Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

PAPER • OPEN ACCESS

A simulation framework for preclinical proton irradiation workflow

To cite this article: Justin Malimban *et al* 2024 *Phys. Med. Biol.* **69** 215040

IOP Publishing Phys. Med. Biol. 69 (2024) 215040 https://doi.org/10.1088/1361-6560/ad897f

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

OPEN ACCESS

RECEIVED
27 June 2024

REVISED
2 October 2024

ACCEPTED FOR PUBLICATION
21 October 2024

PUBLISHED
4 November 2024

PAPER

A simulation framework for preclinical proton irradiation workflow

Justin Malimban[1,*], Felix Ludwig[1], Danny Lathouwers[2], Marius Staring[3,4], Frank Verhaegen[5] and Sytze Brandenburg[1]

[1] Department of Radiation Oncology and Particle Therapy Research Center, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
[2] Department of Radiation Science and Technology, Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands
[3] Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands
[4] Department of Radiation Oncology, Leiden University Medical Center, Leiden, The Netherlands
[5] Department of Radiation Oncology (MAASTRO), Research Institute for Oncology & Reproduction, Maastricht University Medical Center, Maastricht, The Netherlands
[*] Author to whom any correspondence should be addressed.

E-mail: j.malimban@umcg.nl

Keywords: micro-CT simulation, single-energy CT, dual-energy CT, stopping power, preclinical proton dose calculations, Monte Carlo

Supplementary material for this article is available online

Abstract

Objective. The integration of proton beamlines with x-ray imaging/irradiation platforms has opened up possibilities for image-guided Bragg peak irradiations in small animals. Such irradiations allow selective targeting of normal tissue substructures and tumours. However, their small size and location pose challenges in designing experiments. This work presents a simulation framework useful for optimizing beamlines, imaging protocols, and design of animal experiments. The usage of the framework is demonstrated, mainly focusing on the imaging part. *Approach.* The fastCAT toolkit was modified with Monte Carlo (MC)-calculated primary and scatter data of a small animal imager for the simulation of micro-CT scans. The simulated CT of a mini-calibration phantom from fastCAT was validated against a full MC TOPAS CT simulation. A realistic beam model of a preclinical proton facility was obtained from beam transport simulations to create irradiation plans in matRad. Simulated CT images of a digital mouse phantom were generated using single-energy CT (SECT) and dual-energy CT (DECT) protocols and their accuracy in proton stopping power ratio (SPR) estimation and their impact on calculated proton dose distributions in a mouse were evaluated. *Main results.* The CT numbers from fastCAT agree within 11 HU with TOPAS except for materials at the centre of the phantom. Discrepancies for central inserts are caused by beam hardening issues. The root mean square deviation in the SPR for the best SECT (90 kV/Cu) and DECT (50 kV/Al–90 kV/Al) protocols are 3.7% and 1.0%, respectively. Dose distributions calculated for SECT and DECT datasets revealed range shifts <0.1 mm, gamma pass rates (3%/0.1 mm) greater than 99%, and no substantial dosimetric differences for all structures. The outcomes suggest that SECT is sufficient for proton treatment planning in animals. *Significance.* The framework is a useful tool for the development of an optimized experimental configuration without using animals and beam time.

1. Introduction

Over the past decade, there has been a rapid increase in the number of proton therapy facilities around the world (PTCOG 2024). Owing to the characteristic dose fall off after the Bragg peak, proton therapy can deliver a much more conformal dose to the tumour, thereby more effectively sparing surrounding healthy tissues. To fully realize the potential and exploit the benefits of proton therapy, preclinical studies are needed to better understand the biological mechanisms of tumour and normal tissue response and to study differential effects of x-rays and protons.

Several *in vivo* studies have already been conducted to investigate the interplay between radiation induced damage to the heart and lungs (Ghobadi *et al* 2012) as well as inhibition of repair mechanisms by low dose irradiations around the primary high dose areas in the spinal cord (Bijl *et al* 2006). Previous works have also looked into the relative biological effectiveness (RBE) by irradiating mice with different parts of the Bragg curve (Sørensen *et al* 2017, Howard *et al* 2021, Denbeigh *et al* 2024). Other treatment strategies are also being explored such as using an array of narrow, spatially fractionated beams—proton minibeam (Prezado *et al* 2017, 2019) and ultra-high dose rates—FLASH (Beyreuther *et al* 2019, Kim *et al* 2021, Sørensen *et al* 2022) to promote normal tissue sparing. Targeted irradiations of critical normal tissue structures such as stem cell-rich ducts in the parotid glands (van Luijk *et al* 2015) and hippocampus (Suckert *et al* 2021) also suggest that sparing these regions may reduce associated radiation-induced side effects (i.e. xerostomia and neurocognitive dysfunction, respectively).

To get more insight into the radiation response of both normal tissue structures and tumours, highly accurate irradiations must be performed, which can be challenging due to motion, location, and size of the targets in small animals. To create opportunities for these experiments, facilities have started to integrate preclinical x-ray CT imaging and irradiation platforms with proton beamlines to provide the image guidance needed to achieve the required accuracy (Ford *et al* 2017, Kim *et al* 2019, Parodi *et al* 2019, Schneider *et al* 2024). This opens up possibilities for delivering spread-out Bragg peak (SOBP) irradiations, which allow a much better conformity of the dose distributions to the targets. However, designing such experiments is not straightforward. To achieve highly conformal dose distributions, millimeter-sized pencil beams must be delivered. The beam must also be degraded down to energies much lower than a clinical proton beam ($\sim$30 MeV). These require additional components such as collimators and range shifters to be integrated into the beamline. The material chosen and the position of these components with respect to the animal affect the quality of the dose distributions. Furthermore, correct positioning of the Bragg peaks must be ensured and the actual dose distribution in the animal must be accurately determined.

In this work, we present a simulation framework of the preclinical irradiation workflow, which can be used to optimize beam properties, evaluate imaging protocols, and characterize dose distributions prior to performing experiments. The framework provides a quick way to discover sensitivities and weak points of experimental setups so they can be addressed beforehand. This leads to a more efficient workflow and also enhances experiment quality and capacity.

The framework allows generation of realistic x-ray micro-CT images using the fastCAT CBCT simulator (O'Connell and Bazalova-Carter 2021). We have modified the original code to enable creation of CT scans consistent with a small animal imager. This was validated against a full Monte Carlo (MC) x-ray CT simulation in TOPAS (Perl *et al* 2012). Beam transport simulations in the Geant4-based beam delivery simulation (BDSIM) toolkit (Nevay *et al* 2020) were performed to obtain a realistic proton beam model optimized for small animal irradiations. The proton beam model was used to generate the beam data library in matRad (Wieser *et al* 2017) for treatment planning.

We focus on the imaging issues to demonstrate an application of this framework. Single-energy CT (SECT) and dual-energy CT (DECT) calibration methods have already been investigated for preclinical studies but only for x-ray irradiations (Schyns *et al* 2017, Vaniqui *et al* 2017). Here, we extend that to protons by evaluating different published SECT and DECT approaches for estimation of proton stopping power ratios (SPR) and material identification in animal CTs. The uncertainties associated with the calibration methods are assessed by simulating proton radiographs to get range error maps. Their impact on proton dose distributions was also evaluated to establish whether the differences observed between SECT and DECT were large enough to warrant DECT-based proton treatment planning for animals.

2. Materials and methods

2.1. Micro-CBCT model for fastCAT

FastCAT is a cone-beam CT (CBCT) simulation toolkit relying on scatter kernels and detector response functions pre-calculated using MC simulations (O'Connell and Bazalova-Carter 2021). The x-ray source is modelled using SpekPy, which is a python-based program that allows calculation of polychromatic x-ray spectra for a wide range of x-ray tube specifications (Poludniowski *et al* 2021). For kilovoltage beams, which are typically used for small animal CT acquisition, fastCAT offers a 450 μm thick Cesium iodide (CsI) detector. To create CT images, a 3D voxel geometry of the object is imported in the form of a matrix with integers corresponding to the material index. For each of these materials, the linear attenuation coefficient (μ) as a function of energy taken from the NIST XCOM database is used as input in the simulation. The CT parameters such as source-to-axis distance (SAD), source-to-detector distance (SDD), and imaging dose must also be defined. Then, projection images are created using raytracing, which are subsequently used for CT reconstruction in TIGRE (Biguri *et al* 2016).

Table 1. Parameters used in fastCAT based on the X-RAD 225Cx machine.

