR predictions than the SECT method. The lowest error was found for the dual-source DECT scans based SPR prediction. These results indicate that DECT can be used to improve proton treatment by reducing uncertainty margins.

Xie et al. [18] placed their animal tissue samples inside acrylic boxes. These were stored frozen at -20 °C. They were taken out of the freezer only just before the WER and CT measurements and then placed in the freezer again limiting time out of freezer to 5 minutes, to ensure accurate repeatability of the measurements. It was acknowledged that the samples were not completely homogeneous. To minimize the effect of the heterogeneity of the samples, a 2.5 cm thick steel collimator was used with an aperture of 2 mm, leading to a spot size after the collimator of 5 mm. The WER was measured using an MLIC by measuring three times at four different beam energies (130, 140, 150, 160 MeV).

The SECT scans were performed at 120 kVp and the DECT scans at 80/140 kVp. From the SECT data, the SPR was predicted using the stoichiometric calibration. From the DECT data, the ρ_e was calculated for each energy spectrum using the Z_{eff}. The SPR was then calculated using the Bethe-Bloch equation as described in Section III-A1.

A Bland-Altman analysis was used to compare the results of the SECT and DECT with the measured SPR data. This resulted in a difference between SECT-based SPR and the measured SPR of $0.55 \pm 1.94\%$ and for DECT-based SPR and the measured SPR of $0.07 \pm 0.58\%$, indicating a better performance by DECT in predicting the SPR and thus a potential to improve treatment planning for proton therapy.

Niepel et al. [23] have designed PMMA phantom shells where the animal tissue samples as well as a container with VIP gel can be placed in. The tissue samples were collected fresh from the butcher a few hours before the measurement. The inner dimension of the sample boxes is $4 \times 4 \times 13.4\ cm$. The boxes were filled with the various tissue samples and then filled up with purified water to prevent air filling up the holes.

The WER measurements were firstly performed using an adjustable water column with peakfinder at 80% dose in the distal fall-off with an energy of 200 MeV. Considering that these measurements are one-dimensional, they also serve as benchmark for the VIP gel measurements. The gel dosimetry offers a three-dimensional result with a high spatial resolution.

Calibration measurements were performed with a phantom with known inserts (Gammex RMI, Sun Nuclear Corporation, Melbourne, FL, USA). The $CTDI_{vol}$ values were not matched for the SECT and the DECT scans, elaborate scanning protocols can be found in the paper. Eight different methods were used to predict the SPR from the dual source DECT data. These all outperformed the SECT method. In addition, it was showed that sequential DECT on a conventional SECT scanner also provided better results over the SECT results, but not as good as the dual source DECT results.

Whereas the previously stated papers investigated whether DECT was suitable as planning CT for proton therapy, Sarkar et al. [17] specifically looked into the use of DECT with the previously explained DirectSPR software tool for treatment planning, comparing it with conventional SECT results. For the SECT analysis, the Bethe-Bloch equation was used to calculate the SPR.

Boxes of $10 \times 10 \times 15\ cm$ were made of perspex. The animal tissues were added and the boxes were filled up with water. The WER was measured using an MLIC at 90% dose in the distal fall-off at an energy of 164.8 MeV. It was stated that difference between the SECT-based SPR predictions and the measurements was 2.2 mm, whereas the difference was only 0.1 mm for the SPR predictions using DirectSPR. This does however seem to be the signed difference. For the SECT-based SPR predictions, the absolute difference is also 2.2 mm, since all predictions were lower than the measured SPR. For the DirectSPR-based predictions on the other hand, the absolute difference is 0.3 mm. It can still be concluded that the DECT-based DirectSPR predictions lead to much better results than the conventional SECT-based SPR predictions.

This section has demonstrated that the methods as explained in Section III have been used to experimentally verify new DECT techniques for SPR prediction. In addition, it has been shown how DECT has improved SPR predictions compared with SECT.

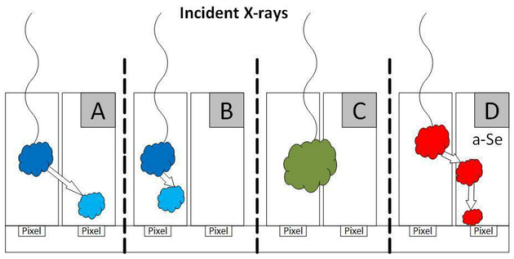


Fig. 6: (A) Photoelectric absorption with k-fluorescence emission escaping target pixel, (B) Photoelectric absorption with k-fluorescence emission in same pixel, (C) charge sharing at pixel boundaries and (D) Compton scatter. From [33]

V. Photon-Counting CT

Firstly, the physics underlying the PCCT technique will be discussed in this section, along with its pros and cons. Subsequently, literature concerning the application of PCCT as a planning CT for proton therapy will be discussed.

