Theoretical and experimental analysis of photon counting detector CT for proton stopping power prediction

Vicki T. Taasti^{a)} and David C. Hansen

Department of Medical Physics, Aarhus University Hospital, Aarhus, Denmark

Gregory J. Michalak

Department of Radiology, Mayo Clinic, Rochester, MN, USA

Amanda J. Deisher and Jon J. Kruse

Department of Radiation Oncology, Mayo Clinic, Rochester, MN, USA

Ludvig P. Muren and Jørgen B. B. Petersen

Department of Medical Physics, Aarhus University Hospital, Aarhus, Denmark

Cynthia H. McCollough

Department of Radiology, Mayo Clinic, Rochester, MN, USA

(Received 14 November 2017; revised 25 July 2018; accepted for publication 31 August 2018; published 1 October 2018)

Purpose: Photon counting detectors (PCDs) are being introduced in advanced x-ray computed tomography (CT) scanners. From a single PCD-CT acquisition, multiple images can be reconstructed, each based on only a part of the original x-ray spectrum. In this study, we investigated whether PCD-CT can be used to estimate stopping power ratios (SPRs) for proton therapy treatment planning, both by comparing to other SPR methods proposed for single energy CT (SECT) and dual energy CT (DECT) as well as to experimental measurements.

Methods: A previously developed DECT-based SPR estimation method was adapted to PCD-CT data, by adjusting the estimation equations to allow for more energy spectra. The method was calibrated directly on noisy data to increase the robustness toward image noise. The new PCD SPR estimation method was tested in theoretical calculations as well as in an experimental setup, using both four and two energy bin PCD-CT images, and through comparison to two other SPR methods proposed for SECT and DECT. These two methods were also evaluated on PCD-CT images, full spectrum (one-bin) or two-bin images, respectively. In a theoretical framework, we evaluated the effect of patient-specific tissue variations (density and elemental composition) and image noise on the SPR accuracy; the latter effect was assessed by applying three different noise levels (low, medium, and high noise). SPR estimates derived using real PCD-CT images were compared to experimentally measured SPRs in nine organic tissue samples, including fat, muscle, and bone tissues.

Results: For the theoretical calculations, the root-mean-square error (RMSE) of the SPR estimation was 0.1% for the new PCD method using both two and four energy bins, compared to 0.2% and 0.7% for the DECT- and SECT-based method, respectively. The PCD method was found to be very robust toward CT image noise, with a RMSE of 2.7% when high noise was added to the CT numbers. Introducing tissue variations, the RMSE only increased to 0.5%; even when adding high image noise to the changed tissues, the RMSE stayed within 3.1%. In the experimental measurements, the RMSE over the nine tissue samples was 0.8% when using two energy bins, and 1.0% for the four-bin images.

Conclusions: In all tested cases, the new PCD method produced similar or better results than the SECT- and DECT-based methods, showing an overall improvement of the SPR accuracy. This study thus demonstrated that PCD-CT scans will be a qualified candidate for SPR estimations. © 2018 American Association of Physicists in Medicine [<https://doi.org/10.1002/mp.13173>]

Key words: CT noise, experimental verification, noise robustness, photon counting detector CT, proton stopping power ratio

1. INTRODUCTION

Several recent studies have investigated how to improve stopping power ratio (SPR) estimations with the aim of improving the accuracy of CT-based dose calculations in proton therapy. These studies include investigations on how to

perform a robust calibration of the conversion from CT numbers in conventional single-energy CT (SECT) to SPR,¹ the influence of beam hardening (BH),² the use of dual-energy CT (DECT),³ as well as the use of proton CT.^{4–6} It has been shown in several studies^{7–19} that DECT gives more accurate SPR estimates than SECT using the stoichiometric method

proposed by Schneider et al.¹⁹. In spite of that, concerns have been raised that DECT is more influenced by image noise than SECT.²⁰

Photon counting detectors (PCDs) can improve image quality and spatial resolution compared to energy-integrating detectors (EIDs) used in conventional SECT and DECT scanners.²¹ Furthermore, PCDs have the ability to provide energy information, which allows for obtaining multiple CT images at different effective energies in a single CT acquisition.²² Ideally, the energy bins for PCDs would be fully separated without spectral overlap. However, in real PCDs, this is not the case due to detrimental physical effects such as charge sharing, K-edge escape, and pulse pileup.²³

Lalonde and Bouchard,²⁴ Lalonde et al.²⁵, and Shen et al.²⁶ have recently proposed new SPR methods for PCD-CT. Based on computer simulations or multiple regular SECT scans, all three studies found that the root-mean-square error (RMSE) for SPR estimation decreased when the number of energy bins increased.^{24–26}

In a previous study, we proposed a DECT-based SPR method which was constructed to be robust toward noise.²⁷ In the present study, we update this method to be applicable to PCD-CT images. This new method can be used for varying numbers of energy bins or energy spectra; it can thus also be applied for regular DECT images. The aim of this study was to explore the adaptation of this method to PCD-CT data and also to compare it to established SECT- and DECT-based SPR methods, in theoretical as well as experimental evaluations. These SECT- and DECT-based SPR methods will be evaluated on PCD-CT images using either the full energy spectrum (one-bin) images or the two-bin images.

2. MATERIALS AND METHODS

2.A. PCD-CT scanner system

In this study, we used the Siemens SOMATOM CounT research PCD-CT system (Siemens Healthineers, Forchheim, Germany).^{28,29} This PCD-CT scanner can acquire either four-bin or two-bin PCD-CT images, which are denoted as Chess Mode and Macro Mode, due to the subpixel arrangements of the detector (see fig. 1B of Ref. [28]). The x-ray tube of this CT scanner is the same as in the Siemens SOMATOM Definition Flash CT scanner. In this study, we used the 140 kVp x-ray spectrum without extra tin filtration, since this provided the broadest x-ray energy spectrum, whereby the energy bin thresholds could be more widely spread. We investigated both the Macro Mode (two energy bins) and the Chess Mode (four energy bins); the same computer code was used to produce both sets of results.

All CT scans/CT number computations used in this study were based on this PCD-CT scanner system. The SECT- and DECT-based SPR methods used in the comparisons were therefore based on Macro Mode acquisitions, reconstructed either based on the full energy spectrum (effectively one energy bin) or the two energy bins.