Parameter	Value
Focal spot	0.4 mm based on IEC 336
Inherent filtration	0.8 mm Be
Anode	Tungsten (W)
Anode angle	20°
Source-to-axis distance (SAD)	303.4 mm
Source-to-detector distance (SDD)	622.9 mm
Number of projections	360

In this way, fastCAT is capable of generating realistic CT images in minutes. However, in its original form it is not suitable for micro-CT simulations since its base data (i.e. primary and scatter kernels) were modelled after a clinical CBCT machine, and the phantoms available are in the human scale. Therefore, we performed separate MC simulations in TOPAS to generate base data for micro-CT acquisitions using the X-RAD 225Cx (Precision x-ray Inc, Madison, CT) as the reference CT model. Table 1 gives the CT parameters for this small animal imager.

The following subsections detail the changes we have made in fastCAT and the MC validation of the modified version. The physics list and range cuts used in TOPAS can be found in table S1 in the supplementary file.

2.1.1. Primary and scattered x-ray modelling in TOPAS

CT images in fastCAT are created by first calculating forward projections using raytracing in TIGRE for 18 discrete energies (i.e. 10–100 keV in increments of 10 keV, 300–900 keV in increments of 200 keV, and 1, 2, 4 and 6 MeV). Each projection is turned into an intensity image by using the primary field and then, the scatter contribution is subsequently added. The final intensity image is obtained by weighting the 18 projection images by the x-ray spectrum and energy deposition efficiency.

We calculated the primary field and scatter contribution for our micro-CT model following the simulation procedure described in the paper of O'Connell and Bazalova–Carter but using the CT parameters given in table 1. The primary field was obtained by irradiating the detector without an object. The source was modelled as a cone beam with a Gaussian focal spot size with $\sigma = 0.4$ mm. The *SurfaceTrackCount* scorer in TOPAS was used to count the x-rays incident on a 2D air slab (512×512, pixel size $= 0.15$ mm). A radial profile of the x-ray intensity was taken and the same curve as in the original fastCAT paper was fitted on the resulting profile. Since the shape of the primary field does not change much with energy, the simulation was only performed for 90 keV x-rays and the same profile was used for the other energies.

The phantom-specific scatter kernels were generated for a spherical water phantom with radius of 15 mm for imaging a mouse. These were calculated for all energies required by fastCAT. Similar to the primary field, a radial profile exists for each energy. The scatter contribution is assumed to be independent of the projection angle.

Since the filtration materials were not explicitly modelled in fastCAT, additional MC simulations of a single CT projection were performed to estimate the contribution of the scattered photons generated by the filters. This evaluation was done for the two filters available in X-RAD 225Cx: 2 mm aluminium (Al) and 0.32 mm copper (Cu). The contribution of scattered photons reaching the detector is minimal, with 0.18% for Al and 0.23% for Cu. Given these values, the exclusion of the filters from the simulation does not influence the results.

2.1.2. Image noise

FastCAT applies Poisson noise to the final intensity image to create a more realistic CT. To generate CTs at different noise levels (or dose levels), fastCAT scales the raw intensity and the noise using the ratio of the user-defined total particle fluence and the reference value (i.e. number of particles for which the base data was calculated for in the MC simulation). Here, the dose in a spherical water phantom was scored using the *DoseToMedium* scorer in TOPAS for a given number of x-rays. This gives a relation that allows to assess image noise as a function of imaging dose.

2.1.3. Phantoms

Two types of phantoms were used in this work: a mouse-sized CT calibration phantom (SmART Scientific Solutions BV, Maastricht, the Netherlands) and the mathematical MOBY mouse phantom (Segars *et al* 2004). The mini-calibration phantom, which contains tissue-equivalent materials of known elemental composition (supplementary table S2), was used for the CT HU calibration. It has a diameter of 30 mm and

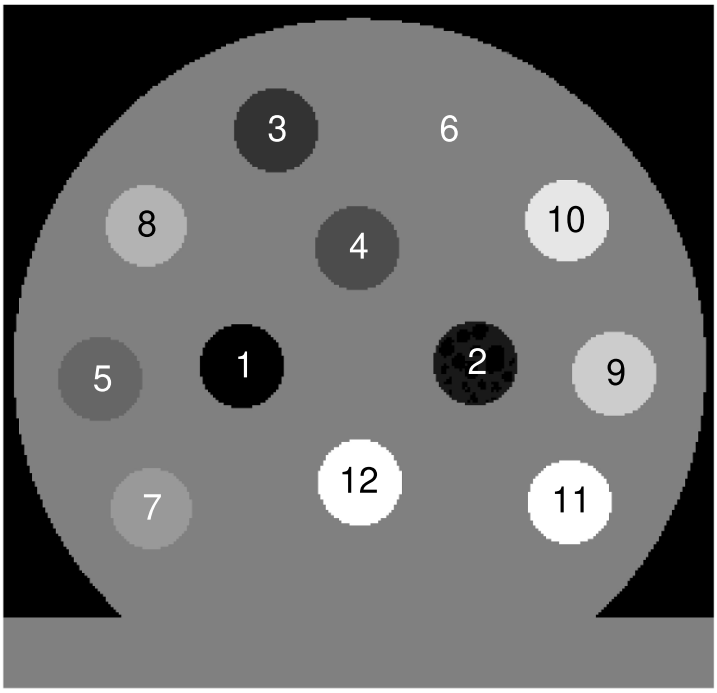

Figure 1. Schematic diagram of the mini-calibration phantom. The numbers correspond to materials in table 2.

Table 2. Effective atomic number $\left(Z_{\text{eff}} = \sqrt[m]{\frac{\sum_i w_i \frac{Z_i}{A_i} Z_i^m}{\sum_i w_i \frac{Z_i}{A_i}}},\ m = 3.3\right)$ (Landry *et al* 2013), relative electron density (ρ_e), and stopping power ratio (*SPR*) relative to water (calculated using the Bethe Bloch equation for 100 MeV protons) of the Gammex materials. Values for the lung insert are not given as it is a highly heterogeneous material.

#	Material	Z_{eff}	ρ_e	*SPR*
1	Air	7.71	0.001	0.001
2	LN-450 lung	—	—	—
3	AP6 adipose	6.21	0.928	0.947
4	SR2 brain	6.09	1.047	1.075
5	BR12 breast	6.93	0.956	0.972
6	Solid water	7.74	0.992	1.005
7	LV1 liver	7.74	1.064	1.078
8	IB3 inner bone	10.42	1.086	1.082
9	B200 bone	10.42	1.103	1.099
10	CB2-30% $CaCO_3$	10.90	1.276	1.270
11	CB2-50% $CaCO_3$	12.54	1.469	1.436
12	SB3 cortical bone	13.64	1.695	1.631

holds 11 cylindrical tissue-equivalent inserts (Gammex Inc., WI, USA) with a diameter of 3.5 mm. Figure 1 shows the schematic diagram of the phantom, and table 2 gives the corresponding material properties.

The MOBY digital phantom was used to generate simulated CT images of the head and thorax of a mouse as shown in figure 2. Artificial spherical tumours were created in the brain and lungs with diameters 2 mm and 3 mm, respectively. Reference human tissue compositions were assigned to the phantom. However, it is important to note that this material assignment is arbitrary, and users have the option to incorporate murine tissue compositions, such as those derived in the work of Schyns *et al* (2019). The material assignment for each organ is given in table 3, and the corresponding elemental composition can be found in supplementary table S3.

Using the known material compositions of the phantoms, a ground truth image was created that was used for the evaluation of the CT-based simulations. For example, dose distributions calculated on CT images were validated against the dose computed for an image in which the actual materials are assigned.

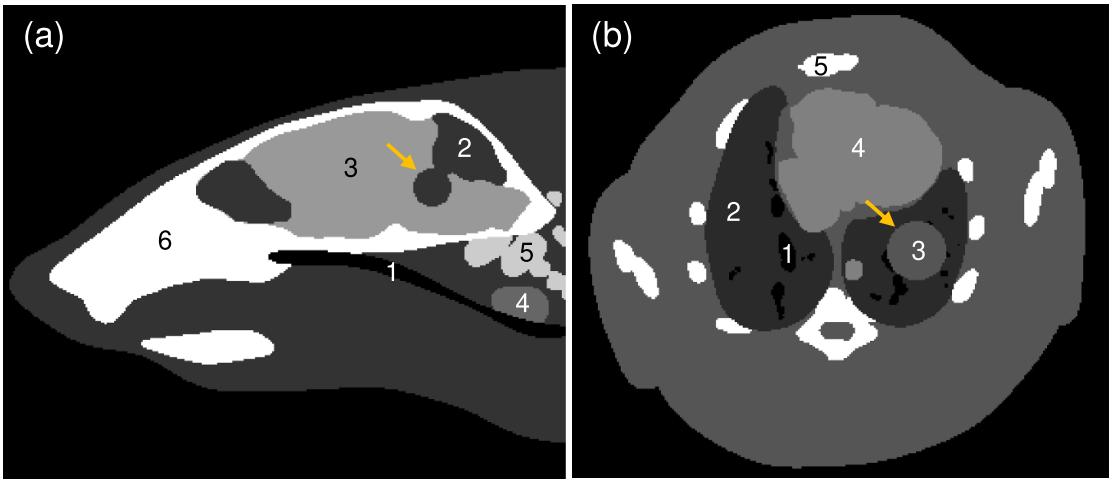

Figure 2. (a) Head and (b) thorax of the MOBY phantom. The numbers correspond to the materials assigned to each organ given in table 3. The spherical tumours with diameters 2 mm (brain) and 3 mm (lung) artificially added in the phantom are indicated by the yellow arrow.