A. Physics

A conventional (SECT or DECT) scanner uses an energy-integrating detector (EID), which generates light from the incoming photons with a scintillator. This light is then converted into an electrical signal with photo-diodes. All the incoming pulses are integrated over a certain time interval to generate a signal [31].

A PCCT scanner, on the other hand, uses a photon-counting detector (PCD). This detector is made of a semiconductor that directly converts the incoming photons into electrical pulses. All pulses that surpass the threshold level will be registered as an incoming photon. Additionally, several higher threshold levels can be set to sort the photons in so-called energy bins to acquire spectral information [32]. This allows reconstructing at specific energy levels (e.g. by selecting only photons in the bin around 70 keV), which is the same as the earlier explained VMI.

This new CT technique has its own signal degrading effects. Firstly, a few of those effects can lead to photons ending up in the wrong pixel, which is regarded as crosstalk (Fig. 6). The first two subfigures show photoelectric absorption with k-fluorescence. The incident photons kick out electrons of the detector material, this leads to characteristic X-rays being emitted [31]. Another signal degrading effect is charge sharing, where the charge cloud of an individual photon is shared by two detector pixels, which leads to part of the original energy being registered by two pixels instead of one. In the last subfigure, the effect of Compton scatter can be seen.

In addition, when two photons land on the detector in a very short time span, the second one might not be registered due to pulse pile-up. A detector element has a certain registration and recovery time before it can detect the next photon [32]. Which means that when a photon reaches the detector during this short period, it is not registered. This was a large issue during the development of the PCD. It can not entirely be overcome and should be dealt with in image processing.

B. PCCT for proton therapy

This new technique has shown improvements compared to conventional SECT and DECT in terms of spatial resolution [31] and spectral information. However, little research is available on whether PCCT can function as planning CT to improve SPR prediction accuracy. Increasing the amount of bins is still an issue, since it leads to higher levels of noise per bin and more overlap between the bins. In addition, further research should be performed on improving projection-based corrections for beam hardening and scattering [7].

Taasti et al. [34] performed a theoretical evaluation of a new PCCT-based SPR prediction method. They found a lower root mean square (RMS) error for the PCCT-based method (0.1%) than for SECT-based (0.7%) and DECT-based (0.2%) methods. In addition, measurements were performed on organic tissue samples to compare PCCT-based SPR predictions with SECT-based and DECT-based SPR predictions. The same samples are used as in the study by Taasti et al. [20] that is highlighted in Section IV-B. For these experiments, RMS errors of 0.8% and 1.0% were found for PCCT scans with respectively two and four energy bins.

Saito [35] looked into the conversion from HU to ρ using PCCT. It was found that the optimal reconstruction energy was 70 keV for a tube voltage of 140 kVp, since this yielded the lowest standard deviation and RMS error. However, the expected performance was not met due to the PCCT-specific signal degrading effects as explained in Section V-A. According to Saito, further research is necessary into accounting for these effects before PCCT can improve the conversion from HU to RED.

Simard et al. [36] compared the performance of PCCT (four energy bins) with DECT in a simulation environment to assess its potential to improve treatment planning for radiotherapy. This was done by predicting the electron density and the SPR for proton therapy. A new method for eigentissue decomposition was proposed and compared to conventional 'maximum likelihood material decomposition'. These methods were simulated on an ideal case and a virtual patient with noise and beam-hardening artifacts. For both cases, the RMS error of the PCCT-based predictions was lower than those of DECT-based predictions. Simard et al. concluded that in a realistic simulation environment, the accuracy of proton therapy parameters can be increased when using PCCT with four energy bins.

A year later, Simard et al [37] did PCCT experimental measurements on two human tissue substitute phantoms. The RMS error was reduced from 1.09% to 0.89% when using PCCT instead of DECT. The RMS error of the SPR was reduced from 1.92% to 0.89% when using PCCT instead of DECT. This study highlights the improvements PCCT offers for material decomposition, which is necessary to further improve the accuracy of proton therapy.

