RESEARCH ARTICLE

MEDICAL PHYSICS

Comparative accuracy and resolution assessment of two prototype proton computed tomography scanners

George Dedes[1] | Hubertus Drosten[1] | Stefanie Götz[1] | Jannis Dickmann[1] | Christina Sarosiek[2] | Mark Pankuch[3] | Nils Krah[4] | Simon Rit[4] | Vladimir Bashkirov[5] | Reinhard W. Schulte[5] | Robert P. Johnson[6] | Katia Parodi[1] | Ethan DeJongh[7] | Guillaume Landry[1,8,9]

[1]Department of Medical Physics, Faculty of Physics, Ludwig-Maximilians-Universität München, München, Germany

[2]Department of Physics, Northern Illinois University, DeKalb, Illinois, USA

[3]Northwestern Medicine Chicago Proton Center, Warrenville, Illinois, USA

[4]University of Lyon, INSA-Lyon, Université Claude Bernard Lyon 1, Lyon, France

[5]Division of Biomedical Engineering Sciences, Loma Linda University, Loma Linda, California, USA

[6]Department of Physics, U.C. Santa Cruz, Santa Cruz, California, USA

[7]ProtonVDA LLC, Naperville, Illinois, USA

[8]Department of Radiation Oncology, University Hospital, LMU Munich, Munich, Germany

[9]German Cancer Consortium (DKTK), Munich, Germany

Correspondence
G. Dedes, Department of Medical Physics, Faculty of Physics, Ludwig-Maximilians-Universität München, Am Coulombwall 1, 85748 Garching b., München, Germany.
Email: g.dedes@lmu.de

Abstract

Background: Improving the accuracy of relative stopping power (RSP) in proton therapy may allow reducing range margins. Proton computed tomography (pCT) has been shown to provide state-of-the-art RSP accuracy estimation, and various scanner prototypes have recently been built. The different approaches used in scanner design are expected to impact spatial resolution and RSP accuracy.

Purpose: The goal of this study was to perform the first direct comparison, in terms of spatial resolution and RSP accuracy, of two pCT prototype scanners installed at the same facility and by using the same image reconstruction algorithm.

Methods: A phantom containing cylindrical inserts of known RSP was scanned at the phase-II pCT prototype of the U.S. pCT collaboration and at the commercially oriented ProtonVDA scanner. Following distance-driven binning filtered backprojection reconstruction, the radial edge spread function of high-density inserts was used to estimate the spatial resolution. RSP accuracy was evaluated by the mean absolute percent error (MAPE) over the inserts. No direct imaging dose estimation was possible, which prevented a comparison of the two scanners in terms of RSP noise.

Results: In terms of RSP accuracy, both scanners achieved the same MAPE of 0.72% when excluding the porous sinus insert from the evaluation. The ProtonVDA scanner reached a better overall MAPE when all inserts and the body of the phantom were accounted for (0.81%), compared to the phase-II scanner (1.14%). The spatial resolution with the phase-II scanner was found to be 0.61 lp/mm, while for the ProtonVDA scanner somewhat lower at 0.46 lp/mm.

Conclusions: The comparison between two prototype pCT scanners operated in the same clinical facility showed that they both fulfill the requirement of an RSP accuracy of about 1%. Their spatial resolution performance reflects the different design choices of either a scanner with full tracking capabilities (phase-II) or of a more compact tracker system, which only provides the positions of protons but not their directions (ProtonVDA).

KEYWORDS
design choices, image quality, proton CT, RSP, spatial resolution

1 | INTRODUCTION

Radiation therapy with external proton beams offers the potential for highly conformal dose distributions with the possibility of healthy tissue sparing. The point where protons stop in the patient is dependent on the stopping power of the tissue they traverse. This is commonly expressed as relative to water and denoted as relative stopping power (RSP). An accurate three dimensional (3D) patient RSP image in treatment position is needed for accurate proton dose calculation. Errors in the 3D patient RSP image will cause proton range prediction errors, thus introducing range uncertainties that entail the use of safety margins.[1] Currently, RSP images are obtained by converting X-ray linear attenuation coefficients, acquired with single energy X-ray computed tomography (CT) imaging. This is often based on a stoichiometric calibration,[2] which leads to institution-specific margin recipes, for example, 3.5% plus 1 mm at the Massachusetts General Hospital or similar recipes at other proton centers.[1,3]

Besides CT imaging for treatment planning, the use of volumetric image guidance for proton therapy is becoming well established.[4] Beyond patient positioning, there is a need for adaptive replanning based on accurate RSP estimation in 3D, when the patient is in treatment position at isocenter, which allows reduced planning margins. The leading example for improved RSP estimation is dual-energy X-ray CT (DECT), which after undergoing extensive phantom[5–14] and animal tissue validation,[15–19] has seen clinical implementation for treatment planning,[20–22] and which is now appearing in treatment rooms, but requires the patient to be moved away from isocenter for verification.[23]

An alternative is to use the proton beam itself for imaging and to acquire aproton CT (pCT) scan of the patient on the treatment table. While X-ray interactions are sparse, multiple Coulomb scattering (MCS) of protons means that straight-path approximations in the reconstruction algorithm lead to poor spatial resolution.[24] A curved path approximating the true proton path can be estimated from measurements of the location of each proton before and after the object, and can be refined with information about the direction of the proton, inferred from two locations separated by air or materials with a low scattering power. The RSP line integrals are evaluated along these curved paths,[25] and thus, a pCT scanner equipped with tracking modules and an energy detector allows the use of proton tracking information in image reconstruction. pCT reconstruction algorithms account for curved paths, either using iterative reconstruction[26–28] or filtered-backprojection.[29–31]

Several prototype pCT scanners have been presented in the literature,[32–38] and an early simulation study suggested that pCT may provide competitive ideal RSP accuracy compared to DECT.[39] Many of these scanners have been used to image phantoms, and experimental RSP accuracy is reported better than 1.6%,[37] 1.4%,[40] or 0.74%,[38] and was shown to be equivalent to state-of-the-art DECT in a comparison using known RSP tissue mimicking phantoms[41] with a mean absolute percentage error (MAPE) of 0.6%. Since various proton tracking pCT prototypes are in the early development stage, it is yet unclear what the optimal system design is. Two of the most advanced scanners currently in operation are the pCT collaboration's phase-II scanner[35] and the commercially oriented ProtonVDA (pVDA) scanner.[42–44] The phase-II scanner features silicon-strip tracking modules registering both the location and direction of protons before and after the object, and a five-stage energy detector. The pVDA scanner uses scintillating fibers tracking modules, which register only position with initial direction vectors derived from the beam geometry, and a single-stage energy detector requiring variation of the proton beam energy. Given the differences in particle tracking and energy detection between the scanners, and since they are both available at the Northwestern Medicine Chicago Proton Center, this study aimed at performing the first comparison of two particle tracking pCT scanner prototypes.