Table 3. Material assignment to organs in the head and thorax of the MOBY phantom. Materials 6–8 in the thorax are not shown in the image slice in figure 2(b). The elemental composition of each material is given in supplementary table S3.

#	Head		Thorax	
	Organs	Material assignment	Organs	Material assignment
1	Esophagus	Air[a]	Airsacs	Air[a]
2	Body, cerebellum, olfactory bulb, tumour	Soft tissue[a]	Lung	Lung-inflated[c]
3	Cerebral cortex, brainstem, striatum, rest of the brain	Brain[b]	Body, tumour	Soft tissue[a]
4	Thyroid	Thyroid[c]	Heart	Heart[b]
5	Spinal cord	Vertebral column (C4)[d]	Ribs, spine, bones	Ribs (2nd, 6th)[d]
6	Skull	Cranium[d]	Liver	Liver[b]
7			Stomach	Stomach[c]
8			Gall bladder	Gall bladder[b]

[a] Geant4 material database (www.fe.infn.it/u/paterno/Geant4_tutorial/slides_further/Geometry/G4_Nist_Materials.pdf).
[b] ICRP Publication 110 (Menzel *et al* 2009).
[c] Woodard and White (1986).
[d] White *et al* (1987).

2.1.4. Validation of x-ray CT imaging with TOPAS

A full MC CBCT simulation of the mini-calibration phantom, including x-ray scattering, was performed in TOPAS to validate the micro-CT model in fastCAT. The same CT geometry was used with the voxelized model of the mini-calibration phantom as the object and the energy spectrum of 90 kVp x-rays filtered with 2 mm aluminium taken from SpekPy as the source. Projection images were taken at 1° interval over a 360° rotation for 9×10^8 x-rays per projection. To compare fastCAT and TOPAS, the mean Hounsfield Units (HU) of Gammex materials were extracted from a region of interest (ROI) with a diameter of 2.5 mm centred at each rod averaged over 5 mm thickness at the centre of the phantom body.

2.2. X-ray CT-based conversion to proton stopping power

Figure 3 shows the workflow developed to perform the HU to SPR calibration from which material and density assignments were derived for the MC simulations of proton irradiations. First, simulated CT images of the mini-calibration phantom were generated with fastCAT using imaging protocols typically used for small animal imaging. Scans at tube potentials 50 kVp and 90 kVp filtered with 2 mm aluminium were created as it was determined to be the best preclinical DECT combination in a previous study (Schyns *et al* 2017). We also looked into using stronger filtration for the high kV beam to reduce the overlap between the low and high kV spectra for DECT (supplementary figure S1). For this, copper with thickness of 0.32 mm was used, which is a filter also available in the X-RAD 225Cx machine. The 50 kVp + 2 mm Al, 90 kVp + 2 mm Al and 90 kVp + 0.32 mm Cu protocols will be referred to as 50 kV/Al, 90 kV/Al and 90 kV/Cu, respectively, for the rest of the paper.

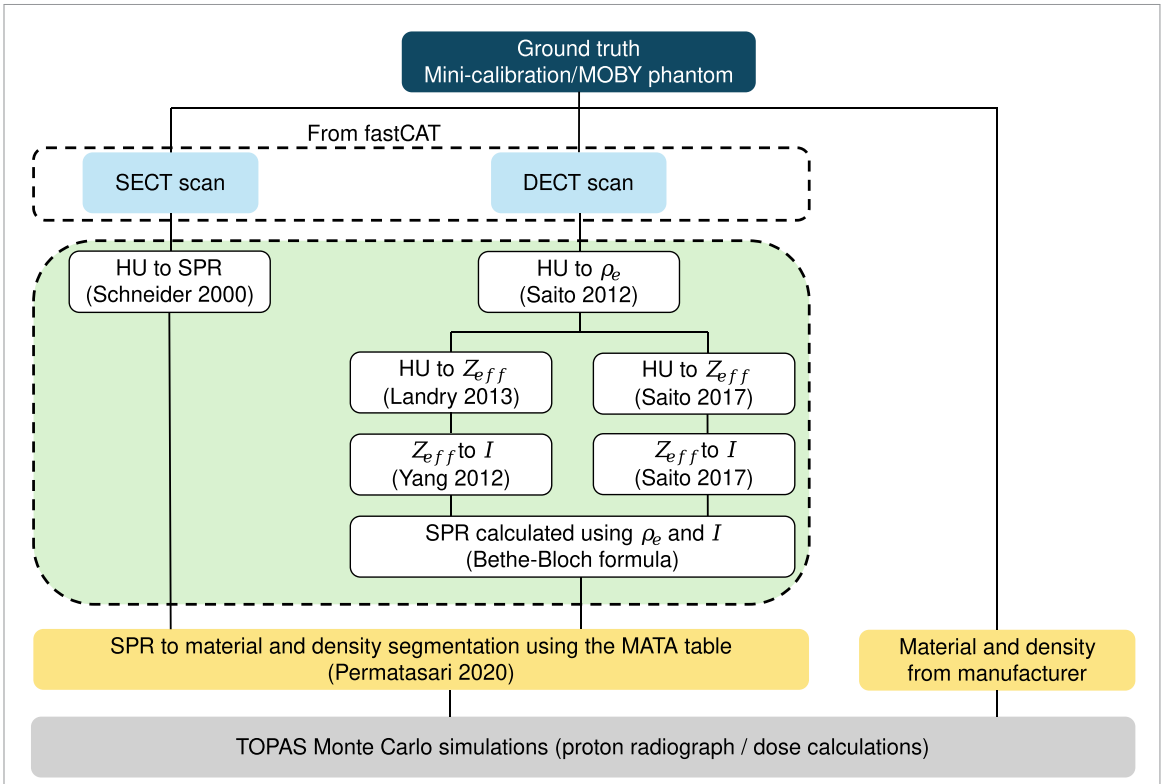

Figure 3. Overview of the simulation framework used in this work. Simulated CT scans of preclinical phantoms were generated with fastCAT. The CT scans are then converted to SPR images through SECT and DECT calibration methods. The MATA table is used to convert SPR datasets to material and density maps, which is the input required for Monte Carlo simulations. Comparisons can be made with the ground truth image, which is the phantom with tissue composition used to generate the CT scans.

The mean HU values of 10 Gammex materials (i.e. materials 3-12 in table 2) given in table S4 in the supplementary file were used to fit parameters for SECT/DECT-to-SPR conversion algorithms. The values were obtained for the same ROI as described in section 2.1.4. The LN-450 lung rod was excluded because it spanned a large HU range due to the clearly noticeable air pockets in the CT images.

2.2.1. Single-energy CT (SECT)

Schneider's stoichiometric calibration method was used to obtain the HU-to-SPR conversion for SECT (Schneider *et al* 2000). The fit parameters k_1 and k_2 were derived using least squares fitting on the Gammex tissue substitutes for the three imaging protocols implemented in this study. The *scipy.optimize.least_squares* function in python was used for the minimization, and the fitting was constrained to positive k-values to avoid physically meaningless parameters. To check our fit procedure, the measured HU values of different phantom materials given in Schneider *et al* (2000) were also analyzed, resulting in the same k-values as reported by them.

Once k-values were determined, the CT numbers of reference human tissues given in Woodard and White (1986) and White *et al* (1987) were calculated. Their corresponding SPR for 100 MeV protons were calculated using the Bethe–Bloch equation given below

$$SPR = \rho_e \frac{\ln\left(\frac{2m_e c^2\beta^2}{I(1-\beta^2)}\right) - \beta^2}{\ln\left(\frac{2m_e c^2\beta^2}{I_w(1-\beta^2)}\right) - \beta^2} \quad (1)$$

where ρ_e is the relative electron density, m_e is the electron mass, c is the speed of light, $\beta = \frac{v}{c}$ is ratio between the speed of the proton and speed of light ($\beta = 0.428$ for 100 MeV protons), I is the mean excitation energy of the material calculated using the Bragg additivity rule, and I_w is the mean excitation energy of water taken as 78 eV (Bär *et al* 2018a). Then, the calibration curve was created by performing linear fits between the predicted HU and SPR on the lung, soft tissue, and bone regions. The k-values resulting from the stoichiometric fit and the root-mean-square deviation of the predicted CT numbers of Gammex materials for the three scan protocols are given in table 4, and an example of the SECT calibration curve is shown in figure 4.