2.B. SPR estimation method

The PCD-CT-based SPR estimation method proposed in this study was based on the DECT-based SPR estimation method proposed by Taasti et al.²⁷ In the present study, the DECT-based SPR method was adapted to be used for multi-energy CT images with different numbers of energy bins, including two energy bins or two full energy spectra as in regular DECT imaging. This new method will be denoted the PCD method.

As the method by Taasti et al.,²⁷ this method is purely empirical. It was based on a Taylor expansion for the CT numbers and with no theoretical assumptions. However, even more freedom was given in this method, with the cost of more fitting parameters as the number of fitting parameters was allowed to increase with the number of energy bins. The hypothesis was that the more fitting parameters, the better the calibration could account for differences in the tissues. More fitting parameters were used for soft tissues, since more variation was found among this tissue group. The method was further updated by adding an extra equation specifically for lung tissues. The three SPR estimation equations we propose are:

$$\text{SPR}_{\text{est}}^{\text{lung}} = \sum_{i=1}^{N_{\text{bin}}} u_i \cdot x_i^{\text{lung}} \quad (1a)$$

$$\begin{aligned} \text{SPR}_{\text{est}}^{\text{soft}} = & \sum_{i=1}^{N_{\text{bin}}} \left(u_i \cdot x_i^{\text{soft}} + u_i^2 \cdot x_{i+N_{\text{bin}}}^{\text{soft}} \right) \\ & + x_{2 \cdot N_{\text{bin}}+1}^{\text{soft}} \cdot \sum_{i=1}^{N_{\text{bin}}} u_i^3 + x_{2 \cdot N_{\text{bin}}+2}^{\text{soft}} \end{aligned} \quad (1b)$$

$$\begin{aligned} \text{SPR}_{\text{est}}^{\text{bone}} = & \sum_{i=1}^{N_{\text{bin}}} \left(u_i \cdot x_i^{\text{bone}} + u_i^2 \cdot x_{i+N_{\text{bin}}}^{\text{bone}} \right) \\ & + x_{2 \cdot N_{\text{bin}}+1}^{\text{bone}} \end{aligned} \quad (1c)$$

Here, N_{bin} is the number of energy bins (in this study, $N_{\text{bin}} = 4$ for Chess Mode and $N_{\text{bin}} = 2$ for Macro Mode). The x_j -parameters are the fitting parameters, and u_i is the so-called reduced CT number for energy bin i , to be defined in Eq. (4) below.

2.C. Calibration procedure

The calibration of this method is a three-step process. First, the effective energies for the used energy spectra were found based on CT scans of a calibration phantom with known elemental composition and density. These effective energies were then used to calculate CT numbers for a set of reference human tissues, denoted calibration tissues. Second, the calibration tissues were categorized as lung, soft tissue, or bone and a nearest neighbor classification model was fitted. Third, the fitting parameters for the SPR estimation equations [Eq. (1)] were found based on the calibration tissues.

FIG. 1. Detected energy spectra used for the theoretical calculations. (These spectra were not used in the experimental evaluation, and we do not assume the actual spectra used in the experimental evaluation to be the same as these theoretical energy spectra). (a) The four energy bin spectra for the Chess Mode. (b) The two energy bin spectra used for Macro mode. The energy thresholds for each energy bin are shown in the figure legends, along with the mean energy, E_m , and the fraction of detected photons (relative to the photons emitted from the x-ray source) for each energy bin. [Color figure can be viewed at wileyonlinelibrary.com]

The Gammex Cone-Beam Electron Density Phantom (Gammex, Inc., Middleton, WI) was used as calibration phantom in this study. This phantom is an updated version of the Gammex 467 CT electron density phantom. It contained 12 insert materials ranging from low-density lung to cortical bone; seven inserts were categorized as soft tissues and five as bone tissues. Specifications on the density and the elemental weight fractions were provided by the manufacturer.

As calibration tissues, we used the 71 reference human tissues from the Woodard and White publications.^{30,31} To this list, we added five extra lung tissues of density 0.20, 0.30, 0.35, 0.40, and 0.45 g/cm³ to cover inhale to exhale lung. These extra lung tissues had the same elemental composition as the original lung tissue. We also added teeth as described in ICRP Report 110.³² In total, 77 tissues were used for the calibration.

2.C.1. Step 1: CT number estimation

Theoretical CT number calculation — for evaluation tissues: For the CT number calculation in the theoretical evaluations, we used the measured 140 kVp x-ray source energy spectrum of the Siemens SOMATOM Definition Flash and the detector response for the SOMATOM Count PCD scanner provided by Siemens Healthineers. The x-ray energy spectrum was attenuated by 25 cm of water to simulate beam hardening through a patient and weighted by the detector response. The energy bin configuration was chosen to obtain an equal number of detected photons in each energy bin. The resulting detected energy spectra, $S_{E,i}$, for each energy bin can be seen in Fig. 1.

The CT numbers, $\mathcal{H}_i$, were calculated theoretically:

$$\mathcal{H}_i^{\text{theo}} = \left(\frac{\langle \mu \rangle_i}{\langle \mu_w \rangle_i} - 1 \right) \cdot 1000 \text{ HU} \quad (2)$$

where μ and μ_w are the linear attenuation values for tissue and water, respectively. Here, $\langle \dots \rangle$ denotes averaging over the x-ray energy spectrum and the detector response function, given as:

$$\langle \mu \rangle_i = \rho \sum_E \left(\sum_l \omega_l \left(\frac{\mu}{\rho} \right)_{l,E} \right) S_{E,i} \quad (3)$$

here, ρ is the mass density, and ω_l is the weight fraction for element l in the material. The mass attenuation coefficients, $(\mu/\rho)_{l,E}$, for each energy, E , in steps of 1 keV, were taken from the XCOM database.³³

CT number estimation: The CT numbers for the calibration tissues were calculated based on the effective energy for each energy bin. Two separate effective energies were found for each energy bin, one for soft (and lung) tissues and one for bone tissues. For each effective energy, a set of linear fitting parameters (A and B) was found using linear least squares (see also Refs. [27,34]). The effective energy was defined as the energy which maximized the coefficient of determination, R^2 , for the last part of the following equation:

$$u_i \equiv \frac{\mu(E_{i,\text{eff}}^{tg})}{\mu_w(E_{i,\text{eff}}^{tg})} = \frac{\mathcal{H}_i - B_i^{tg}}{A_i^{tg}} \quad (4)$$

$E_{i,\text{eff}}^{tg}$ is the effective energy for each tissue group, tg , that is, for soft and bone tissues, and A_i^{tg} and B_i^{tg} are the corresponding fitting parameters. This fit was based on the CT numbers for the Gammex phantom (in the theoretical evaluation, the CT numbers ($\mathcal{H}_i$) for the Gammex materials were calculated theoretically as described above).