Hu et al. [38] performed comparable measurements with Simard et al. [37] on a calibration (Gammex) and a validation (CIRS) phantom with PCCT with four energy bins. Using an optimal pair of VMIs at 60 and 180 keV, the RMS error of the SPR was 1.27% and 0.71% for the calibration and

the validation phantom respectively. This SPR accuracy was comparable to DECT in literature.

Larsson [39] trained a neural network to convert PCCT images to SPR maps. This network was tested on a head phantom and yielded RMS errors of 0.54-1.25%. These results are comparable to those of DECT. However, the author states that the model could be more robust than DECT-based models, since the percentage errors were small while realistic noise levels were used.

VI. Project Outlook

A. Discussion & Conclusion

Section I provided an introduction on proton therapy and highlights the importance of research towards improvements in accuracy of SPR predictions based on different methods and types of CT scanners. Section II explained the used manners of retrieving literature. Section III described the different methods of predicting and measuring the SPR. Furthermore, the necessity of a regularized guide for the HU to SPR conversion was explained. Section IV elaborately discussed literature on the comparison between the performance of DECT and SECT on SPR predictions. Section V introduced PCCT and explains its benefits and drawbacks. In addition, literature is discussed on the potential of PCCT for proton therapy.

Firstly, an important finding is the large variety between proton treatment clinics in workflow and specifically in HLUT definition, as shown by Peters et al. [9]. Therefore, using their guide on a general HLUT definition would be a logical step and should be taken into account in further research.

In addition, Section IV has shown the improvements brought by the use of DECT instead of SECT for SPR predictions. All of these studies indicated that DECT has the potential to improve proton treatment quality by decreasing the error in SPR predictions. In 2015 the first DECT scanner was used in the clinic as planning CT [14]

Another improvement in SPR prediction accuracy in the field has been caused by the introduction of DirectSPR, which uses the spectral information from DECT scans. The studies [17, 27, 40] that used this software to perform SPR predictions unanimously showed improvement in SPR prediction accuracy compared to SECT-based SPR predictions. Since these improvements were larger than the improvements caused by the use of DECT instead of SECT without DirectSPR, it can be concluded that DirectSPR solely can also reduce the error in SPR predictions.

The discussed studies in Section V on the potential of PCCT for proton therapy mainly performed theoretical simulations or phantom measurements. These studies showed improvements by achieving lower error margins on the SPR predictions. Besides this, no case exists to our knowledge where PCCT-based SPR predictions with and without DirectSPR were compared. Lastly, the one study [34] that performed actual animal tissue measurements, did not use DirectSPR.

Combining the insights that PCCT has shown improvements in reducing the errors of the SPR predictions and that DirectSPR has done the same when applied to DECT scans, it can be expected that the combination of PCCT with DirectSPR could reduce the errors of SPR predictions even further. For that reason, the research gap lies in the fact that no (animal) experiments have been performed using PCCT for SPR predictions with use of the DirectSPR software tool. Therefore, the research question is as follows "*To what extend can the implementation of Photon-Counting CT with DirectSPR prediction enhance proton therapy treatment quality by improving the SPR predictions?*

B. Research plan

In order to fill this identified research gap and answer the posed research question, experiments, similar as described in previous literature, will be carried out. This includes imaging of animal tissues with SECT, DECT as well as PCCT. According to these three datasets, SPR predictions will be performed, using DirectSPR prediction methods as well as conventional SPR prediction methods. The animal tissue samples will then be irradiated with a proton beam and a depth-dose curve will be recorded with a multi layer ionization chamber. Using the WER, the SPR of the tissue can be calculated, which validates the SPR predictions.

In addition, the same measurements and DirectSPR predictions will be performed with two standard quality assurance phantoms. since these were not found in literature. However, this lies outside of the scope of this research plan.

Acknowledgment

I would like to thank my HollandPTC supervisor Sophie Huijskens, my Erasmus MC supervisor Marcel van Straten and my TU Delft supervisor Marlies Goorden.

References

[1] Delaney, G., Jacob, S., Featherstone, C., & Barton, M. (2005). The role of radiotherapy in cancer treatment. *Cancer*, *104*(6), 1129–1137.