Table 4. Fit coefficients (k_1, k_2) derived from stoichiometric calibration and root-mean-square deviation ($RMSD_{HU}$) of the predicted CT numbers obtained for Gammex materials for imaging protocols: 50 kV/Al, 90 kV/Al, and 90 kV/Cu.

	50 kV/Al	90 kV/Al	90 kV/Cu
k_1	4.33×10^{-14}	1.28×10^{-19}	3.79×10^{-8}
k_2	2.19×10^{-4}	1.23×10^{-4}	8.05×10^{-5}
$RMSD_{HU}$	76	53	12

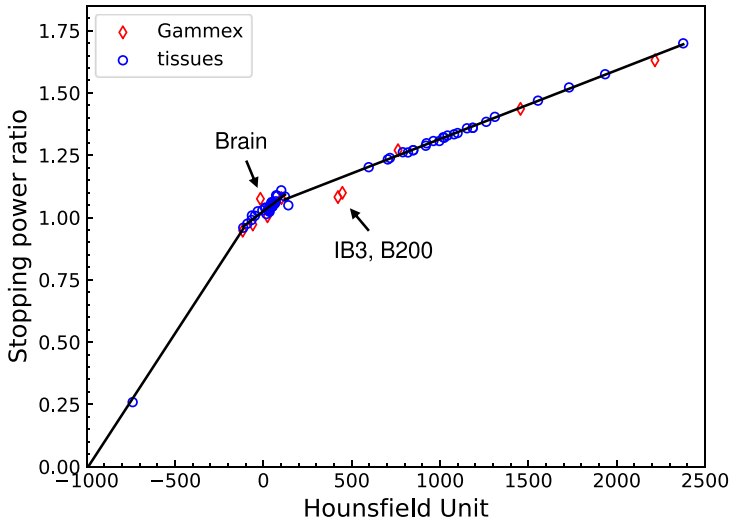

Figure 4. Stoichiometric calibration curve for 90 kV/Cu. The diamonds correspond to Gammex materials with CT numbers from fastCAT simulation, whereas the circles represent datapoints for reference human tissues with CT numbers calculated from the calibration (i.e. using the k_1 and k_2 parameters).

2.2.2. Dual-energy CT (DECT)

The 50 kV/Al–90 kV/Al and 50 kV/Al–90 kV/Cu DECT image pairs were used to extract the relative electron density (ρ_e) and effective atomic number (Z_{eff}). The Z_{eff} was subsequently used to derive the mean excitation energy (I), which, together with ρ_e, is a quantity needed for the calculation of the SPR. Saito's (2012) approach was used to derive the ρ_e from the weighted subtraction of the low and high kV CT numbers. Two different methods were employed to extract the Z_{eff} from the DECT scans. The first is Landry *et al*'s (2013) model which has been implemented in previous small animal DECT studies for x-ray irradiation (Schyns *et al* 2017, Vaniqui *et al* 2017). It uses the ratio of the attenuation coefficients at low and high x-ray energies ($\mu^{\text{low}}_{\text{high}}$) to obtain the Z_{eff} image. Similar to the stoichiometric method, the fit coefficients were obtained using the *scipy.optimize.least_squares* function, wherein the initial estimates for the $\mu^{\text{low}}_{\text{high}}$ were taken from the NIST XCOM database for water at the effective energies of the x-ray spectra (i.e. 25.5 keV, 30.8 keV, and 49.7 keV for 50 kV/Al, 90 kV/Al, and 90 kV/Cu, respectively) and an Z_{eff} of 7.48. This method was combined with Yang *et al*'s (2010) parameterization of $\ln I$ as a function of Z_{eff}. The second is Saito and Sagara's (2017a) method, which was found to be the superior DECT model in a recent study (Niepel *et al* 2021). They proposed a simple formulation that relates the Z_{eff} to the low energy CT numbers and ρ_e. The I-values were calculated using their parameterization of $\ln \frac{I}{I_w}$ as a function of Z_{eff}. For both methods to derive the Z_{eff}, $m = 3.3$ was used (Landry *et al* 2013). The fitting methods were validated by using the values from tables 1, 3, and table 1 given in Saito (2012), Landry *et al* (2013), and Saito and Sagara (2017a), respectively. The largest deviation for ρ_e and Z_{eff} from our calibration are 0.1% and 1.6% from the original values, respectively.

Table 5 gives the fit coefficients for each method derived from the DECT scans used in this work. The fit results for Yang *et al* (2010) and Saito and Sagara's (2017b) parameterization of I are provided in table 6. These coefficients were obtained from fitting of the I and Z_{eff} of reference human tissues (Woodard and White 1986, White *et al* 1987), which were calculated using Bragg's additivity rule.

2.2.3. SPR to material conversion

To perform MC simulations, the elemental composition and mass density of each voxel in the CT are needed instead of the SPR. The SECT- and DECT-based SPR predictions were converted to composition and density using the MATerial Assignment (MATA) table of Permatasari *et al* (2020). The MATA table contains 40 reference human tissues each of which falls in a predefined SPR interval to facilitate material assignment. The

Table 5. Fit coefficients obtained from the DECT images of the mini-calibration phantom following section 2.2.2.

Method	Parameter	50 kV/Al–90 kV/Al	50 kV/Al–90 kV/Cu
Saito (ρ_e)	α	1.68	0.76
	a	0.981	0.976
	b	0.982	1.003
Landry (Z_{eff})	A_{50kV}	1	1
	B_{50kV}	3.923×10^{-3}	1.921×10^{-18}
	C_{50kV}	3.641×10^{-4}	6.508×10^{-4}
	A_{90kV}	1.241	1.402
	B_{90kV}	2.120×10^{-3}	4.586×10^{-21}
	C_{90kV}	2.440×10^{-4}	3.462×10^{-4}
Saito and Sagara (Z_{eff})	γ_L	2.92	2.88

Table 6. Fit coefficients for the parameterization of the mean excitation energy (I).

Method	Parameter	Soft tissue	Bone
Yang	a	0.124	0.100
	b	3.377	3.329
Saito and Sagara	c_1	0.316	0.075
	c_0	0.054	0.115

density is determined using a linear relationship with the SPR. In this way, calibration-specific methods to generate material and density maps from the SPR dataset are avoided.

2.2.4. SPR accuracy

The SPRs of Gammex materials predicted from both SECT and DECT calibration were compared to the theoretical values (SPR_t) calculated using the Bethe-Bloch equation for 100 MeV protons (table 2). For each insert, the mean SPR (SPR_{mean}) was extracted over a cylindrical ROI (same as section 2.1.4) and the deviation was calculated as follows

$$\Delta SPR\ (\%) = \frac{SPR_{mean} - SPR_t}{SPR_t} \times 100. \tag{2}$$

The root mean square deviation ($RMSD$) as given below

$$RMSD\ (\%) = \sqrt{\frac{\sum_{i=1}^{N} (\Delta SPR_i)^2}{N}} \tag{3}$$

was also estimated for each calibration procedure for the $N = 10$ inserts considered in this work.

To get an idea on how the SPRs obtained with the different calibration methods contribute to range errors, proton radiographs of CTs converted to SPR maps using the SECT and DECT methods were calculated and compared to the proton radiograph of the ground truth phantom (i.e. geometry with actual elemental composition assigned). Only the SECT (90 kV/Cu) and DECT (50 kV/Al–90 kV/Al) protocols that showed the least SPR deviation for Gammex materials were used in the comparison. The MOBY thorax was chosen because it is a highly heterogeneous region with air/soft tissue/bone interfaces. The simulated proton radiographs were obtained in TOPAS by irradiating the CT with a 100 MeV parallel proton beam and scoring the residual kinetic energy of protons (KE_{res}) exiting the volume on a 2D air slab placed directly behind it.

A separate simulation was also performed to obtain the calibration curve relating the residual kinetic energy to the water equivalent thickness (WET). Following the same simulation setup, the KE_{res} was scored for water slabs of increasing thickness in increments of 1 mm. The average value in a 1 cm ROI at the centre of the scorer was used for the WET calibration. The resulting calibration curve is given in supplementary figure S2.

The final analysis is expressed in terms of the range error, which is quantified as the pixel-by-pixel difference of the WET map of the calibrated CT from the WET map of the ground truth phantom ($\Delta WET = WET_{CT} - WET_{GT}$).