To find $E_{i,\text{eff}}^{tg}$, we simply tested all energies in a range from 30 to 160 keV in steps of 1 keV to find the best fit. That is, $E_{i,\text{eff}}^{tg}$ was chosen as the energy for which R^2 was largest. Moreover, the linear attenuation coefficient was again calculated based on the XCOM database, $\mu_{\text{insert}}(E) = \rho_{\text{insert}} \sum_l \omega_l (\mu/\rho)_{l,E}$.

The accuracy of this CT numbers estimation was investigated both in theoretical and the experimental evaluation. In the theoretical evaluation, the RMSE for the calibration

tissues was between 0.004% and 0.024% for each energy bin in Chess and Macro mode. In the experimental evaluation, the CT number accuracy was evaluated for the Gammex inserts, see Table S1.I in Data S1.

2.C.2. Step 2: tissue classification

The calibration tissues were categorized as lung, soft, or bone tissues based on their density and calcium content. Lung tissues were defined as tissues with a mass density below 0.6 g/cm³ and an elemental weight fraction of calcium below 1.5%, soft tissues were defined as tissues with a mass density above 0.6 g/cm³ and a calcium content below 1.5%, while bone tissues were defined as all tissues with a calcium content above 1.5%.

This tissue categorization resulted in six lung tissues for the calibration of Eq. (1a), 46 soft tissues for the calibration of Eq. (1b), and 25 bone tissues for the calibration of Eq. (1c).

The tissue categorization of unknown materials was based on the CT numbers for the material using a nearest neighbor classification³⁵ with three neighbors. All calculations in this study were performed using MATLAB (The MathWorks Inc., Natick, MA), and this model was fitted using the *fitcknn* function. The model was fitted based on the CT numbers for either the calibration tissues or the calibration tissues with added noise (see Section 2.E.2 below), and in both cases supplemented by tissues with applied density or elemental composition variations (Section 2.E.3).

2.C.3. Step 3: Calibration of the SPR estimation equations

To find the fitting parameters $x = \{x_j^{tg}\}$ in Eq. (1), a regularized least-squares optimization was performed using the *lsqnonlin* function on the following equation:

$$\min_x \left(\sum_{j=1}^{N_{tg}} r(x)^2 + \lambda \cdot \sum_{i=1}^{N_{x,tg}} x_i^2 \right) \quad (5)$$

Here, N_{tg} is the number of calibration tissues in tg , and $N_{x,tg}$ is the number of fitting parameters for tg while r is a residual term for the estimated SPR values. For lung tissues, an absolute residual term was used, $r(x) = \text{SPR}_{\text{est},j}^{tg}(x) - \text{SPR}_{\text{theo},j}$, while a relative residual term was used for soft and bone tissues, $r(x) = (\text{SPR}_{\text{est},j}^{tg}(x) - \text{SPR}_{\text{theo},j}) / \text{SPR}_{\text{theo},j}$. λ is a regularization parameter which was used to enhance stability of the fit; we used $\lambda = 10^{-2}$.

To increase the robustness of the method toward CT image noise, the method was also calibrated on noisy data, as suggested by Sietsma and Dow.³⁶ Noise was randomly sampled from a normal distribution and added to the CT numbers for the 77 calibration tissues. Both the noise-free CT numbers and 10,000 noise realizations of each tissue were included in the calibration; in total, 770,077 tissues were used for these calibrations. A large number of noise realizations were used to ensure low variability for the fitting parameters between different runs of the calibration process.

2.D. SPR estimation for unknown materials

In the SPR estimation process for an unknown material, the material was first categorized as lung, soft tissue, or bone tissue applying the nearest neighbor classification. After tissue classification, the reduced CT numbers, u_i , were calculated from Eq. (4) based on the CT number, $\mathcal{H}_i$, for the material, and lastly, the SPR estimate was calculated from Eqs. (1a), (1b) or (1c). Any negative SPR estimate was set to zero.

2.E. Theoretical evaluation of the SPR estimation method

Only the original list of 71 tissues was used for theoretical evaluation of the SPR estimation, with the CT numbers calculated theoretically using Eqs. (2) and (3).

All results were compared to both a SPR method proposed for DECT by Hünemohr *et al.*,⁸ and the stoichiometric method proposed for SECT by Schneider *et al.*¹⁹ The stoichiometric method is considered a reference SPR method, since it is used clinically by many proton centers for treatment planning dose calculations.³⁷ The implementation of the Hünemohr and the stoichiometric method is described in Data S1.

2.E.1. Evaluation of the calibration fit

The SPR methods were first tested on the 71 standard evaluation tissues (with no tissue changes or image noise added) to evaluate the consistency of the new PCD-CT-based method, and to test if the Hünemohr and the stoichiometric methods were correctly implemented.

2.E.2. Image noise evaluation

To test the robustness of the PCD method toward CT image noise, noise sampled from a normal distribution was added to the theoretical CT numbers of the evaluation tissues. For each tissue, 5,000 noise realizations were added, resulting in 355,000 tissues used in the evaluation process.

The standard deviation of the image noise distribution for each energy bin depended on the number of energy bins, $\sigma_{\text{bin}} = \sqrt{N_{\text{bin}}} \cdot \sigma_{FS}$, as in Lalonde *et al.*²⁵ Here, subscript *FS* denotes the full spectrum. This was done to simulate the same total dose for each CT configuration (full spectrum (one-bin), two-bin, and four-bin). We used three different noise levels $\sigma_{FS} = [10, 20, 40]$ HU. For the high noise level, this resulted in a standard deviation for the four-bin images of 80 HU, which was comparable to what was seen for real PCD-CT images.