[2] Paganetti, H., Beltran, C., Both, S., Dong, L., Flanz, J., Furutani, K., Grassberger, C., Grosshans, D. R., Knopf, A.-c., Langendijk, J. A., Nystrom, H., Parodi, K., Raaymakers, B. W., Richter, C., Sawakuchi, G. O., Schippers, M., Shaitelman, S. F., Teo, B. K. K., Unkelbach, J., ... Lomax, T. (2021). Roadmap : proton therapy physics and biology. *Physics in Medicine & Biology*, *66*, 1–61.

[3] Radiology Key. (2017). Particle Therapy in the Third Millennium: Current Status and Future Outlook. https://radiologykey.com/particle-therapy-in-the-third-millennium-current-status-and-future-outlook/

[4] Taasti, V. T., Bäumer, C., Dahlgren, C. V., Deisher, A. J., Ellerbrock, M., Free, J., Gora, J., Kozera, A., Lomax, A. J., De Marzi, L., Molinelli, S., Kevin Teo, B. K., Wohlfahrt, P., Petersen, J. B., Muren, L. P., Hansen, D. C., & Richter, C. (2018). Inter-centre variability of CT-based stopping-power prediction in particle therapy: Survey-based evaluation. *Physics and Imaging in Radiation Oncology*, *6*, 25–30.

[5] Goitein, M. (1985). Calculation of the uncertainty in the dose delivered during radiation therapy. *Medical Physics*, *12*(5), 608–612.

[6] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, *372*(71).

[7] Wohlfahrt, P., & Richter, C. (2020). Proton therapy special feature: Review article status and innovations in pre-treatment CT imaging for proton therapy. *British Journal of Radiology*, *93*(July), 1–14.

[8] Salvat, F. (2022). PHYSICAL REVIEW A 106 , 032809 (2022) Bethe stopping-power formula and its corrections. *American Physical Society*, *106*.

[9] Peters, N., Trier Taasti, V., Ackermann, B., Bolsi, A., Vallhagen Dahlgren, C., Ellerbrock, M., Fracchiolla, F., Gomà, C., Góra, J., Cambraia Lopes, P., Rinaldi, I., Salvo, K., Sojat Tarp, I., Vai, A., Bortfeld, T., Lomax, A., Richter, C., & Wohlfahrt, P. (2023). Consensus guide on CT-based prediction of stopping-power ratio using a Hounsfield look-up table for proton therapy. *Radiotherapy and Oncology*, *184*, 109675.

[10] McGarry, C. K., Grattan, L. J., Ivory, A. M., Leek, F., Liney, G. P., Liu, Y., Miloro, P., Rai, R., Robinson, A. P., Shih, A. J., Zeqiri, B., & Clark, C. H. (2020). Tissue mimicking materials for imaging and therapy phantoms: a review. *Physics in Medicine & Biology*, *65*(23), 23TR01.

[11] Schneider, U., Pedroni, E., & Lomax, A. (1996). The calibration of CT Hounsfield units for radiotherapy treatment planning. *Physics in Medicine & Biology*, *41*, 111–124.

[12] Wohlfahrt, P., Möhler, C., Greilich, S., & Richter, C. (2017). Comment on: Dosimetric comparison of stopping-power calibration with dual-energy CT and single-energy CT in proton therapy treatment planning [Med. Phys. 43(6), 2845-2854 (2016)]. *Medical Physics*, *44*(10), 5533–5536.

[13] Goitein, M. (1977). The measurement of tissue heterodensity to guide charged particle radiotherapy. *International Journal of Radiation Oncology, Biology, Physics*, *3*, 27–33.

[14] Wohlfahrt, P., Möhler, C., Hietschold, V., Menkel, S., Greilich, S., Krause, M., Baumann, M., Enghardt, W., & Richter, C. (2017). Clinical Implementation of Dual-energy CT for Proton Treatment Planning on Pseudo-monoenergetic CT scans. *International Journal of Radiation Oncology*Biology*Physics*, *97*(2), 427–434.

[15] Peters, N., Wohlfahrt, P., Dahlgren, C. V., de Marzi, L., Ellerbrock, M., Fracchiolla, F., Free, J., Gomà, C., Góra, J., Jensen, M. F., Kajdrowicz, T., Mackay, R., Molinelli, S., Rinaldi, I., Rompokos, V., Siewert, D., van der Tol, P., Vermeren, X., Nyström, H., ... Richter, C. (2021). Experimental assessment of inter-centre variation in stopping-power and range prediction in particle therapy. *Radiotherapy and Oncology*, *163*, 7–13.