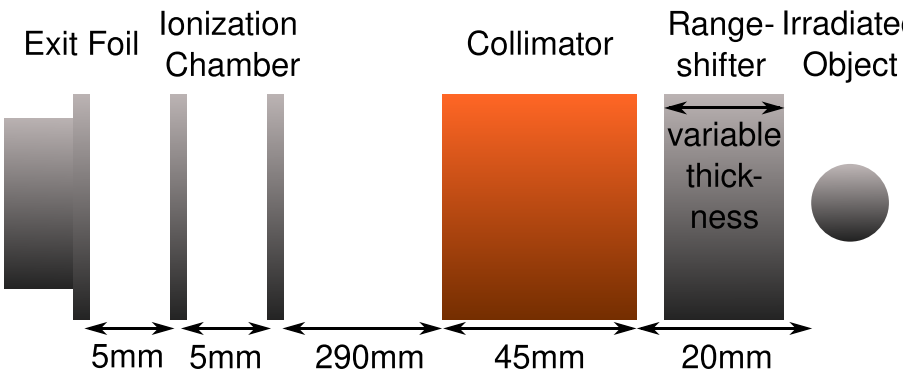

Figure 5. Schematic diagram of the final part of the IMPACT beamline.

2.3. Dose calculations

To assess the impact of SECT and DECT calibration methods on the dose distribution in animals, irradiation plans were created using the open-source matRad treatment planning system (TPS) (Wieser *et al* 2017). Section 2.3.1 details the creation of the matRad beam model for small animal irradiations used in this work. Section 2.3.2 describes the parameters used to create a pencil beam scanning (PBS) plan for a tumour in the middle of a mouse brain, and section 2.3.3 explains the criteria for dosimetric evaluation.

2.3.1. Beam modelling

MatRad requires a database containing the integral depth dose (IDD) curves and lateral size of the beam as a function of depth for a series of proton energies. A realistic model of a 66.5 MeV proton beam for pencil beam irradiations at the IMPACT beamline for small animal radiation biology research at UMCG-PARTREC, which is currently under construction, was created following these steps:

1. The settings of the beam line magnets were optimized with the ion optics code TRANSPORT (Rohrer 2007), requiring a dispersion free beam waist in both transverse planes at the centre of the 1 mm diameter, 45 mm length brass collimator used to shape the beam. The initial transverse emittance in the calculations was obtained from measurements close to the cyclotron exit.
2. The magnet settings obtained from the TRANSPORT calculation and the initial emittance were then used as input for MC particle tracking simulations with the Geant4-based BDSIM toolkit (Nevay *et al* 2020). The initial phase space file containing the phase space coordinates (i.e. XYZ position, XY momentum components, and kinetic energy) of the individual particles included the transverse-longitudinal correlations introduced by the extraction system of the cyclotron. In the simulation, the interactions of the particles with the vacuum exit window, the foils of the ionization chamber, the air traversed by the particles, and the collimator are taken into account as shown in figure 5. The output of the simulation is a phase space file that contains the phase space coordinates of the particles that have passed the collimator. The lateral penumbra (20%–80%) of the beam at the collimator exit was 0.33 mm, while its divergence was 16 mrad. This phase space file was used as the source for a subsequent simulation in TOPAS, wherein the 3D total energy deposit in water was scored for 1×10^6 primaries. From this 3D distribution, the quantities needed for the beam model were derived.
3. To generate data for lower energies without performing additional simulations, we pulled back the Bragg curve of the 66.5 MeV proton beam in increments of 0.2 mm. This method gives the correct beam properties provided the range shifter is right in front of the irradiated object and is water equivalent in terms of multiple scattering.

2.3.2. Irradiation plans

The irradiation plan was created for the MOBY phantom with a 2 mm diameter spherical tumour in the brain as shown in figure 2(a). The planning target volume (PTV) was created by adding a safety margin of 0.2 mm around the tumour to account for setup uncertainties (Verhaegen *et al* 2011, Ford *et al* 2017). The PBS plan was made following these steps:

1. First, an initial plan was calculated using the pencil beam algorithm (PBA) of matRad. Since the PBA requires the SPR for the computation of WET, the theoretical SPR (i.e. SPR calculated with the

Bethe–Bloch formula using the elemental composition assigned to the organs) was assigned to the MOBY head phantom and used as the input in the calculation. This initial PBA plan evaluation was used to determine the optimal beam arrangement in order to achieve a prescribed dose of 10 Gy in the PTV. A lateral spot spacing of 0.8 mm was implemented and a SOBP was created by superimposing several Bragg peaks spaced 0.8 mm apart. These parameters were determined from an initial optimization on a target of the same size in water. For the spot fluence optimization, squared dose deviation from the prescribed dose was used for the PTV and no dose constraints were imposed to organs-at-risk (OAR) such as the brain. The optimization was based on the RBE-weighted dose, which was calculated by applying a constant RBE of 1.1 to the physical dose. The doses presented onwards always refer to the RBE-weighted dose.

2. Second, the dose distribution of each pencil beam was recomputed in TOPAS. The phase space file for 66.5 MeV protons obtained in section 2.3.1 was used as the source in the simulation. To match the beam energies in the PBA plan, a range shifter in the form of water slabs was explicitly simulated in TOPAS with thicknesses equal to the difference between the range of the original 66.5 MeV beam and the range in water of the proton beam at each energy layer. The dose was scored in the ground truth MOBY phantom using the *DoseToMedium* scorer. The ground truth is the geometry with material compositions given in table 3 assigned (i.e. the same materials used to generate the corresponding CTs in fastCAT).
3. The initial PBA doses were then replaced by the MC-calculated ones. The spot weights were re-optimized in matRad based on the new MC pencil beam doses.

Using the optimized plan, the final dose distribution for the ground truth MOBY head phantom was calculated in TOPAS. The same plan was delivered to the SECT and DECT images to demonstrate how the calibration methods affect the dose distribution. All dose distributions were calculated on a 0.1 mm isotropic dose grid (voxel size of the CT).

2.3.3. Dosimetric evaluation

To assess the accuracy of the dose distributions, dose volume histograms (DVHs) were calculated. The criteria for dose coverage is $V_{95} \geq 98\%$ in the PTV. This requires that at least 98% of the PTV receives 95% of the prescribed dose, which is 9.5 Gy. Local differences were also investigated for OARs by looking into the mean (D_{mean}) and maximum (D_{max}) doses received by the brain (i.e. whole brain excluding the PTV) and cranium. The CT-based dose distributions were also compared to the ground truth by performing 3D global gamma analysis with dose difference (DD)/distance to agreement (DTA) set to 3%/0.1 mm. The range shift ($\Delta R = R_{80,\text{CT}} - R_{80,\text{GT}}$) defined as the difference in the distal range at 80% of the maximum dose was also calculated for 366 line dose profiles over the target area along the beam direction.

3. Results

3.1. Simulated CTs from fastCAT

Figures 6(a) and (b) show the simulated CT images of the mini-calibration phantom generated with fastCAT and with a full CBCT simulation in TOPAS, respectively. The comparison of the line profiles is given in figures 6(c) and (d). Qualitatively, the contrast is very similar between the two CT images. This observation is supported by the good agreement in the mean HU of the Gammex inserts extracted from fastCAT and TOPAS as shown in table 7. The difference in mean HU values between fastCAT and TOPAS were within 11 HU except for the brain (27 HU) and SB3 cortical bone (63 HU). The TOPAS simulations resulted in lower HU values for these materials. For SB3 cortical bone, the cupping artifact is more pronounced in TOPAS leading to a much larger difference. Notably, these two inserts are located in the middle of the phantom. To check whether there is a position dependence of the extracted HU values, an additional simulation was performed wherein the brain and SB3 cortical bone were moved to the outer ring in both the fastCAT and TOPAS simulations (i.e. brain and SB3 cortical bone were switched with the liver and CB2-50%, respectively). The HU difference reduced to 15 HU and 24 HU for the brain and SB3 cortical bone, respectively. In contrast, the deviation for the liver and CB2-50% increased to 15 HU and 26 HU, respectively.