For the calibration tissues, the standard deviation for the added noise was set to half the standard deviation used for the evaluation tissues. The fitting parameters used for the lung tissues [in Eq. (1)] were though always obtained from the noise-free calibration (see Data S1, Section S1).

2.E.3. Tissue variations

The performance of the SPR estimation for tissues differing from the calibration tissues was also tested by applying three tissue variations to the original evaluation tissues. For all tissues, the density was changed; for soft tissues (including lung tissues), the elemental weight fraction of hydrogen was changed relative to the carbon or the oxygen fraction, and for bone tissues, the calcium fraction was changed relative to the carbon or the oxygen fraction, similar to previous studies.^{7,27,34} The percentage of change was randomly sampled from a normal distribution with a standard deviation of 2%. For each type of variation, 1,000 changes were applied, giving in total 213,000 tissues for this evaluation. For this evaluation, the fitting parameters from the noise-free calibration were used.

In a separate evaluation, noise was added to these changed tissues. The same three noise levels, as applied in the noise evaluation above, were used. For this evaluation, the noisy fitting parameters were used, but as before, for lung tissues, the noise-free fitting parameters were used. In this evaluation, 20 noise realizations were added to each changed evaluation tissue.

2.F. Experimental evaluation based on real PCD-CT scans of organic tissues

To evaluate the new PCD method experimentally, we used the proton pencil beam measurement data for organic tissues described in Taasti et al.¹². PCD-CT scans were acquired of nine of the tissue samples 1 or 2 days after the proton measurements. The tissue samples included seven soft tissues contained in $5 \times 5 \times 8 \text{ cm}^3$ 3D-printed plastic boxes, blood contained in a flask, and a beef femur cut into solid block, all preparation details can be found in Taasti et al.¹². The tissue samples were submerged into a 35 cm water tank to simulate a patient geometry, one at a time, and CT scanned at 140 kVp. They were scanned using both Chess Mode with a threshold configuration of $E_{\text{th}} = [25, 45, 66, 84] \text{ keV}$, and Macro Mode, $E_{\text{th}} = [25, 65] \text{ keV}$. For both scan settings, we used a dose of $\text{CTDI}_{\text{vol},32\text{cm}} = 16.3 \text{ mGy}$ (except for the blood and bone sample where $\text{CTDI}_{\text{vol},32\text{cm}} = 30.8 \text{ mGy}$). The SECT-based SPR estimation was performed on the full spectrum PCD-CT scans, denoted threshold low images,²⁸ based on the Macro Mode acquisition to have the highest dose efficiency, (see Section 4 and Ref. ²⁸). The DECT-based method proposed by Hünemohr et al. was performed on Macro Mode PCD-CT scans. All PCD-CT images acquired in this study were reconstructed with a D34 kernel, with bone beam hardening correction (BHC).

The PCD-CT scanner had a field of view (FOV) of 27.5 cm; therefore, the entire 35 cm water tank did not fit in the FOV. To avoid truncation artifacts, a data completion scan (DCS) was acquired using the EID subsystem.²¹ The DCS scan had a FOV of 50 cm, which was large enough to encompass the water tank, and no artifacts were seen in the reconstructed PCD-CT images.

PCD-CT scans were acquired of the Gammex phantom using the same scan mode and threshold configurations as for the tissue samples. These scans were used to fit Eq. (4) to find $E_{i,\text{eff}}^{\text{tg}}$, A_i^{tg} and B_i^{tg} . To ensure the same BH effect as for the water tanks, the entire torso-shaped phantom was scanned (lateral dimension = 40 cm), but only the CT numbers for the inserts placed in the inner head-sized part of the phantom (diameter = 20 cm) were used in the calibration of the effective energies.

The fitting parameters and effective energies used in the experimental evaluation are listed in the Data S1, Tables S1.I, and S1.II.

The SPR estimates for the tissue samples were calculated pixel-by-pixel from Eq. (1) and the mean and standard deviation of the SPR were calculated over a square volume-of-interest (VOI) well within the tissue samples, covering multiple image slices. VOI sizes are stated in Table S1.III in Data S1. Most tissues were found to be very homogeneous.¹² The SPR estimates were compared to measured SPR values, which were taken as the ground truth in this evaluation. The SPR values were measured in proton pencil beams with an energy of 185.2 MeV using a multilayer ionization chamber, further details on the experimental setup and uncertainty calculations for the SPR measurements were described by Taasti et al.¹²

The noise added to the calibration tissues for the fitting procedure of Eq. (1) used for the tissue samples was based on half of the average of standard deviations for the CT numbers over the VOIs in the tissue samples' PCD-CT scans. For the four-bin images, the average noise level was $\bar{\sigma}(\text{HU}) = [112, 111, 102, 80] \text{ HU}$, and for two-bin images, it was $\bar{\sigma}(\text{HU}) = [41, 41] \text{ HU}$. $\bar{\sigma}$ was divided by two. We found that adding only half the noise in the calibration improved the results for these experimental evaluations.

3. RESULTS

3.A. Theoretical calculations

3.A.1. Calibration fits

When no tissue change or image noise was added to the evaluation tissues, the RMSE (mean error) was 0.1% (0.0%) for the PCD method for both four and two bins, compared to 0.2% (−0.1%) for the Hünemohr method and 0.7% (0.0%) for the stoichiometric method. This shows that all methods were appropriately implemented.

3.A.2. Noise evaluation

As expected, the RMSEs increased with the noise level (Fig. 2 and Table I), that is, the higher image noise added to the evaluation tissues, the higher SPR errors were seen. The errors for the Hünemohr method were much larger than for the three other SPR methods (see Table I).