[16] Bär, E., Lalonde, A., Zhang, R., Jee, K. W., Yang, K., Sharp, G., Liu, B., Royle, G., Bouchard, H., & Lu, H. M. (2018). Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples. *Medical Physics*, *45*(1), 48–59.

[17] Sarkar, V., Paxton, A., Su, F., Price, R., Nelson, G., Szegedi, M., James, S. S., & Salter, B. J. (2023). An evaluation of the use of DirectSPR images for proton planning in the RayStation treatment planning software. *Journal of Applied Clinical Medical Physics*, *24*(5).

[18] Xie, Y., Ainsley, C., Yin, L., Zou, W., McDonough, J., Solberg, T. D., Lin, A., & Teo, B. K. K. (2018). Ex vivo validation of a stoichiometric dual energy CT proton stopping power ratio calibration. *Physics in Medicine & Biology*, *63*(5), 055016.

[19] IBA Dosimetry. (n.d.). IBA Dosimetry: Zebra. https://www.iba-dosimetry.com/product/zebra

[20] Taasti, V. T., Michalak, G. J., Hansen, D. C., Deisher, A. J., Kruse, J. J., Krauss, B., Muren, L. P., Petersen, J. B., & McCollough, C. H. (2018). Validation of proton stopping power ratio estimation based on dual energy CT using fresh tissue samples. *Physics in Medicine & Biology*, *63*(1), 1–12.

[21] Möhler, C., Russ, T., Wohlfahrt, P., Elter, A., Runz, A., Richter, C., & Greilich, S. (2018). Experimental verification of stopping-power prediction from single- and dual-energy computed tomography in biological tissues. *Physics in Medicine & Biology*, *63*, 025001.

[22] Wohlfahrt, P. (2018). *Dual-Energy Computed Tomography for Accurate Stopping-Power Prediction in Proton*

Treatment Planning [Doctoral dissertation]. http://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa2-317554

[23] Niepel, K. B., Stanislawski, M., Wuerl, M., Doerringer, F., Pinto, M., Dietrich, O., Ertl-Wagner, B., Lalonde, A., Bouchard, H., Pappas, E., Yohannes, I., Hillbrand, M., Landry, G., & Parodi, K. (2021). Animal tissue-based quantitative comparison of dual-energy CT to SPR conversion methods using high-resolution gel dosimetry. *Physics in Medicine & Biology Biology*, *66*(7), 075009.

[24] Wohlfahrt, P., Möhler, C., Stützer, K., Greilich, S., & Richter, C. (2017). Dual-energy CT based proton range prediction in head and pelvic tumor patients. *Radiotherapy and Oncology*, *125*(3), 526–533.

[25] Taasti, V. T., Muren, L. P., Jensen, K., Petersen, J. B. B., Thygesen, J., Tietze, A., Grau, C., & Hansen, D. C. (2018). Comparison of single and dual energy CT for stopping power determination in proton therapy of head and neck cancer. *Physics and Imaging in Radiation Oncology*, *6*, 14–19.

[26] Polf, J. C., Mille, M. M., Mossahebi, S., Chen, H., Maggi, P., & Chen-Mayer, H. (2019). Determination of proton stopping power ratio with dual-energy CT in 3D-printed tissue/air cavity surrogates. *Medical Physics*, *46*(7), 3245–3253.

[27] Peters, N., Wohlfahrt, P., Hofmann, C., Möhler, C., Menkel, S., Tschiche, M., Krause, M., Troost, E. G., Enghardt, W., & Richter, C. (2022). Reduction of clinical safety margins in proton therapy enabled by the clinical implementation of dual-energy CT for direct stopping-power prediction. *Radiotherapy and Oncology*, *166*, 71–78.

[28] Bourque, A. E., Carrier, J. F., & Bouchard, H. (2014). A stoichiometric calibration method for dual energy computed tomography. *Physics in Medicine & Biology*, *59*, 2059–2088.

[29] Lalonde, A., Bär, E., & Bouchard, H. (2017). A Bayesian approach to solve proton stopping powers from noisy multi-energy CT data. *Medical physics*, *44*(10), 5293–5302.