2 | MATERIALS AND METHODS

To compare the performance of the new commercial pVDA system against the well-studied phase-II scanner in terms of spatial resolution and RSP accuracy, we scanned a phantom of known RSP with both systems. In order to focus on the hardware differences between the two scanners, preconditioning of the data before image reconstruction and the reconstruction algorithm were the same for both scanners. Details about the scanners, the data processing, the reconstruction algorithm, and the scanned object are given in the subsequent sections.

2.1 | Phase-II prototype scanner

The phase-II scanner[34] is depicted in Figure 1, left panel, and was designed by the pCT collaboration (at the Baylor University, Loma Linda University, and the University of California at Santa Cruz). The particle tracking system consists of two tracking modules,[35] one before and one after the object, measuring position and direction information, as well as a scintillating detector[45] measuring the residual energy.

Each of the two tracking modules[35] consists of two tracking planes separated by 50 mm, each measuring the coordinates of every incoming proton. Using the two position measurements and the direction vector, the most likely path (MLP) can be estimated.[25] Coordinates are determined as hit locations of two adjacent planes of single-sided silicon strip detectors with lateral and vertical strip orientation, respectively, to provide
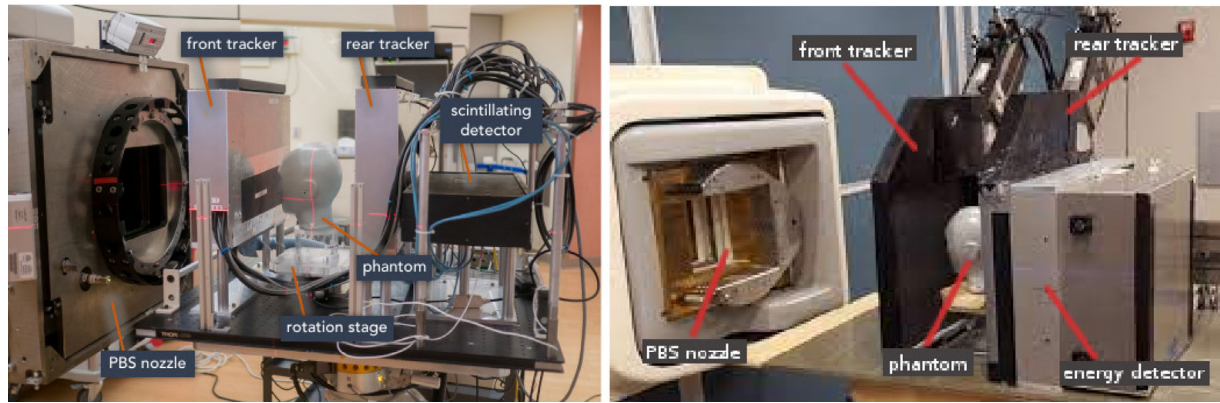
FIGURE 1 (Left) Phase-II scanner with various parts of the system indicated (taken from Reference 48 - © Institute of Physics and Engineering in Medicine. Reproduced by permission of IOP Publishing. All rights reserved). (Right) Photo of the pVDA scanner with its main parts labeled

two-dimensional coordinate measurements. Laterally, each tracking plane consists of four modules that are glued together, resulting in an insensitive gap of 0.6 mm.

The energy detector[45] consists of five longitudinal segments of UPS-923A[46] (RSP = 1.038), which are individually wrapped in reflective material and coupled to a photomultiplier. Each segment is referred to as a *stage* and their longitudinal depth is 51 mm. Each incident proton can produce up to five readings in the analog-digital-converters (ADCs) of the photo–multipliers. The segmented design with five stages was chosen to minimize water-equivalent path length (WEPL) noise and to make it almost independent of the scanned object.

The scanner has been shown to produce a near 100% detection efficiency for count rates up to 1 MHz with homogeneous proton fluence.[35] For operation of the scanner with scanned pencil beams, the local count rate increases and a pileup-free operation was demonstrated for count rates up to 400 kHz.[47,48] For the purpose of this study, a 200 MeV broad proton beam was utilized with the phase-II scanner. More details about the beam characteristics are given in Section 2.6

2.2 | pVDA scanner

The second pCT scanner under investigation is the prototype[43] of ProtonVDA LLC (Naperville, IL, USA) depicted in Figure 1, right panel. Its more compact design is intended for future clinical use. The system has two tracking modules, one upstream and one downstream the scanned object. Unlike the phase-II scanner, these tracking modules only consist of a single tracking plane and no direction information is acquired directly.

Tracking is realized through scintillating fibers. Each tracking module consists of two layers of fibers, one oriented horizontally and one vertically to provide two-dimensional coordinate measurements. The system can only be operated with scanned pencil beams due to the

multiplexing of the tracking fibers: multiple nonadjacent fibers are connected to one of many photomultipliers and an exact and unambiguous location is determined using the approximate location of the pencil beam.

The energy detector is realized as a 130 mm thick monolithic scintillator block. The scintillation light is measured by 16 individual photomultipliers, which are distributed laterally across the surface of the detector. Due to the limited depth of the sensor, only thin objects can be scanned with a single proton energy. To scan thicker objects, scans of several incident proton energies are combined, rejecting protons that stopped in the object or completely penetrated the scintillator. The system is designed for count rates of up to 10 MHz;[43] in this experiment, it was operated reliably at count rates of $1-2$ MHz.

2.3 | Data processing of the phase-II scanner

The data processing steps of the phase-II scanner, before data conditioning for reconstruction and image reconstruction itself, have been described by Schultze et al.[49] The acquired data from the tracking detector and the energy detector are converted to coordinates in the reference system of the trackers and WEPL values for proton histories that are not rejected for various reasons, for example, incomplete tracking data or suspected pileup events.