The RMSE for the SPR estimates for each individual evaluation tissue followed the same trend using the PCD

FIG. 2. Relative SPR errors for the 71 evaluation tissues with added noise, using the PCD method (A: four bin, B: two bin) and the stoichiometric method (D). RMSEs were calculated for the three noise realizations for each tissue. The tissues were sorted according to their density (lowest to the left and highest to the right). Note the larger scale on the y-axis for the lung tissue at the left-hand side of the figure. The vertical dotted line between tissue numbers 46 and 47 separates the soft tissues (to the left of the line) and bone tissues (to the right of the line). The absolute SPR errors are shown in Fig. S4.A in Data S1. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE I. SPR errors for the 71 evaluation tissues with noise added to their CT numbers. The results presented are the RMSE, the mean error, and the interval from the 5% percentile to the 95% percentile of the error distributions. For brevity, the names of the methods will be abbreviated using the following symbols: A: PCD method — four bin, B: PCD method — two bin, C: Hünemohr's method, D: Stoichiometric method.

Noise level	RMSE (%)	Mean error (%)	[5%,95%] (%)
A			
Low	1.0	0.0	[−1.5,1.7]
Medium	1.6	0.1	[−2.1,2.5]
High	2.7	−0.1	[−3.6,3.3]
B			
Low	1.0	0.0	[−1.5,1.6]
Medium	1.6	0.0	[−2.1,2.3]
High	2.7	−0.0	[−3.4,3.3]
C			
Low	3.0	−0.0	[−4.5,4.9]
Medium	5.8	0.2	[−8.3,9.9]
High	10.5	0.2	[−15.9,16.3]
D			
Low	1.0	0.0	[−1.6,1.4]
Medium	1.5	0.0	[−2.2,2.2]
High	2.9	0.1	[−3.8,4.2]

method with both four and two bins and using the stoichiometric method (Fig. 2). In all cases, bone tissues were less affected by image noise than soft and lung tissues. For the highest noise level, the PCD method estimated the SPR of soft tissues better than the stoichiometric method. Note here that a higher noise was added to each bin for the two and four energy bins than for the full spectrum, as the total dose for each CT

configuration was kept the same (by adjusting the standard deviation of the noise distributions, see Section 2.E.2).

At the highest noise level, the RMSE over all tissues was only 2.7% for the PCD method using both four and two bins. In contrast, the Hünemohr method was found to be highly negatively impacted by image noise, and at the highest noise level, the RMSE for this method was 10.5%. An investigation of the effect of image noise on the Hünemohr method is presented in Data S1, Section S.2.

3.A.3. Tissue variations

The stoichiometric method was found to be most sensitive to tissue variations. The RMSE for the SPR estimates of the changed tissues with no noise added was 1.5% using this method (Table II). In contrast, for the PCD method, the RMSE was only 0.5% even though changes well above 4% were applied for some tissues, and still no bias was seen (mean error of around 0.0%, Table II). The Hünemohr method produced slightly better results in this evaluation, giving a RMSE of 0.4%.

After adding noise to the CT numbers for the changed evaluation tissues, the RMSEs again increased with the noise level (Fig. 3). Larger RMSEs for the SPR estimates were seen than when no tissue changes were applied, but the RMSEs for the individual tissues were more alike (compare Figs. 2 and 3). Again, the Hünemohr method was found to be very sensitive to image noise (Table II).

The absolute SPR errors corresponding to the results shown in Figs. 2 and 3 are shown in Fig. S4.A and S4.B in Data S1.

3.A.4. Tissue classification

We compared the tissue categorization estimated from the nearest neighbor classification to the theoretical categorization. For the unchanged tissues, misclassifications were only seen when adding the highest noise level (for one or six tissues for four and two bin, respectively). When changing the tissues, misclassification was seen for all noise levels,

TABLE II. SPR errors for the changed evaluation tissues with and without noise added to their CT numbers. Abbreviations for the names of the methods: A: PCD method — four bin, B: PCD method — two bin, C: HÜnemohr's method, D: Stoichiometric method.

Noise level	RMSE (%)	Mean error (%)	[5%,95%] (%)
A			
No noise	0.5	0.0	[−0.9,0.9]
Low	1.6	0.1	[−2.4,2.7]
Medium	2.1	0.1	[−3.0,3.4]
High	3.1	−0.1	[−4.3,4.2]
B			
No noise	0.5	−0.0	[−0.9,0.9]
Low	1.6	0.1	[−2.4,2.7]
Medium	2.1	0.1	[−3.0,3.3]
High	3.0	−0.0	[−4.2,4.2]
C			
No noise	0.4	−0.1	[−0.8,0.6]
Low	3.0	0.0	[−4.5,5.0]
Medium	5.8	0.2	[−8.3,9.9]
High	10.6	0.2	[−15.9,16.4]
D			
No noise	1.5	0.0	[−2.4,2.6]
Low	1.6	0.0	[−2.6,2.8]
Medium	2.0	0.1	[−3.0,3.2]
High	3.2	0.1	[−4.3,4.7]

including no noise. However, at most 0.005% of the tissues were misclassified, corresponding to 219 (242) tissues, at the highest noise level for four (two) bins.

3.B. Experimental validation

In the experimental evaluation based on real organic tissue samples, the best results were seen for the PCD method using two-bin images and the stoichiometric method, which both gave a RMSE for the SPR estimates of 0.8% over the nine tissue samples (Table III). The RMSE for the PCD method on four-bin images was slightly higher (RMSE = 1.0%), but still the SPR errors for four out of the nine tissue samples were within one standard deviation of the measured SPR (Fig. 4).

The Hünemohr method produced the largest errors, RMSE = 1.3%, however, the higher RMSE for this method was mainly dominated by the high relative SPR error for the femur bone sample of 3.2% (Fig. 4).

4. DISCUSSION

In this study, we have proposed a new SPR estimation method for PCD-CT images independently of the number of energy bins. The method was investigated theoretically in different scenarios encountered in real CT scans, including simulation of imaging noise and tissue variations. Through comparison with SPR estimation methods proposed for SECT and DECT as well as experimental measurements, we have shown that the proposed method gave good results when an appropriate calibration method was applied. For all theoretical evaluations, the mean SPR errors were close to 0%, showing that the method was unbiased. The RMSEs were also small in all tested cases. The fitting error was RMSE = 0.1%, and in the medium noise case, the RMSE was 1.3% for the standard tissues and 2.1% for the changed tissues.