[30] Yang, M., Virshup, G., Clayton, J., Zhu, X. R., Mohan, R., & Dong, L. (2010). Theoretical variance analysis of single- and dual-energy computed tomography methods for calculating proton stopping power ratios of biological tissues. *Physics in Medicine & Biology*, *55*, 1343–1362.

[31] Flohr, T., Petersilka, M., Henning, A., Ulzheimer, S., Ferda, J., & Schmidt, B. (2020). Photon-counting CT review. *Physica Medica*, *79*, 126–136.

[32] Willemink, M. J., Persson, M., Pourmorteza, A., Pelc, N. J., & Fleischmann, D. (2018). Photon-counting CT: Technical principles and clinical prospects. *Radiology*, *289*(2), 293–312.

[33] Stavro, J., Goldan, A. H., & Zhao, W. (2016). SWAD: inherent photon counting performance of amorphous selenium multi-well avalanche detector. *Medical Imaging 2016: Physics of Medical Imaging*, *9783*.

[34] Taasti, V. T., Hansen, D. C., Michalak, G. J., Deisher, A. J., Kruse, J. J., Muren, L. P., Petersen, J. B., & McCollough, C. H. (2018). Theoretical and experimental analysis of photon counting detector CT for proton stopping power prediction. *Medical Physics*, *45*(11), 5186–5196.

[35] Saito, M. (2019). Simulation of photon-counting detectors for conversion of dual-energy-subtracted computed tomography number to electron density. *Radiological Physics and Technology*, *12*(1), 105–117.

[36] Simard, M., Lapointe, A., Lalonde, A., Bahig, H., & Bouchard, H. (2019). The potential of photon-counting CT for quantitative contrast-enhanced imaging in radiotherapy. *Physics in Medicine & Biology*, *64*, 115020.

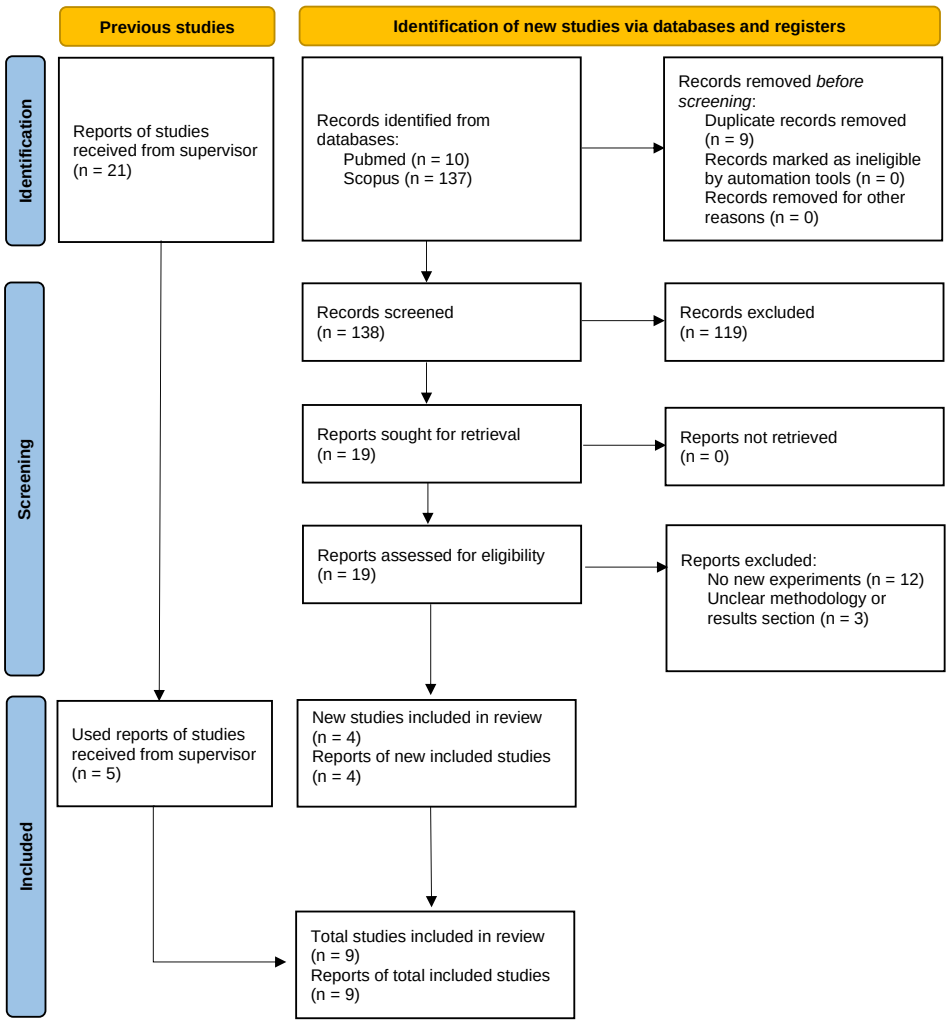
[37] Simard, M., Panta, R. K., Bell, S. T., Butler, A. P., & Bouchard, H. (2020). Quantitative imaging performance of MARS spectral photon-counting CT for radiotherapy. *Medical Physics*, *47*(8), 3423–3434.