To generate WEPL measurements from the ADC readout, the phase-II scanner makes use of an elaborate calibration procedure based on data from a double-wedged polystyrene phantom (RSP= 1.03) complemented by one to four rectangular polystyrene blocks.[50] A detailed description of the calibration procedure is given in several previous publications, including Bashkirov et al.,[45] Piersimoni et al.,[50] Dedes et al.,[51] and Schultze et al.[49]

The calibration runs cover the complete WEPL range of the detector. They are used to establish two-dimensional histograms binned by WEPL and energy deposit to the stopping stage. The stopping stage is defined as the furthest stage where the energy deposit exceeded a certain threshold (usually 1 MeV) and the WEPL can be inferred from the tracking information and the geometry of the calibration phantom. From this histogram, for each energy deposit, the most frequent WEPL is found, which allows establishing an energy-to-WEPL calibration curve for each of the five stages of the energy detector. These calibration curves are then applied to subsequent imaging runs to convert the ADC measurements of an unknown object to WEPL values. Finally, the WEPL values are combined with the four coordinate measurements in the tracker and yield the proton-by-proton WEPL, upstream/downstream position, and upstream/downstream direction information used in the reconstruction algorithm, as described by Schultze et al.[49]

During this particular experiment, noise on the energy detector data was slightly elevated compared to past scans. This effect was manifested as increased RSP artifacts compared to what was previously observed[41] for the energy threshold of 1 MeV to select the stopping stage. To reduce the ring artifacts, which intersect the inserts and affect the RSP accuracy, the stage threshold was varied from 1.5 to 3 MeV in 0.25 MeV steps, in addition to the default threshold value. The optimal threshold value was found to be 1.5 MeV, yielding images with ring artifacts of comparable amplitude to what was previously reported in the published literature.[41]

2.4 | Data processing of the pVDA scanner

The pVDA scanner is calibrated without the use of a dedicated phantom and instead utilizes multiple acquisitions at different incident beam energies. The calibration procedure converts the photomultiplier signals to residual proton range in WEPL. Protons are delivered at 44 energies ranging from 100.8 to 164.7 MeV, with their range reduced by a 6.5 cm water-equivalent material uniform absorber, covering proton ranges between 1.25 and 12 cm. 1.25 cm is roughly the minimum proton range needed to traverse the preceding detector material, enter the scintillator, and produce a signal just above the noise threshold. Each incident energy smaller than the maximum energy corresponds to a reduction in range and produces energy deposits approximately proportional to the equivalent WEPL in the energy detector. Via this calibration procedure, for each of the 44 proton beam energies, the corresponding WEPL in the detector is mapped to the ADC number produced by the detector. The exact procedure, including the correction of the spatial dependence of the signal with

TABLE 1 WEPL sampling for the pVDA scanner

Energy (MeV)	WEPL (mm)	WEPL for reconstruction (mm)
118	0–90	0–90
160	50–125	50–125
187	100–200	125–200

Note: The first column indicates the energies used. The second column lists the WEPL interval covered by each energy. The third column shows the WEPL range finally used for reconstruction, after removing the overlap between the two higher energies.

respect to the locations of the energy deposit and the photomultiplier, is described in detail in DeJongh et al.[43]

Since the system's energy detector is too thin to cover the desired WEPL range directly, data from multiple energies are merged. The calibration needs to be performed only once, since calibrated WEPLs can be shifted by the difference in range of a lower incident energy. For the imaging runs presented in this work, three incident energies (118, 160, and 187 MeV) were acquired throughout the field of view, which resulted in an increased imaging dose, but could be avoided in case prior knowledge of the scanned object is available.[52] Each of the beam energies covered a specific WEPL range in the object scanned in this experiment, with some overlapping WEPLs between them, as indicated in Table 1. The first overlap region extends from approximately 50 to about 90 mm and is covered by 118 and 160 MeV beams, respectively. The second overlap region spans 100 − 125 mm and was sampled by 160 and 187 MeV beams. For that particular dataset, it was found that using protons from 187 MeV beam together with protons from the other beams for sampling the same WEPL regions caused considerable ring-shaped image artifacts. Therefore, the use of 187 MeV was reduced to the bare minimum, namely, for WEPLs larger than 125 mm. With this filtration, the number of protons available for reconstruction was reduced by 38% compared to the initial dataset.

The proton angle at the front tracker is determined from the location of the proton hit and the distance from the steering magnets in the beamline.[43] The rear tracker angle is determined by an MLP calculation. During that procedure, the most likely combination of exit position from the object and exit angle, given the upstream tracker position, upstream angle, and downstream tracker position, is estimated on a proton-by-proton basis.

2.5 | Phantom

The analyses performed in this work were based on scans of a custom-built phantom with known RSP and geometry shown in Figure 2. The phantom's body
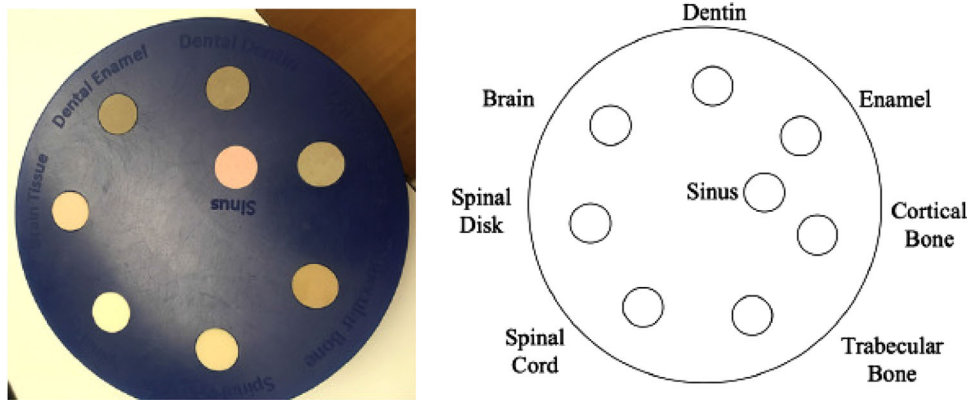
FIGURE 2   (Left) Photo of the cylindrical phantom made of plastic body and containing eight plastic tissue equivalent inserts. (Right) Drawing of the phantom with the inserts labeled. It should be noted that the labels of the enamel and dentin inserts as seen in the photograph have been mistakenly swapped in the manufacturing process. The correct positions are shown in the drawing

consists of blue wax with an RSP of 0.980. It is 40 mm thick and has a diameter of 180 mm. It contains eight inserts made of tissue-equivalent plastic materials, each with a diameter of 18 mm and thickness of 40 mm, listed in Figure 2. The reference RSP (RSP_{ref}) of each insert was determined in another study[42] using a multilayer ionization chamber.