3.2. SPR accuracy of Gammex materials

Figure 7 shows the deviation of the SECT- and DECT-predicted SPR of tissue-equivalent materials from the theoretical value (i.e. SPR calculated using the Bethe–Bloch equation). The corresponding root mean square deviation (RMSD) calculated for each protocol is given in table 8. The brain, IB3 inner bone, and B200 bone inserts exhibited the largest deviation (>5%) for SECT. The datapoints for these three materials can be seen to be farther away from the calibration curve (figure 4) leading to divergent results. Overall, the CT scan

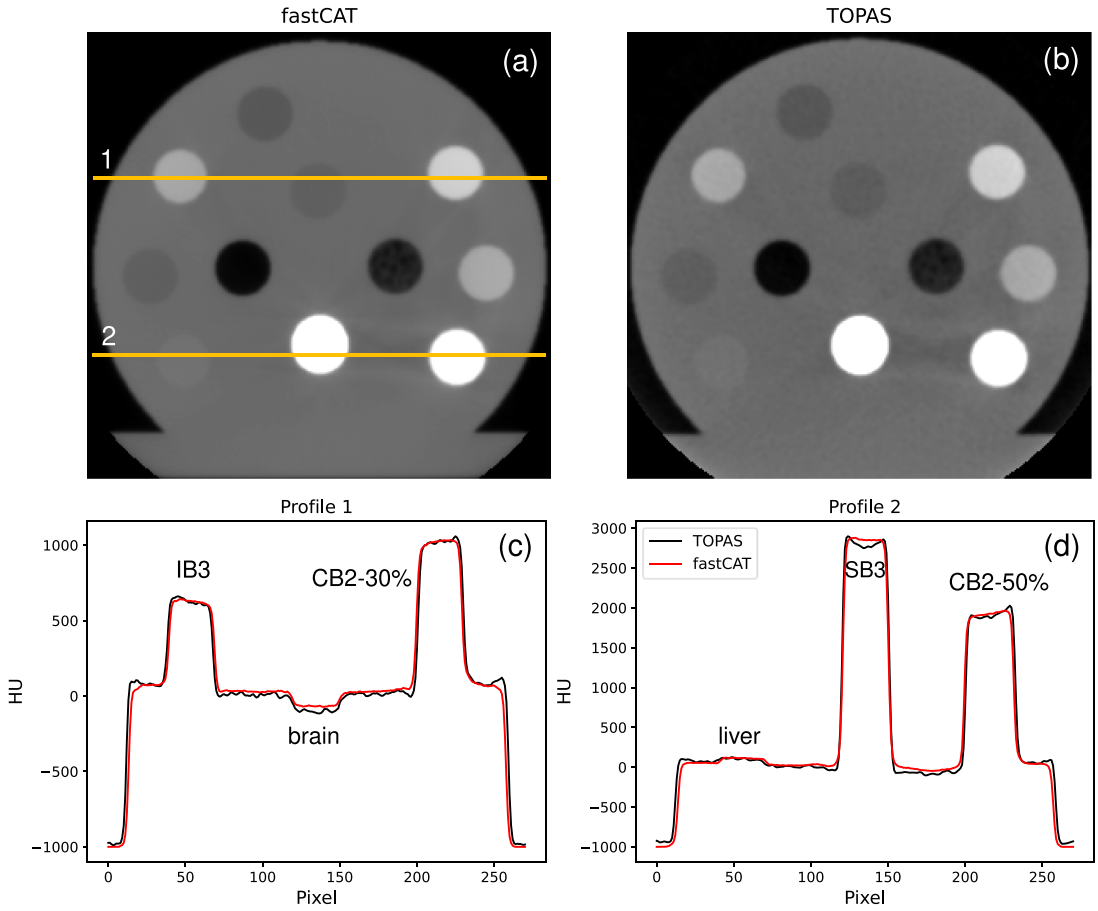

Figure 6. Simulated CT images of the mini-calibration phantom from (a) fastCAT and (b) TOPAS averaged over ten slices for 90 kV/Al. The line profiles indicated by the yellow lines going through IB3/brain/CB2-30% and liver/SB3/CB2-50% inserts are plotted in (c) and (d), respectively.

Table 7. CT numbers of Gammex inserts from TOPAS and fastCAT simulations. The last column lists the difference in Hounsfield units.

Material	HU (TOPAS)	HU (fastCAT)	ΔHU
Adipose	−154	−143	11
Brain	−93	−66	27
Breast	−75	−68	7
Solid Water	40	47	7
Liver	109	115	6
IB3	635	628	7
B200	653	647	6
CB2-30%	1033	1031	2
CB2-50%	1920	1928	8
SB3	2780	2843	64

taken with the 90 kV/Cu yielded smaller SPR deviations and the lowest RMSD among the SECT imaging protocols implemented in this work.

On the other hand, between the two DECT configurations, the 50 kV/Al–90 kV/Cu image pair performed worse than 50 kV/Al–90 kV/Al combination as illustrated in figure 7 and table 8 despite having better separation in the low and high energy spectra. The 50 kV/Al–90 kV/Al DECT yielded prediction errors within ±2%, while 50 kV/Al–90 kV/Cu DECT had larger deviations up to −5.2% for IB3 inner bone with Saito and Sagara's method. Compared to the best SECT case, the 50 kV/Al–90 kV/Al DECT also exhibited better results, which is consistent with findings using clinical imaging protocols (Hudobivnik *et al* 2016, Bär *et al* 2017, 2018b). The two DECT approaches (Landry, Saito and Sagara) at 50 kV/Al–90 kV/Al performed similarly with RMSD of 0.9% and 1.0%, respectively. The 90 kV/Cu SECT and 50 kV/Al–90 kV/Al DECT imaging protocols have then been used to generate SECT and DECT scans of the animal phantom for dose calculations.

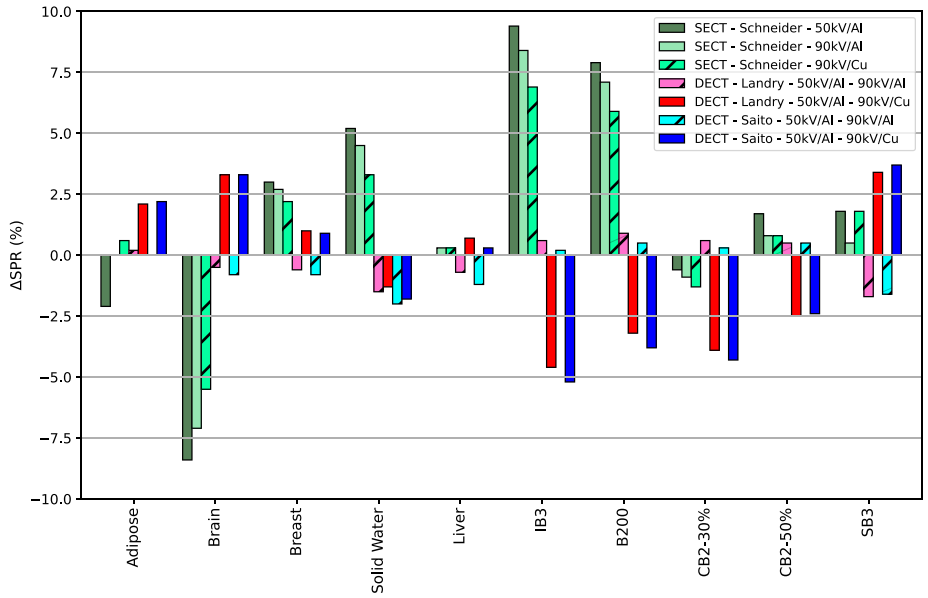

Figure 7. Deviation of the SECT- and DECT-based SPR predictions from theoretical values for Gammex materials. Three imaging protocols were evaluated for SECT: 50 kV/Al, 90 kV/Al, and 90 kV/Cu, whereas two combinations were implemented for DECT: 50 kV/Al–90 kV/Al and 50 kV/Al–90 kV/Cu.

Table 8. Root mean square deviation (RMSD) of SPR values of Gammex materials.

	RMSD (%)		RMSD (%)	
Protocol	SECT	Protocol	DECT Landry	DECT Saito
50 kV/Al	5.2	50 kV/Al–90 kV/Al	0.9	1.0
90 kV/Al	4.5	50 kV/Al–90 kV/Cu	2.9	3.1
90 kV/Cu	3.7			

3.3. Evaluation of range errors in the MOBY thorax

Figure 8 shows the distribution of the range differences for the MOBY thorax between the ground truth and the SECT and DECT calibration methods. The range errors are within ± 1 mm with SECT having larger errors and larger width of the distribution than DECT. The mean $\pm$ standard deviations of the range errors are 0.41 ± 0.16 mm, -0.23 ± 0.12 mm and -0.24 ± 0.12 mm for SECT, DECT-Landry and DECT-Saito, respectively. SECT calibration systematically gives positive range errors, which indicates that WET values are overestimated as a direct consequence of higher SPR predictions than the ground truth. On the contrary, DECT approaches lean towards negative range errors due to overall underestimation of the SPR.

3.4. Impact of CT calibration methods to the dose distribution

Figure 9(a) presents the proton dose distribution for the treatment plan optimized on the ground truth MOBY phantom. To gauge the uncertainty introduced by the calibration methods to the dose calculation, the same plan was delivered to all other cases. The gamma index maps calculated for each calibration method are shown in figure 9(b). Gamma values greater than 1 appear mostly at the distal edge of the PTV and along transitions between soft tissue and bone. Nevertheless, gamma analysis of the SECT-, DECT-Landry-, and DECT-Saito-based dose distributions using 3%/0.1 mm criteria gives passing rates of 99.8%, 99.5% and 99.3%, respectively, which shows that there are no major dosimetric differences with the ground truth.