FIG. 3. Relative SPR errors for the 71 evaluation tissues with tissue variations and added noise. RMSEs were calculated for the three noise realizations for each tissue. The absolute SPR errors are shown in Fig. S4.B in Data S1. [Color figure can be viewed at wileyonlinelibrary.com]

FIG. 4. SPR errors for the organic tissue samples, relative to the measured SPR values. The gray-shaded areas show the relative standard deviation for the measured SPR values, that is, $\sigma(\text{SPR}_{\text{meas}})/\text{SPR}_{\text{meas}} \cdot 100\%$. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE III. Results for the experimental evaluation. The RMSE over the nine tissue samples are given along with the standard deviation (σ) for the RMSE calculated using error propagation.

SPR method	RMSE (%)	$\sigma(\text{RMSE})$ (%)
PCD method – 4 bin	1.0	0.26
PCD method – 2 bin	0.8	0.33
Hünemohr's method	1.3	0.36
Stoichiometric method	0.8	0.30

For the theoretical evaluations, similar results were obtained for four and two energy bins. However, in the experimental evaluation, better results were seen for the two-bin images than for the four-bin images (Fig. 4). For the four-bin PCD-CT scans, the threshold configuration was not optimal, since the noise was unevenly distributed between the energy bins (higher noise was found in bin 1 and 2 than in bin 3 and 4, Section 2.F). The energy bin thresholds were optimized based on a smaller phantom and reoptimization for an object of the same size as the water tank was not performed. This indicates a dependency of phantom size on the threshold configurations, and it also implies that separate calibrations would be needed for different patient sizes, for example, head vs body.

From Table S1.I in Data S1, it is seen that our proposed CT number estimation method can result in effective energies higher than the maximal energy of the x-ray spectrum ($E_{\text{eff}}^{\text{bone}} = 158$ keV for bin 4 of Chess Mode, for a 140 kVp x-ray source spectrum). This does not make physical sense, but no restrictions were imposed on the effective energies, as these should not be seen as mean energies of the x-ray spectrum, but rather the energy where the measured CT numbers

and the estimated CT numbers matched the best. We tested several CT number estimation methods and found this new method based on two separate effective energies to perform the best (Table S5.I, Data S1) which is the reason for using it despite this behavior.

As the reconstruction algorithm also influences the measured CT numbers, it would have been of interest to investigate different reconstruction algorithms, including reconstruction algorithms which did not include a bone BHC. This was unfortunately not done in this study, which is a limitation of the study. There might be a need for a more in-depth optimization of the BHC for the PCD-CT reconstruction based on acquisition-specific parametrization. This could potentially explain the high effective energy for bone tissues in bin 4, if the bone BHC algorithm overcorrected the actual BH seen in the individual energy bins. As energy bin 4 of Chess Mode is very narrow and has a high mean energy, not much BH would be expected and an overcorrection in the reconstruction could have been obtained. Note that this behavior of “unphysical” effective energies was not seen for the Macro Mode. It was also not seen when tested for regular DECT spectra.^{18,38} Furthermore, it should be noted that the effective energies are not directly used in the SPR estimation method, but they are only used to find the A and B fitting parameters of Eq. (4).

The nearest neighbor classification resulted in very few misclassifications. We therefore deemed that using a model based on only three nearest neighbors was sufficient.

For four energy bins, a total of 23 fitting parameters were needed for Eq. (1). For more energy bins (as can be obtained using, e.g., the PCD-CT scanner described by Persson *et al.*³⁹), it can be assumed that the fit will be more robust if energy bins are merged, for example, by averaging the CT numbers for two or more energy bins.

It has been shown that the linear behavior of the stoichiometric method makes it very robust to noise. Especially for soft tissues, Bär et al. found the stoichiometric method to outperform DECT-based SPR methods in the case of high noise.²⁰ The PCD method, proposed in this study, had the same noise performance as the stoichiometric method, and for high noise, the PCD method even performed slightly better than the stoichiometric method for soft tissues (Fig. 2).

A disadvantage of the SECT-based stoichiometric method is its inflexibility to tissue variations.^{7,27} In this study, a RMSE of 1.5% was obtained for the stoichiometric method when applying tissue changes but no noise. The DECT-based Hünemohr method gave a RMSE of 0.4% (Table II). This superiority of the Hünemohr method was though lost when adding image noise (Tables I and II). In contrast, the PCD method produced better or equal result, compared to the Hünemohr and the stoichiometric method in all cases tested.

For the PCD method, it was found that calibrating on noisy data improved the robustness to noise. For the tissue samples, the noise in their PCD-CT images was directly measured. Furthermore, it could be speculated that, for a given scan protocol, the noise level would be the same for all patients of similar size, so the same set of fitting parameters could be used.

A high noise was seen in the PCD-CT images of the tissue samples. For the two-bin images, the noise level was ~ 40 HU in each energy bin (Section 2.F). This could potentially explain the higher RMSE for the Hünemohr method (Fig. 4), as it was shown in the theoretical evaluations that noise had a large negative impact on this method. However, the method by Hünemohr et al. was found to produce good results for DECT images in the study by Taasti et al.¹² This could potentially imply that SPR estimation based on this research prototype PCD-CT scanner could be improved, if the detector was further improved to reduce the noise levels.

In the experimental evaluation, we obtained a RMSE of 0.8% for the two-bin images. This result can be compared to the result for the Hünemohr method (RMSE = 1.3%) as the same image data were used for these two SPR methods, showing favorable results with the new PCD method. Lower errors using the Hünemohr method were found in our previous paper, where the RMSE were between 0.9% and 1.5% for the different DECT acquisition techniques evaluated.¹² This variation in the RMSEs shows that this SPR method is dependent on the CT energy spectra. Three further studies have evaluated DECT-based SPR estimation on fresh tissue samples.^{13–15}

It should be noted that the implementation of the Hünemohr method used in this study is not identical to Siemens *syngo.via* RhoZ software, as this commercial software includes a noise reduction step, and no noise reduction was included in this study. Furthermore, the used noise levels in this study are higher than typically seen in DECT scans.