2.6 | Datasets

The phase-II scanner data of the cylindrical phantom were acquired at the Northwestern Medicine Chicago Proton center using a 200 MeV wobbled beam of approximately 1.7 cm size (1 σ at isocenter). The acquisition was obtained in a continuous rotation mode, which means that the phantom was rotating throughout the beam-on time and the data were later on split into 360 projections, each one covering an arc of 1°, using the timestamp of each proton and the known rotation speed (1 *rpm*) of the stage. The scan time was approximately 5 min, during which about 3.6×10^8 protons were acquired at a 1 MHz rate. From the 360 projections, 90 projections at 4° steps were kept and used in the reconstruction. This choice was made in order to have the same number of projections and at the same angular steps as in the pVDA dataset, which was acquired in a step-and-shoot mode. The data were processed with a calibration curve obtained on the same day with the same beam characteristics as the imaging run. After cuts, in the processing of the raw data, the mean proton fluence in the object was 30 protons/mm^2 per projection. As described in Section 2.3, the energy detector stage threshold was optimized to suppress image artifacts. Prior to reconstruction, the data were binned in 2 mm × 2 mm sized-pixels at the front tracker and 3-σ cuts were applied on the angle and WEPL distribution in each pixel.

The pVDA scanner data of the same phantom were acquired in the same facility in a step-and-shoot mode (90 projections at 4° angular steps), by scanning a 0.4 − 0.8 cm size (energy dependent and quantified as 1 σ at isocenter) pencil beam over an area large enough to cover the phantom. Three incident energies (118 , 160, and 187 MeV) were utilized to sample the WEPL range of the phantom taking into account the limited thickness of the energy detector. The beam-on time, considering all three energies, was approximately 90 s, during which about 10^8 protons were acquired at a rate between 1 and 2 MHz. This does not account for the overhead time for rotating the phantom between two projections (phantom rotation has been automated after that experiment). Additionally, about 10–20 s were needed for beam energy switching. The three incident energies combined covered a WEPL range from 0 to approximately 200 mm (as shown in Table 1). To avoid inadequate WEPL sampling, the WEPL interval covered by each energy overlaps with that covered by another energy. The extent of this overlap was tuned to provide the best possible image quality, as previously described in Section 2.4. After the initial processing and merging, a mean fluence in the object of approximately 20 protons/mm^2 per projection was forwarded to the reconstruction. Similarly to the procedure followed for the phase-II scanner, prior to reconstruction, the data were binned in 2 mm × 2 mm sized-pixels at the front tracker and 3-σ cuts were applied on their angular and WEPL distribution in each pixel.

2.7 | Reconstruction algorithm

Following the processing of the raw scanner data and the filtering on angle and WEPL, the volumetric RSP images were reconstructed using a dedicated filtered backprojection algorithm. A detailed description of the

concept of the algorithm is given in Rit et al.[29] To account for the curved proton trajectories due to MCS, the path of every proton is estimated by an MLP formulation.[25] The MLP estimate is based on the position and direction information from the tracker in the case of the phase-II scanner, and on the positions from the tracker and the angles deduced in the processing of the data in the case of the pVDA scanner.

The proton-by-proton data were binned in projection images with 1 mm × 1 mm pixels for use with the RSP accuracy analysis and with 0.2 mm × 0.2 mm pixels in the case of the spatial resolution determination. The projections were then filtered and backprojected. In the last step, the pCT images were reconstructed as RSP maps in a grid of 1 mm × 1 mm × 1 mm for RSP analysis or 0.2 mm × 0.2 mm × 1 mm for spatial resolution. The reconstruction was applied to both scanner datasets without change of parameters.

2.8 | RSP accuracy and spatial resolution

The RSP accuracy (percent RSP error) was quantified separately in each insert. A circular region-of-interest (ROI) with a radius of 7 mm (80% of the insert radius to avoid insert boundaries) was applied around the insert centers and the mean reconstructed RSP (RSP_{mean}) in each ROI was calculated. The RSP accuracy (RSP_{acc}) was then defined as:

$$RSP_{acc} = \frac{RSP_{mean} - RSP_{ref}}{RSP_{ref}} \cdot 100\%, \quad (1)$$

where RSP_{ref} denotes the reference RSP of each insert. The uncertainty of the RSP_{acc} per insert is derived from the uncertainty on the RSP_{ref} and from the uncertainty on the RSP_{mean}, making use of error propagation. The former was defined as the uncertainty from a 0.1 mm error in the sample thickness estimation during the multilayer ionization chamber measurement. The latter was based on the standard error of the mean (SEM), calculated over the values of N voxels in the ROI:

$$SEM = \frac{\sigma_{rec}}{\sqrt{N}}, \quad (2)$$

where σ_{rec} is the standard deviation of the reconstructed RSP values of the N voxels in the ROI.

In addition to the RSP for each insert, the MAPE achieved for each scanner for that particular phantom was also calculated as:

$$MAPE = \frac{\sum_{i=1}^{n} |RSP_{acc,i}|}{n}, \quad (3)$$

where n is the total number of inserts and $RSP_{acc,i}$ is the percent error for every insert i as calculated from Equation (1). The uncertainty on the MAPE was calculated via an error propagation based on the individual insert accuracy.

The spatial resolution was evaluated by calculating the modulation transfer function (MTF) from the radial edge-spread function (ESF)[53] of the three inserts with the highest RSP (enamel, cortical bone, and dentin). After taking the average of 20 slices to reduce noise in the images and determining the insert centers using a threshold followed by a center of mass calculation, we followed the approach of Krah et al.[54] and Khellaf et al.[31] and modeled the ESF as an error function with parameters A, μ, σ, and C.

$$ESF = A \cdot \frac{1}{2}\left[1 + erf\left(\frac{r - \mu}{\sqrt{2}\sigma}\right)\right] + C, \quad (4)$$

where r is the distance of the sampled points to the insert center. Using the parameters obtained by the fit of the ESF to the inserts, the frequency of the MTF at the 10% level, used here as the metric for spatial resolution, is then given by:

$$f_{MTF_{10\%}} = \sqrt{\frac{\ln 10}{2}} \cdot \frac{1}{\pi \sigma}. \quad (5)$$

The uncertainty on the $f_{MTF_{10\%}}$ is estimated by propagating the uncertainty on σ from the fit described in Equations (4) and (5).

3 | RESULTS

3.1 | RSP accuracy and spatial resolution

For the comparison of the performance of the two scanners, the RSP_{acc} and spatial resolution evaluations were applied on the pCT images (Figure 3a for phase-II and Figure 3b for pVDA scanners, respectively). Ring-shaped artifacts are present in both images and to some extent, affect the quantified RSP_{acc}. The difference image in Figure 3c shows that artifacts did not appear at the same location in both images. A slight residual misalignment is the cause of large differences at the edges of some inserts. Undersampling streak artifacts are also observed in images from both scanners, and are attributed to the limited number of projections (90) used. We confirmed that for the phase-II scanner, these vanish when 360 projections are used (not shown). The RSP_{acc} results are listed in Table 2-phase-II and Table 3-pVDA.