Figure 10(a) shows a representative dose profile for each case. Both SECT and DECT methods show an underestimation of the dose in regions with high density gradient (tissue/bone interface) as a consequence of CT image blurring at material transitions. This difference is also depicted by γ values > 1 for brain or soft tissue voxels close to the cranium as shown in figure 9(b). Analysis of 366 dose profiles revealed that SECT systematically induces negative range shifts with a median of -0.04 mm, while DECT-Landry and DECT-Saito result in positive range shifts with median values of 0.05 mm and 0.06 mm, respectively. These observations are consistent with the results in section 3.3 indicating that SECT leads to overestimation of the WET, whereas DECT underestimates it. As shown in figure 10(b), the range shifts are below the CT voxel size of 0.1 mm and can be considered negligible.

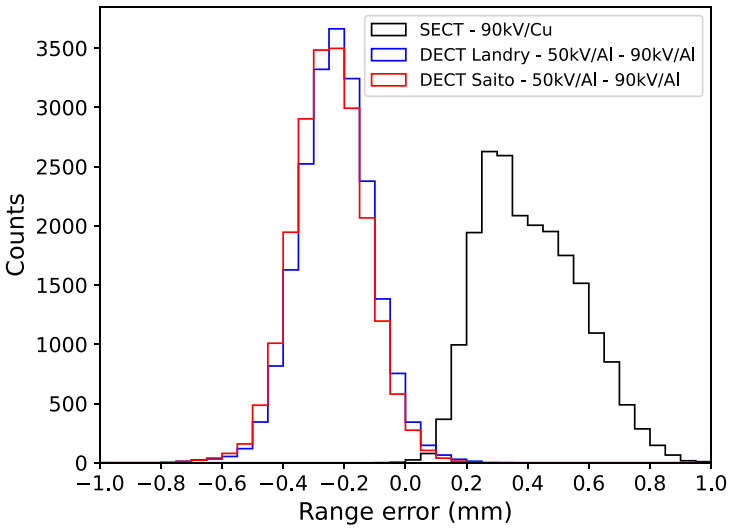

Figure 8. Distribution of the range errors ($\Delta WET = WET_{CT} - WET_{GT}$) for the MOBY thorax resulting from SECT and DECT calibration.

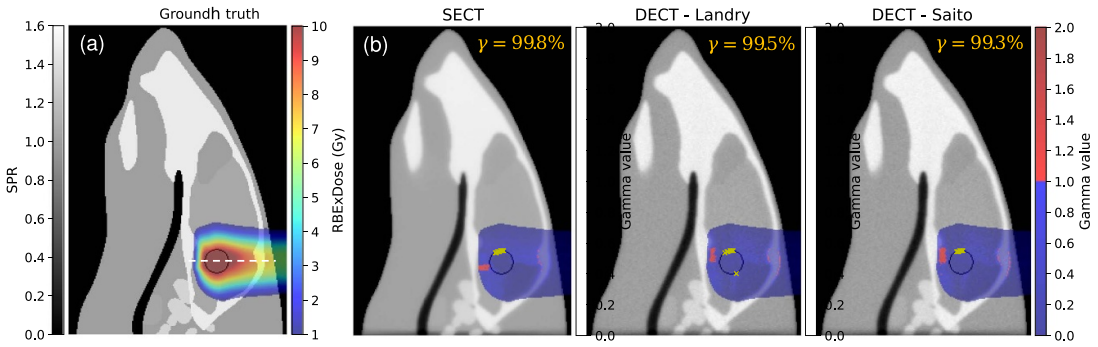

Figure 9. (a) Dose distribution for the treatment plan optimized on the ground truth MOBY head phantom. (b) Gamma index maps (from left to right) for the same plan delivered to SECT, DECT-Landry, and DECT-Saito calibrated images. Tolerance limits (DD/DTA) of 3%/0.1 mm were used for the gamma evaluation. The images onto which the dose distribution and gamma index maps were overlaid are the theoretical SPR (GT) and the SPR distribution resulting from the calibration (SECT and DECT). The black circle illustrates the PTV, and the white line indicates where dose profiles in figure 10(a) were extracted. The yellow 'x' marks show the region in the PTV that failed the dose coverage criteria (i.e. $D_{PTV} < 9.5\,Gy$). DD: dose difference, DTA: distance-to-agreement.

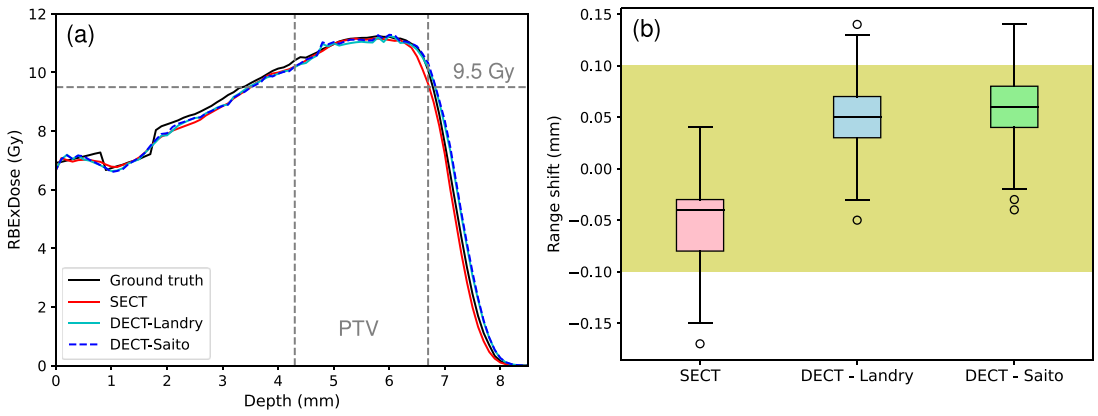

Figure 10. (a) Dose profiles in the beam direction taken at the centre of the PTV indicated by the white line in figure 9(a). (b) Range shift ($\Delta R = R_{80,CT} - R_{80,GT}$) distribution for SECT and DECT calibration methods. Yellow shade indicates the region within one voxel size of the CT (±0.1 mm).

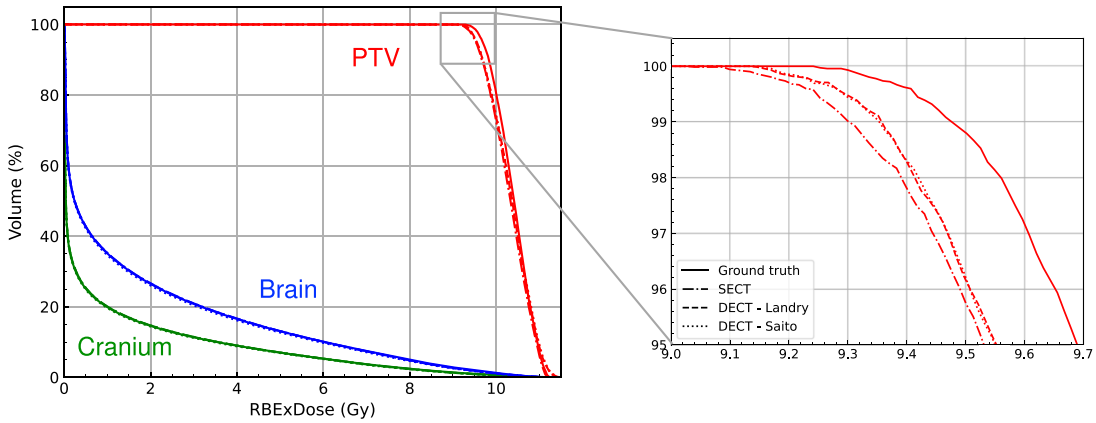

Figure 11. Dose volume histograms calculated for the PTV (prescribed dose: 10 Gy), OAR-brain, and cranium. The solid, dash-dotted, dashed, and dotted lines correspond to the ground truth, SECT, DECT-Landry, and DECT-Saito methods, respectively.

Table 9. Mean (D_{mean}) and maximum (D_{max}) doses, and percent volume that received 95% of the prescribed dose (V_{95}). The last one is only for the PTV.