Increasing the number of energy bins while keeping the dose constant will always increase the noise levels, since fewer photons will be detected in each energy bin. Moreover, for this particular detector design, only half of the subpixels were used for each of the energy bins in the four-bin mode, reducing the dose efficiency,²⁸ which leads to even higher

noise levels. One strategy to compensate for this would be the application of spectral denoising techniques such as multienergy nonlocal means.⁴⁰

The motivation for this study was the clear theoretical advantages of PCDs which would make PCD-CT interesting for SPR estimation, as an alternative to DECT. Currently, six different acquisition strategies for DECT imaging exist, dual source, kVp-switching between projections, kVp-switching between gantry rotations, two consecutive scans, split-filter, and dual-layer detectors.⁴¹ However, all of these techniques have different inherent disadvantages.⁴² These include limited FOV, poor spectral separation, and susceptibility to motion artifacts. The FOV of the investigated PCD-CT scanner is even smaller than for the current dual source models; however, this is not an inherent limitation for PCD-CT scanners in general. The FOV could be extended by using a larger detector module.⁴³ The two latter limitations of DECT are not present in PCD-CT scanners, as the spectral separation, to some extent, can be optimized for the specific imaging task at hand by carefully choosing the energy bin thresholds. Motion artifacts cannot occur as energy binning takes place at the detector level, which is also the case for dual-layer detectors.

Further advantages of PCD-CT scanning compared to DECT scanning include the ability of PCDs to exclude electronic noise and that multienergy imaging can be achieved.²⁸ Also, PCDs and dual-layer detectors are the ideal candidates for projection-based methods, since the projections from the different energy bins/energy spectra are perfectly aligned and acquired without temporal separation,⁴⁴ which is not the case for any of the other DECT systems. Moreover, PCDs can provide a better contrast for high-Z materials, since they weight the energy spectrum by the photon counts instead of the photon energy, which is the case for EIDs in SECT and DECT scanners.²⁸ PCDs therefore have some technical advantages whereby PCD-CT scanners using two energy bins potentially could supersede DECT scanners and eliminate most of the inherent limitations for current DECT imaging.

The separation of the individual energy bin spectra for a PCD is not ideal, and for some energy bins, the overlap can be nearly complete. The optimal threshold levels will depend on the size and attenuation properties of the scanned object, but the optimal threshold settings can be approximated beforehand if the x-ray source spectrum and detector response are known.⁴⁵

Based on this study, we consider PCD-CT to be a qualified candidate for SPR estimation in proton therapy treatment planning. Nevertheless, the PCD-CT scanner technology is not fully mature although much research is going on in the field and we therefore foresee that this technology can be valuable to improve the accuracy of the SPR estimation and overcome some of the inherent limitations of current DECT scanners.

5. CONCLUSION

In the theoretical evaluations, the proposed PCD-CT-based SPR estimation method produced good results for all tested

cases and was shown to be very robust toward CT image noise, with a RMSE of 2.7% in the high noise case. Low RMSEs were also obtained in the experimental measurements — the RMSEs were 1.0% and 0.8% for four and two bin PCD-CT scans, respectively.

CONFLICTS OF INTEREST

CHM, grant support from Siemens Healthcare.

ACKNOWLEDGMENTS

Research reported in this article was supported by the National Institutes of Health under award numbers C06 RR018898 and R01 EB016966 and by the Danish Cancer Society (grant no. R90-A5992). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The equipment and concepts described in this work are based on a research device and are not commercially available.

^{a)}Author to whom correspondence should be addressed. Electronic mail: taastiv@mskcc.org.