In the case of the phase-II scanner (Table 2), the RSP_{acc} varied from −0.29% for the brain insert up to −4.5% for the sinus. Except for the sinus, the enamel insert (−1.03%), and the phantom body (−1.33%), the
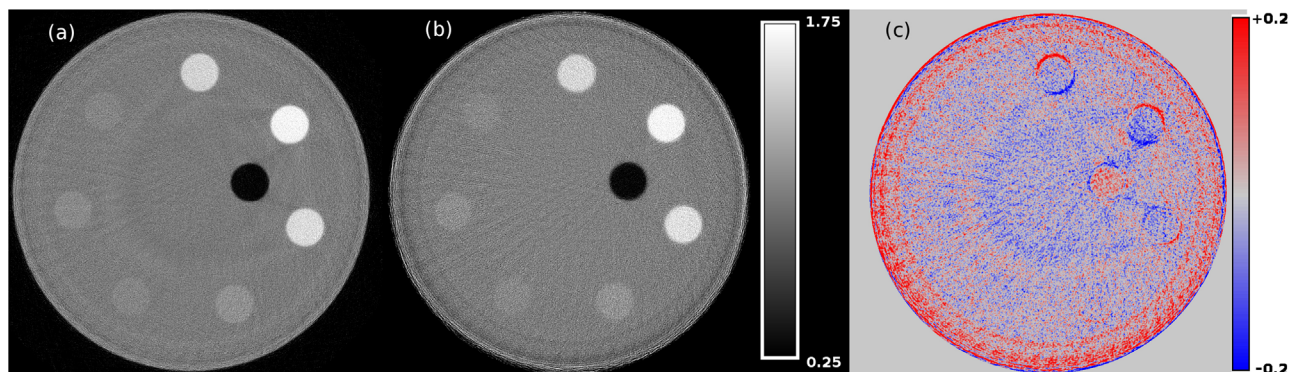
FIGURE 3 Reconstructed RSP images with 0.2 mm × 0.2 mm pixel size from (a) the phase-II and (b) the pVDA scanner. The average of 20 central slices was taken to reduce noise and the images are shown with an RSP level of 1.0 and window of 1.5. In (c), the difference (b)–(a) is shown with an RSP level of 0 and window of 0.4

TABLE 2 Results of the RSP analysis for the phase-II scanner data

Insert	RSP_{ref}	RSP_{mean}	RSP_{acc}/%	ROI size/voxels
Cortical bone	1.555 ± 0.004	1.543 ± 0.0006	-0.77 ± 0.25	3120
Trabecular bone	1.100 ± 0.003	1.095 ± 0.0006	-0.45 ± 0.25	3060
Spinal disc	1.070 ± 0.003	1.066 ± 0.0005	-0.37 ± 0.25	3080
Enamel	1.755 ± 0.004	1.737 ± 0.0005	-1.03 ± 0.25	3060
Dentin	1.495 ± 0.004	1.481 ± 0.0006	-0.94 ± 0.25	3020
Sinus	0.200 ± 0.005	0.191 ± 0.0005	-4.50 ± 0.35	3040
Phantom body	0.980 ± 0.002	0.967 ± 0.0001	-1.33 ± 0.25	46 100
Spinal cord	1.040 ± 0.003	1.034 ± 0.0005	-0.58 ± 0.25	3060
Brain	1.040 ± 0.003	1.037 ± 0.0005	-0.29 ± 0.25	3100
MAPE (all values)			1.14 ± 0.09	
MAPE (w/o body)			1.12 ± 0.09	
MAPE (w/o sinus)			0.72 ± 0.09	
MAPE (w/o body and sinus)			0.63 ± 0.10	

Note: The table lists for each insert the RSP_{ref}, the RSP_{mean}, the RSP_{acc}, and the size of the ROI in voxels. For different subsets of inserts, the MAPE is also listed. The uncertainty on each value was obtained as explained in Section 2.8.

RSP_{acc} was better than 1%. The MAPE over all inserts and the phantom body for the phase-II scanner image was 1.14%. Excluding the phantom body, the MAPE marginally improved to 1.12%. Excluding the sinus, the MAPE reached 0.72%. Finally, excluding both the sinus and the body of the phantom from the RSP_{acc} quantification yielded a MAPE of 0.63%.

For the pVDA scanner (Table 3), the RSP_{acc} varied from 0.1% for the spinal cord and brain inserts up to −2.4% for the phantom body. Except for the body of the phantom, the only other insert for which the RSP_{acc} was worse than 1% was the sinus with −1.5%. The MAPE over all inserts and the phantom body was 0.81%. When excluding the phantom body, the MAPE improved to 0.61%. Excluding the sinus, the MAPE reached 0.72%. When excluding both the sinus and the body of the phantom, the MAPE became 0.48%. The RSP_{acc} per

insert and the MAPE are displayed in Figure 4 for both scanners.

The spatial resolution as estimated by the radial edge profile of the three higher RSP inserts –enamel, cortical bone, and dentin– for both scanners is reported in Table 4. The phase-II scanner yields a spatial resolution of 0.59–0.62lp/mm, while the analysis for the pVDA scanner resulted in a slightly lower spatial resolution of 0.44–0.48lp/mm.

4 | DISCUSSION

The scope of this study was to compare, in terms of RSP_{acc} and spatial resolution, two pCT scanners of different design available for testing in the same clinical facility. The data were collected for the same object on

TABLE 3 Results of the RSP analysis for the pVDA scanner data

Insert	RSP_{ref}	RSP_{mean}	RSP_{acc}/%	ROI size/voxels
Cortical bone	1.555 ± 0.004	1.569 ± 0.001	0.90 ± 0.26	3060
Trabecular bone	1.100 ± 0.003	1.103 ± 0.001	0.27 ± 0.27	3100
Spinal disc	1.070 ± 0.003	1.073 ± 0.001	0.28 ± 0.28	3100
Enamel	1.755 ± 0.004	1.771 ± 0.001	0.91 ± 0.26	3120
Dentin	1.495 ± 0.004	1.507 ± 0.001	0.80 ± 0.27	3120
Sinus	0.200 ± 0.005	0.197 ± 0.001	-1.50 ± 0.65	3120
Phantom body	0.980 ± 0.002	0.957 ± 0.001	-2.40 ± 0.29	448 860
Spinal cord	1.040 ± 0.003	1.041 ± 0.001	0.10 ± 0.28	3060
Brain	1.040 ± 0.003	1.041 ± 0.001	0.10 ± 0.28	3080
MAPE (all values)			0.81 ± 0.11	
MAPE (w/o body)			0.61 ± 0.12	
MAPE (w/o sinus)			0.72 ± 0.10	
MAPE (w/o body and sinus)			0.48 ± 0.10	