[38] Hu, G., Niepel, K., Risch, F., Kurz, C., Würl, M., Kröncke, T., Schwarz, F., Parodi, K., & Landry, G. (2022). Assessment of quantitative information for radiation therapy at a first-generation clinical photon-counting computed tomography scanner. *Frontiers in Oncology*, *12*(September), 1–10.

[39] Larsson, K. (2023). *Improving proton therapy planning with photon-counting spectral computed tomography* [Doctoral dissertation].

[40] Longarino, F. K., Herpel, C., Tessonnier, T., Mein, S., Ackermann, B., Debus, J., Schwindling, F. S., Stiller, W., & Mairani, A. (2023). Dual-energy CT-based stopping power prediction for dental materials in particle therapy. *Journal of Applied Clinical Medical Physics*, *24*(8), 1–16.

APPENDIX A
PRISMA FLOW CHART



APPENDIX B
OVERVIEW OF EXPERIMENTAL RESEARCH ON DECT FOR SPR PREDICTIONS WITH ANIMAL TISSUES

TABLE I: Overview of experimental research on DECT for SPR predictions with animal tissues

	Animal tissue: Beef	Animal tissue: Pork	Proton beam energy	Proton beam measurement	SPR prediction methods	Results
Bar et al.	Adipose, bone	Adipose, muscle, liver, kidney, brain, heart, blood and lung	195 MeV	Dose extinction method	SECT: HLUT DECT: 1. Bethe-Bloch (Rho and I-value) 2. Bethe-Bloch (elemental compositions)	RMS error for soft tissues and bones: SECT: 0.53% and 1. 37% DECT: 1. 0.19% and 1.06% 2. 0.38% and 1.06%
Mohler et al.	Bone marrow, cortical bone	Back fat (smoked), belly fat (fresh), loin (fresh), fillet (fresh), thick flank (fresh), liver (fresh), kidney (fresh), brain (fresh), heart (fresh), blood (fresh) and lung (fresh)	200.28 MeV	Water absorber of variable thickness	SECT: HLUT DECT: syngo Rho/Z software	Signed error: SECT: 0.84% +- 0.12% DECT: No significant error
Taasti et al.	Ground beef, kidney (chopped), suet, round (chopped and unchopped) and femur	Loin (chopped and unchopped), liver (chopped), fat (chopped and unchopped), brain (fresh), brain (frozen, thawed) and blood (mixed with heparin)	185.2 MeV	MLIC	SECT: HLUT DECT: syngo Rho/Z software	RMS error: SECT: 2.78% DECT: 0.47%
Xie et al.	Shank and muscle	Brain, kidney and liver	130, 140, 150, 160 MeV	MLIC	SECT: HLUT DECT: Bethe-Bloch (Rho and I-value)	Error: SECT: 0.55 +-1.94% DECT: 0.07 +- 0.58%
Niepel et al.	Shoulder, tongue, liver, filet and fat	Brain	200 MeV	Water absorber of variable thickness	SECT: HLUT DECT: 1. Bethe-Bloch (Rho and I-value) 2. Bethe-Bloch (elemental composition) 3. Direct conversion	Error for muscle, adipose, bone200 and cortical bone SECT: -2.51%, -0.58%, 3.72% and -0.29% DECT: 3. direct conversion: -1.21%, -0.81%, -0.37% and 1.05%
Sarkar et al.	Grounded fat, grounded heart and grounded sirloin	-	164.8 MeV	MLIC	SECT: HLUT DECT: DirectSPR	Signed and absolute difference with WER: SECT: 2.2 mm and 2.2 mm DECT: 0.1 mm and 0.3 mm