REFERENCES

- Ainsley CG, Yeager CM. Practical considerations in the calibration of CT scanners for proton therapy. *J Appl Clin Med Phys*. 2014;15:202–220.
- Wohlfahrt P, Möhler C, Hietschold V, et al. Clinical implementation of dual-energy CT for proton treatment planning on pseudo-monoenergetic CT scans. *Int J Radiat Oncol Biol Phys*. 2017;97:427–434.
- van Elmpt W, Landry G, Das M, Verhaegen F. Dual energy CT in radiotherapy: current applications and future outlook. *Radiother Oncol*. 2016;119:137–144.
- Hansen DC, Sørensen TS, Rit S. Fast reconstruction of low dose proton CT by sinogram interpolation. *Phys Med Biol*. 2016;61:5868–5882.
- Arbor N, Dauvergne D, Dedes G, et al. Monte Carlo comparison of x-ray and proton CT for range calculations of proton therapy beams. *Phys Med Biol*. 2015;60:7585–7599.
- Johnson RP, Bashkirov V, DeWitt L, et al. A fast experimental scanner for proton CT: technical performance and first experience with phantom scans. *IEEE Trans Nucl Sci*. 2016;63:52–60.
- Yang M, Virshup G, Clayton J, Zhu XR, Mohan R, Dong L. Theoretical variance analysis of single- and dual-energy computed tomography methods for calculating proton stopping power ratios of biological tissues. *Phys Med Biol*. 2010;55:1343–1362.
- Hünemohr N, Krauss B, Tremmel C, Ackermann B, Jäkel O, Greilich S. Experimental verification of ion stopping power prediction from dual energy CT data in tissue surrogates. *Phys Med Biol*. 2014;59:83–96.
- Hansen DC, Seco J, Sørensen TS, et al. A simulation study on proton computed tomography (CT) stopping power accuracy using dual energy CT scans as benchmark. *Acta Oncol*. 2015;54:1638–1642.
- Michalak G, Taasti V, Krauss B, Deisher A, Halaweish A, McCollough C. A comparison of relative proton stopping power measurements across patient size using dual- and single-energy CT. *Acta Oncol*. 2017;56:1465–1471.
- Wohlfahrt P, Möhler C, Richter C, Greilich S. Evaluation of stopping-power prediction by dual- and single-energy computed tomography in an anthropomorphic ground-truth phantom. *Int J Radiat Oncol Biol Phys*. 2018;100:244–253.
- Taasti VT, Michalak GJ, Hansen DC, et al. Validation of proton stopping power ratio estimation based on dual energy CT using fresh tissue samples. *Phys Med Biol*. 2018;63:015012.
- Bär E, Lalonde A, Zhang R, et al. Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples. *Med Phys*. 2018;45:48–59.
- Möhler C, Russ T, Wohlfahrt P, et al. Experimental verification of stopping-power prediction from single- and dual-energy computed tomography in biological tissues. *Phys Med Biol*. 2018;63:025001.
- Xie Y, Ainsley C, Yin L, et al. Ex vivo validation of a stoichiometric dual energy CT proton stopping power ratio calibration. *Phys Med Biol*. 2018;63:055016.
- Hudobivnik N, Schwarz F, Johnson T, et al. Comparison of proton therapy treatment planning for head tumors with a pencil beam algorithm on dual and single energy CT images. *Med Phys*. 2016;43:495–504.
- Wohlfahrt P, Möhler C, Stützer K, Greilich S, Richter C. Dual-energy CT based proton range prediction in head and pelvic tumor patients. *Radiother Oncol*. 2017;125:526–533.
- Taasti VT, Muren LP, Jensen K, et al. Comparison of single and dual energy CT for stopping power determination in proton therapy of head and neck cancer. *Phys Imaging Radiat Oncol*. 2018;6:14–19.
- Schneider U, Pedroni E, Lomax A. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Phys Med Biol*. 1996;41:111–124.
- Bär E, Lalonde A, Royle G, Lu H-M, Bouchard H. The potential of dual-energy CT to reduce proton beam range uncertainties. *Med Phys*. 2017;44:2332–2344.
- Yu Z, Leng S, Li Z, et al. How low can we go in radiation dose for the data-completion scan on a research whole-body photon-counting computed tomography system. *J Comput Assist Tomogr*. 2016;40:663–670.
- Yu Z, Leng S, Li Z, McCollough CH. Spectral prior image constrained compressed sensing (spectral PICCS) for photon-counting computed tomography. *Phys Med Biol*. 2016;61:6707–6732.
- Taguchi K, Iwanczyk JS. Vision 20/20: single photon counting x-ray detectors in medical imaging. *Med Phys*. 2013;40:100901.
- Lalonde A, Bouchard H. A general method to derive tissue parameters for Monte Carlo dose calculation with multi-energy CT. *Phys Med Biol*. 2016;61:8044–8069.
- Lalonde A, Bär E, Bouchard H. A Bayesian approach to solve proton stopping powers from noisy multi-energy CT data. *Med Phys*. 2017;44:5293–5302.
- Shen C, Li B, Chen L, Yang M, Lou Y, Jia X. Material elemental decomposition in dual and multi-energy CT via a sparsity-dictionary approach for proton stopping power ratio calculation. *Med Phys*. 2018;45:1491–1503.
- Taasti VT, Petersen JBB, Muren LP, Thygesen J, Hansen DC. A robust empirical parametrization of proton stopping power using dual energy CT. *Med Phys*. 2016;43:5547–5560.
- Yu Z, Leng S, Jorgensen SM, et al. Evaluation of conventional imaging performance in a research whole-body CT system with a photon-counting detector array. *Phys Med Biol*. 2016;61:1572–1595.
- Kappler S, Henning A, Kreisler B, Schoeck F, Stierstorfer K, Flohr T. Photon counting CT at elevated x-ray tube currents: contrast stability, image noise and multi-energy performance. In: *Proc. SPIE*. Vol. 9033; 2014: 90331C.
- Woodard HQ, White DR. The composition of body tissues. *Br J Radiol*. 1986;59:1209–1219.
- White DR, Woodard HQ, Hammond SM. Average soft-tissue and bone models for use in radiotherapy dosimetry. *Br J Radiol*. 1987;60:907–913.
- ICRP. *ICRP Publication 110: Adult Reference Computational Phantoms*. Ann. ICRP. Vol. 39, 2009. <http://www.icrp.org/publication.asp?id=icrp%20publication%20110>. Accessed October 5, 2016.
- Berger MJ, Hubbell JH, Seltzer SM, et al. *XCOR: Photon Cross Section Database (version 1.5)*. Gaithersburg, MD: National Institute of Standards and Technology; 2010. [Online] Available: <http://physics.nist.gov/xcom>.
- Han D, Siebers JV, Williamson JF. A linear, separable two-parameter model for dual energy CT imaging of proton stopping power computation. *Med Phys*. 2016;43:600–612.
- Cover T, Hart P. Nearest neighbor pattern classification. *IEEE Trans Inf Theory*. 1967;13:21–27.
- Sietsma J, Dow RJF. Creating artificial neural networks that generalize. *Neural Netw*. 1991;4:67–79.

37. Taasti, VT, Bäumer C, Dahlgren CV, et al. Inter-centre variability of CT-based stopping-power prediction in particle therapy: Survey-based evaluation. *Phys Imaging Radiat Oncol.* 2018;6:25–30.
38. Taasti VT. *Improving proton range determination using new x-ray computed tomography principles*. PhD thesis, Aarhus University, The Department of Oncology, Aarhus, Denmark; 2017. http://pure.au.dk/portal/files/127077795/PhDDissertation_VickiTrierTaasti.pdf
39. Persson M, Huber B, Karlsson S, et al. Energy-resolved CT imaging with a photon-counting silicon-strip detector. *Phys Med Biol.* 2014;59:6709–6727.
40. Li Z, Leng S, Yu L, Manduca A, McCollough CH. An effective noise reduction method for multi-energy CT images that exploit spatio-spectral features. *Med Phys.* 2017;44:1610–1623.
41. De Cecco CN, Laghi A, Schoepf UJ, Meinel FG (eds). *Dual Energy CT in Oncology*. Cham, Switzerland: Springer International Publishing; 2015.
42. Zou Y, Silver MD. Analysis of fast kV-switching in dual energy CT using a pre-reconstruction decomposition technique. In: *Proc. SPIE*. Vol. 6913; 2008:691313.
43. Yu Z, Leng S, Kappler S, et al. Low-dose performance of a whole-body research photon-counting CT scanner. In: *Proc. SPIE*. Vol. 9783; 2016:97835Q.
44. Vilches-Freixas G, Taasti VT, Muren LP, et al. Comparison of projection-and image-based methods for proton stopping power estimation using dual energy CT. *Phys Imaging Radiat Oncol.* 2017;3:28–36.
45. Li Z, Leng S, Yu Z, Kappler S, McCollough CH. Estimation of signal and noise for a whole-body photon counting research CT system. In: *Medical Imaging 2016: Physics of Medical Imaging*, Vol. 9783. International Society for Optics and Photonics; 2016: 97831F.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Data S1: Theoretical and experimental analysis of photon counting detector CT for proton stopping power prediction.