Note: The table lists for each insert the RSP_{ref}, the RSP_{mean}, the $RSP_{acc,}$ and the size of the ROI in voxels. For different subsets of inserts, the MAPE is also listed. The uncertainty on each value was obtained as explained in Section 2.8.
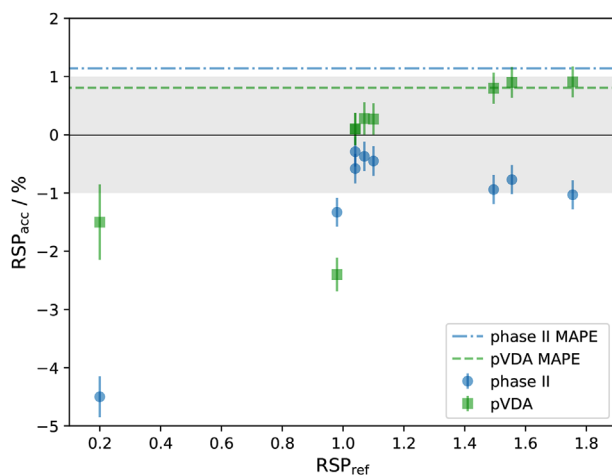

FIGURE 4 RSP accuracy comparison between the phase-II and the pVDA scanners. Each point represents an RSP_{acc} value in % from Tables 2 and 3, with the corresponding uncertainty depicted as the error bar. The shaded area signifies the $\pm 1\%$ accuracy range. The dashed and the dotted-dashed lines indicate the MAPE for each scanner, when accounting for all inserts and the body of the phantom

TABLE 4 $f_{MTF_{10}}$ results for the three highest RSP inserts of the phase-II and pVDA scanners

Insert	$f_{MTF_{10}}^{phase-II}$ (lp/mm)	$f_{MTF_{10}}^{pVDA}$
Cortical bone	0.61 (0.02)	0.47 (0.02)
Dentin	0.62 (0.02)	0.44 (0.02)
Enamel	0.59 (0.01)	0.48 (0.02)
Mean	0.61 (0.01)	0.46 (0.01)

The values and their uncertainties in brackets were obtained as described in Section 2.8.

different days, albeit using different proton beam delivery methods for the two scanners, namely, a wobbled beam for the phase-II scanner and a pencil beam for the pVDA scanner.

No direct imaging dose estimation was possible, which prevented a comparison of the two scanners in terms of RSP noise. This limitation was due to lack of information on the total number of protons delivered by the beamline, which has to be operated in low-fluence mode for both scanners. The fluence is below the threshold where beam diagnostics function well. The number of protons registered by each scanner correlates to dose, but does not reflect lost protons due to nuclear interactions, pile up, or other detector inefficiencies.

As it may be noticed in Figure 3, the pVDA image appears noisier. This is also confirmed by the RSP_{mean} uncertainty per insert shown in Tables 2 and 3 (a factor of 1.7–2 higher for the pVDA scanner), and can be only partially explained by the higher mean proton fluence in the case of the phase-II scanner. Nevertheless, several effects would have to be considered before drawing conclusions about a fluence-based noise comparison. One is the different detector design that could lead to different WEPL or detected energy resolution. This implies that, even for the same mean fluence, different noise levels could be expected for each scanner. Another important aspect is that the pVDA image is composed of scans with three different energies (in general, the lower the energy, the higher the energy straggling into the object). This would have to be considered in any estimation of the noise, and the noise in the pVDA three-energy image might even have a different spatial distribution compared to a pVDA single-energy image. Thus, with the data at our disposal, a comparison based

only on the mean fluence would be inconclusive and the topic certainly deserves a separate investigation.

By design, the pVDA scanner cannot use a single proton energy to cover the WEPL dynamic range required for large objects. Therefore, three different scan energies ($118, 160,$ and $187\,\mathrm{MeV}$) were used and assembled into a single dataset. It was observed that the overlap of the two higher energies resulted in strong ring artifacts present in the RSP image. The optimal RSP image from this dataset was obtained by removing the overlap between the two highest energies and keeping only the WEPLs sampled from the $160\,\mathrm{MeV}$ beam in the region between $100 - 125\,\mathrm{mm}$. The origin of the problematic overlap between the two higher energies remains unclear, and the solution adopted in this manuscript might be applicable and valid for this particular dataset only. In a recent work published by DeJongh and DeJongh,[55] images of the same phantom reconstructed with an iterative reconstruction algorithm show only very faint RSP artifacts, but a more conclusive investigation would require comparison of the two algorithms and quantification of the RSP artifacts versus spatial resolution, especially in the case of high values of the relaxation parameter λ.

For the phase-II scanner, the overall MAPE was slightly above 1% ($1.14 \pm 0.09\%$), while for the pVDA scanner, it was $0.81 \pm 0.11\%$. The quantified $\mathrm{RSP_{acc}}$ of the phase-II scanner was strongly affected by the very large relative error of the sinus insert ($-4.50 \pm 0.35\%$). For the pVDA scanner, the $\mathrm{RSP_{acc}}$ of that insert was also the worst among all inserts ($-1.5 \pm 0.65\%$), but still closer to the desired $\pm 1\%$. When excluding the sinus insert from the analysis, the MAPE for both scanners was 0.72%. It is worth noting that the absolute RSP error $|\mathrm{RSP_{mean}} - \mathrm{RSP_{ref}}|$ on the sinus insert is not generally higher than that of other inserts. For the phase-II scanner, the sinus absolute RSP error is 0.009, while the average absolute RSP error of all inserts is also 0.009 and ranging from 0.003 to 0.018. Similarly, for the pVDA scanner, the sinus absolute RSP error is 0.003, while the average absolute RSP error of all inserts is 0.008 and ranging from 0.001 to 0.023. Furthermore, protons traversing the sinus insert do not have significantly different WEPL compared to protons crossing other inserts, as the majority of the proton path is traveled across the body of the phantom. Therefore, the WEPL measurement error of protons traversing the sinus insert and other inserts is comparable. Thus, the very large relative error of the sinus insert is caused by its small reference RSP, which is the mean value of the air and the trabecular material of the sinus composite. It is conceivable that the systematically lower RSP values measured by both scanners in the case of the sinus insert can be caused by the mean distance traveled in the air component being larger than that traveled in the denser material due to preferentially larger scattering out of the denser material. The next worse $\mathrm{RSP_{acc}}$ for

both scanners was found to be that of the phantom body, with $-1.33 \pm 0.25\%$ and $-2.40 \pm 0.29\%$, respectively. This is to some extent expected as, unlike the small ROIs used for the inserts, the large area in which the body of the phantom $\mathrm{RSP_{acc}}$ was quantified contains most of the observed ring artifacts. For example, quantifying the phantom body RSP in an ROI as large as for the insert at the center of the phantom yielded an accuracy of -0.1% and 1.63% for the phase-II and the pVDA scanners, respectively. When the same sized ROI was moved to a radial distance similar to that of the other inserts, the $\mathrm{RSP_{acc}}$ was -0.92% and -0.1% for the phase-II and the pVDA scanners, respectively. Removing the body of the phantom from the MAPE analysis improves the phase-II scanner accuracy only marginally, from 1.14% to 1.12%, again because it is dominated by the difference in the sinus insert. On the other hand, for the pVDA, the MAPE without the phantom body improves from 0.81% to 0.61%. Finally, a more selective RSP quantification, which excludes both the phantom body and the sinus, yields for both scanners equivalent MAPE of $0.63 \pm 0.10\%$ and $0.48 \pm 0.10\%$. That MAPE of the phase-II scanner (without sinus and phantom body) agrees well with the result from[41] (0.55%), which was obtained from a set of inserts from a different phantom, but covering a similar RSP range and making use of the same reconstruction algorithm.[29]

For both scanners, the spatial resolution obtained from this phantom was within the range expected for pCT systems. For the phase-II scanner, it varied from 0.59 to 0.62 lp/mm. This result is in good agreement with what was reported by Plautz et al.[56] In that work, two quantities where the spatial resolution was quantified were used to define the location in a phantom. These were the circle's minimum chord intersecting the insert and the smallest distance of the insert from the surface of the phantom. For the phantom used in our study, these values were $130 - 150\,\mathrm{mm}$ in WEPL and about $20 - 30\,\mathrm{mm}$, respectively. For the pVDA scanner, the spatial resolution was found to be lower, ranging from 0.44 to 0.48 lp/mm. This can be explained by the fact that the phase-II scanner implements a full tracking capability at the cost of a less compact detector scheme, while the pVDA scanner comprises a more compact tracker design for future clinical use. Furthermore, the lower energies used with the pVDA scanner also affect the scattering power of protons, which is expected to impact spatial resolution.

To test this hypothesis, we made use of the formalism published in Krah et al.[54] to estimate the relative loss of resolution expected from removing downstream position measurements for the phase-II scanner. This led to a decrease of resolution by a factor 1.5, which is in reasonable agreement with the decrease by a factor 1.2 to 1.4 observed for the inserts of Table 4. It is thus likely that the use of position-only trackers downstream, where directions cannot be approximated by the vector

from the source to the position registered by the tracker (as opposed to the upstream tracker), contributes chiefly to the lower spatial resolution of the pVDA scanner. Approximating the pVDA scanner by the phase-II scanner without downstream directions should be valid since both scanners have a distance of approximately 16 cm between trackers and the isocenter.

Whether the difference in spatial resolution affects dosimetric and range prediction accuracy when pCT is used for treatment planning remains to be investigated. Finally, the spatial resolution was evaluated at inserts located at relatively large radii. For more central inserts, where MCS has an increased role, the difference between the two scanners may be reduced and dominated by the accuracy of the MLP.

5 | CONCLUSIONS

The comparison of the phase-II and pVDA particle tracking pCT scanners showed that they provide similar RSP_{acc} (both a MAPE of 0.72%) when excluding the porous sinus insert; when including it, the phase-II scanner performed slightly worse than the pVDA (1.14% vs. 0.81%, respectively). Spatial resolution estimated by three high-density cylindrical inserts was found to be 0.61 lp/mm for the phase-II scanner and slightly lower, at 0.46 lp/mm, for the pVDA scanner. Their spatial resolution performance reflects the different hardware design choices, with the more compact pVDA tracker, which does not provide directions, leading to a slightly lower spatial resolution. The clinical significance of these findings remains to be studied.

CONFLICT OF INTEREST
M. Pankuch and R. W. Schulte are members of the advisory board of ProtonVDA LLC. E. DeJongh is employed by ProtonVDA LLC.

REFERENCES
1. Paganetti H. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Phys Med Biol*. 2012;57:R99-R117.
2. Schneider U, Pedroni E, Lomax A. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Phys Med Biol*. 1996;41:111-124.
3. Yang M, Zhu XR, Park PC, et al. Comprehensive analysis of proton range uncertainties related to patient stopping-power-ratio estimation using the stoichiometric calibration. *Phys Med Biol*. 2012;57:4095-4115.
4. Landry G, Hua C-H. Current state and future applications of radiological image guidance for particle therapy. *Med Phys*. 2018;45:e1086-e1095.
5. 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*. 2013;59:83.
6. Bourque AE, Carrier J-F, Bouchard H. A stoichiometric calibration method for dual energy computed tomography. *Phys Med Biol*. 2014;59:2059-2088.
7. 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.
8. Möhler C, Wohlfahrt P, Richter C, Greilich S. Range prediction for tissue mixtures based on dual-energy CT. *Phys Med Biol*. 2016;61:N268.
9. 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.
10. 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.
11. 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.
12. Saito M, Sagara S. Simplified derivation of stopping power ratio in the human body from dual-energy CT data. *Med Phys*. 2017;44:4179-4187.
13. Almeida IP, Schyns LEJR, Vaniqui A, et al. Monte Carlo proton dose calculations using a radiotherapy specific dual-energy CT scanner for tissue segmentation and range assessment. *Phys Med Biol*. 2018;63:115008.
14. Landry G, Dörringer F, Si-Mohamed S, et al. Technical Note: Relative proton stopping power estimation from virtual monoenergetic images reconstructed from dual-layer computed tomography. *Med Phys*. 2019;46:1821-1828.
15. 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*. 2017;63:015012.
16. 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.
17. 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.
18. 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.
19. Niepel KB, Stanislawski M, Wuerl M, et al. Animal tissue-based quantitative comparison of dual-energy CT to SPR conversion methods using high-resolution gel dosimetry. *Phys Med Biol*. 2021;66:075009.
20. 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.
21. 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.

22. Wohlfahrt P, Möhler C, Enghardt W, et al. Refinement of the Hounsfield look-up table by retrospective application of patient-specific direct proton stopping-power prediction from dual-energy CT. *Med Phys*. 2020;47:1796-1806.

23. Berthold J, Khamfongkhruea C, Petzoldt J, et al. First-in-human validation of CT-based proton range prediction using prompt gamma imaging in prostate cancer treatments. *Int J Radiat Oncol Biol Phys*. 2021;111:1033-1043.

24. Zygmanski P, Gall KP, Rabin MSZ, Rosenthal SJ. The measurement of proton stopping power using proton-cone-beam computed tomography. *Phys Med Biol*. 2000;45:511-528.

25. Schulte RW, Penfold SN, Tafas JT, Schubert KE. A maximum likelihood proton path formalism for application in proton computed tomography. *Med Phys*. 2008;35:4849-4856.

26. Penfold SN, Rosenfeld AB, Schulte RW, Schubert KE. A more accurate reconstruction system matrix for quantitative proton computed tomography. *Med Phys*. 2009;36:4511-4518.

27. Penfold SN, Schulte RW, Censor Y, Rosenfeld AB. Total variation superiorization schemes in proton computed tomography image reconstruction. *Med Phys*. 2010;37:5887-5895.

28. Hansen DC, Bassler N, Sørensen TS, Seco J. The image quality of ion computed tomography at clinical imaging dose levels. *Med Phys*. 2014;41:111908.

29. Rit S, Dedes G, Freud N, Sarrut D, Létang JM. Filtered backprojection proton CT reconstruction along most likely paths. *Med Phys*. 2013;40:031103.

30. Poludniowski G, Allinson NM, Evans PM. Proton computed tomography reconstruction using a backprojection-then-filtering approach. *Phys Med Biol*. 2014;59:7905-7918.

31. Khellaf F, Krah N, Létang JM, Collins-Fekete C-A, Rit S. A comparison of direct reconstruction algorithms in proton computed tomography. *Phys Med Biol*. 2020;65:105010.

32. Takada Y, Kondo K, Marume T, Nagayoshi K, Okada I, Takikawa K. Proton computed tomography with a 250 MeV pulsed beam. *Nucl Instr Meth Phys Res A*. 1988;273:410-422.

33. Coutrakon G, Bashkirov V, Hurley F, et al. Design and construction of the 1st proton CT scanner. In *AIP Conference Proceedings 1525, 327 (2013)*, 2013.

34. Sadrozinski HF-W, Geoghegan T, Harvey E, et al. Operation of the preclinical head scanner for proton CT. *Nucl Instr Meth Phys Res A*. 2016;831:394-399.

35. 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.

36. Pettersen HES, Alme J, Biegun A, et al. Proton tracking in a high-granularity digital tracking calorimeter for proton CT purposes. *Nucl Instr Meth Phys Res A*. 2016;860:51-61.

37. Esposito M, Waltham C, Taylor JT, et al. PRaVDA: the first solid-state system for proton computed tomography. *Physica Med*. 2018;55:149-154.

38. Civinini C, Scaringella M, Brianzi M, et al. Relative stopping power measurements and prosthesis artifacts reduction in proton CT. *Phys Med Biol*. 2020;65:225012.

39. Hansen DC, Sangild Sørensen T, Rit S. Fast reconstruction of low dose proton CT by sinogram interpolation. *Phys Med Biol*. 2016;61:5868-5882.

40. Giacometti V, Bashkirov VA, Piersimoni P, et al. Software platform for simulation of a prototype proton CT scanner. *Med Phys*. 2017;44:1002-1016.

41. Dedes G, Dickmann J, Niepel K, et al. Experimental comparison of proton CT and dual energy X-ray CT for relative stopping power estimation in proton therapy. *Phys Med Biol*. 2019;64:165002.

42. Sarosiek C, DeJongh EA, Coutrakon G, et al. Analysis of characteristics of images acquired with a prototype clinical proton radiography system. *Med Phys*. 2021;48(5):2271-2278.

43. DeJongh EA, DeJongh DF, Polnyi I, et al. Technical Note: A fast and monolithic prototype clinical proton radiography system optimized for pencil beam scanning. *Med Phys*. 2021;48:1356-1364.

44. DeJongh DF, DeJongh EA, Rykalin V, et al. A comparison of proton stopping power measured with proton CT and X-ray CT in fresh postmortem porcine structures. *Med Phys*. 2021;48(12):7998-8009.

45. Bashkirov VA, Schulte RW, Hurley RF, et al. Novel scintillation detector design and performance for proton radiography and computed tomography. *Med Phys*. 2016;43:664-674.

46. Artikov A, Budagov J, Chirikov-Zorin I, et al. Properties of the Ukraine polystyrene-based plastic scintillator UPS 923A. *Nucl Instr & Methods Phys Res Sect A: Accel Spectromet Detect Assoc Equip*. 2005;555:125-131.

47. Dedes G, Johnson RP, Pankuch M, et al. Experimental fluence-modulated proton computed tomography by pencil beam scanning. *Med Phys*. 2018;45:3287-3296.

48. Dickmann J, Sarosiek C, Rykalin V, et al. Experimental realization of dynamic fluence field optimization for proton computed tomography. *Phys Med Biol*. 2020;65:195001.

49. Schultze B, Karbasi P, Sarosiek C, et al. Particle-tracking proton computed tomography–data acquisition, preprocessing, and preconditioning. *IEEE Access*. 2021;9:25946-25958.

50. Piersimoni P, Ramos-Méndez J, Geoghegan T, Bashkirov V, Schulte R, Faddegon B. The effect of beam purity and scanner complexity on proton CT accuracy. *Med Phys*. 2017;44:284-298.

51. Dedes G, Dickmann J, Giacometti V, et al. The role of Monte Carlo simulation in understanding the performance of proton computed tomography. *Z Med Phys*. 2022;32(1):23-38.

52. Dickmann J, Sarosiek C, Rykalin V, et al. Proof of concept image artifact reduction by energy-modulated proton computed tomography (EMpCT). *Physica Med*. 2021;81:237-244.

53. Richard S, Husarik DB, Yadava G, Murphy SN, Samei E. Towards task-based assessment of CT performance: system and object MTF across different reconstruction algorithms. *Med Phys*. 2012;39:4115-4122.

54. Krah N, Khellaf F, Létang JM, Rit S, Rinaldi I. A comprehensive theoretical comparison of proton imaging set-ups in terms of spatial resolution. *Phys Med Biol*. 2018;63:135013.

55. Jongh DFD, Jongh EAD. An iterative least squares method for proton CT image reconstruction. *IEEE Trans Radiat Plasma Med Sci*. 2022;6(3):304-312.u

56. Plautz TE, Bashkirov V, Giacometti V, et al. An evaluation of spatial resolution of a prototype proton CT scanner. *Med Phys*. 2016;43:6291-6300.