Structure Metric	Ground truth	SECT	DECT-Landry	DECT-Saito
PTV				
D_{mean} (Gy)	10.4	10.3	10.4	10.4
D_{max} (Gy)	11.3	11.3	11.5	11.5
V_{95} (%)	98.8	95.9	96.3	96.4
Brain				
D_{mean} (Gy)	1.6	1.5	1.6	1.6
D_{max} (Gy)	10.7	10.5	10.7	10.7
Cranium				
D_{mean} (Gy)	0.2	0.2	0.2	0.2
D_{max} (Gy)	7.5	7.8	7.8	7.8

The comparison of the dose volume histograms (DVHs) is displayed in figure 11, and the dose metrics are given in table 9. From the DVH analysis, it can be deduced that adequate PTV coverage (i.e. $V_{95} \geq 98\%$) was not achieved for any of the dose distributions based on the calibrated CTs. Only 95.9%, 96.3%, and 96.4% of the PTV received at least 9.5 Gy for dose distributions computed with SECT, DECT-Landry, and DECT-Saito images, respectively. Voxels that did not meet the dose coverage criteria are mostly towards the edge of the PTV as shown in figure 9(b). Meanwhile, all calibration methods showed no considerable change in the mean and maximum doses to the surrounding OAR.

4. Discussion

This work presents a framework to simulate preclinical proton irradiations. Using this platform, one can design and optimize preclinical setups and gain insights about how it affects the dose distribution in an animal prior to experiments. A major part of this framework is the fastCAT CBCT simulator. It was modified based on the CT geometry of a small animal CT scanner to allow generation of realistic micro-CBCT images. The validity of the fastCAT micro-CBCT scans was demonstrated by the overall good agreement of the CT numbers with full MC simulations. The large discrepancies observed for the brain and SB3 cortical bone inserts are likely caused by beam hardening not being properly handled in fastCAT. These two inserts are located at the centre of the phantom, where the beam has the hardest spectrum (i.e. beam with a higher mean energy due to low energy x-rays being preferentially absorbed). The attenuation coefficient then becomes lower, which results to lower HU values as demonstrated by TOPAS. Although some degree of beam hardening is present in fastCAT CT images, it is less prominent than in those generated by TOPAS. For instance, the cupping artifact in the SB3 cortical bone is much more pronounced in TOPAS than in fastCAT, further supporting this claim.

In comparison to Vaniqui *et al*'s (2017) measured HU values for the same mini-calibration phantom at 50 kV/Al and 90 kV/Al, fastCAT CT numbers are generally higher. The difference can have a number of

explanations. The reconstruction protocol and the ROI from which the HU values were extracted may have been different. The thickness of the CsI detector and the pixel pitch are not the same. Furthermore, the simulations to create the micro-CBCT fastCAT model were rather simplistic. Collimators and other CT components were not included in the modelling, which could generate additional scatter. Also, the angular distribution of the x-ray intensity was assumed to be uniform while in reality, it is not. A better agreement with experimental values could be obtained by replacing the detector response functions in fastCAT and performing more detailed modelling of the micro-CT scanner. Nevertheless, the general trend of the CT numbers of Gammex materials from fastCAT remains consistent with published values.

The potential use of fastCAT simulated CT scans in the evaluation of SECT- and DECT-based SPR estimation for proton treatment planning in animals was also demonstrated in this study. Previous works have already investigated the feasibility of using DECT, and its impact on preclinical x-ray dose distributions (Schyns *et al* 2017, Vaniqui *et al* 2017). Here, we extended it by exploring another energy combination with better spectral separation (50 kV/Al–90 kV/Cu), using different DECT approaches to predict proton SPRs, and comparing the results to the SECT stoichiometric calibration. The best performing SECT and DECT protocols were also evaluated in terms of range errors (from MC-simulated proton radiographs) and proton dose calculation accuracy in animal CTs.

The SECT stoichiometric calibration has been reported to yield uncertainties up to 3.5% (Yang *et al* 2012). For the SECT imaging protocols (i.e. 50 kV/Al, 90 kV/Al and 90 kV/Cu) investigated in this work, the SPR deviation of Gammex materials from theoretical values is in a similar range except for the brain, IB3, and B200 inserts, where deviations $>5\%$ were obtained. These tissue surrogates are seen to lie away from the calibration curve as shown in figure 4. This suggests that these materials have a poor ability in reproducing CT numbers of biological tissues. Although not as pronounced as in our case, the same three materials also slightly deviate from the HU-to-ρ_e calibration curve by Gomá *et al* (2018) for human size phantoms. Since the SPR is directly proportional to the ρ_e, we expect a similar trend. Overall, the SECT SPR prediction improves as the mean energy of the x-ray spectra increases, with 90 kV/Cu yielding the lowest RMSD.

For DECT, two different acquisition settings were investigated. First, the 50 kV/Al–90 kV/Al combination was selected as it was reported by Schyns *et al* (2017) to be the optimal energy combination for preclinical DECT showing the lowest deviation in the estimated ρ_e and Z_{eff} of tissue equivalent materials. However, their work only compared DECT results for different energy combinations filtered by the same material. To improve spectral separation, we implemented a stronger filtration material for the high kV beam (i.e. 90 kVp filtered with 0.32 mm Cu) in SpekPy for the second DECT image pair (50 kV/Al–90 kV/Cu). The idea is that by reducing the overlap between the x-ray spectra, the attenuation values (or CT numbers) from the high and low kV datasets become more independent from each other, which should improve material discrimination (Krauss *et al* 2015). However, contrary to expectation, the 50 kV/Al–90 kV/Cu performed poorly compared to 50 kV/Al–90 kV/Al. To understand these findings, further studies including experimental measurements are needed.

The 50 kV/Al–90 kV/Al DECT pair achieved better results for Gammex materials compared to 90 kV/Cu SECT with deviations within $\pm 2\%$. At this setting, the DECT conversion algorithms of Landry *et al* (2013) and Saito and Sagara (2017a) were comparable although the latter offers a much simpler implementation. Note that several DECT studies using clinical imaging protocols have shown higher accuracy in SPR estimation than what we have achieved (Hünemohr *et al* 2014, Taasti *et al* 2016, Bär *et al* 2017, Landry *et al* 2019). These studies were usually performed with 80–140 kV pair where tin (Sn) filtration was applied for the high tube voltage. This combination exhibits a large x-ray spectra separation, with the high kV image being dominated by Compton effect. Yang *et al* (2011) and Li *et al* (2017) have demonstrated that increasing the energy separation leads to a reduction in the uncertainties associated with SPR estimation. Although it is desirable to use higher energies for DECT, there are limitations for pre-clinical imaging. For instance, the tube voltage of some micro-CT machines only go up until 100 kV. Moreover, higher energies will also result to poorer contrast in the image.

To get an idea on how the SPR uncertainties associated to the CT calibration methods translate into errors in the proton range, we obtained WET maps through proton radiograph simulations on SECT and DECT images. The mean shift and variation in the proton range for DECT are smaller in magnitude than SECT, reflecting DECT's superiority in tissue characterization. However, it is important to note that performing dual energy CT imaging results in additional radiation dose to the animal. Given that DECT is sensitive to image noise, high imaging doses are needed to achieve reliable results. According to Schyns *et al* (2017), the imaging dose required to obtain acceptable DECT calibration is 30 cGy per scan, resulting in a total dose of 60 cGy.

To assess whether DECT indeed offers a potential gain over SECT in treatment planning for animals, proton dose calculations in a mouse brain were performed on SECT and DECT images. The treatment plans were made in matRad to which we have added a beam data library based on a realistic beam model for small

animal irradiations. The proton beam model was obtained from beam transport simulations of a preclinical proton beamline in BDSIM. The accuracy of the dose distributions was evaluated by comparing it with the one calculated based on the ground truth phantom. It is worth highlighting that this is another advantage of the framework. Since the actual tissue compositions used to generate the CTs in fastCAT are known, the same materials can be assigned to the MOBY geometry to create a ground truth image for dose calculations, which is usually lacking when experimental CTs of animals are used.

Results from proton dose calculations on SECT- and DECT-calibrated CT images showed very small differences with the ground truth. A slight shift in the dose profiles can be observed in figure 10(a), wherein the SECT profile is more upstream and DECT ones are deeper than the ground truth. These observations are in line with the proton radiograph results in which SECT was shown to overestimate the SPR, while DECT tend to underestimate it. However, SECT- and DECT-calibrated scans did not really lead to considerable range errors as the average shift is smaller than the dose grid (0.1 mm). There was also negligible effect on the global dose distribution as shown by the high gamma pass ratios (>99% for both SECT and DECT) and on the local mean doses to anatomical structures (both target and OAR). Overall, the results suggest that SECT is sufficient for CT characterization of tissues in a micro-CT for animal proton treatment planning and that the additional dose required for DECT imaging is not warranted.

To our knowledge, this work presented the first dosimetric evaluation of SECT and DECT calibration methods in the context of proton irradiation of small animals. It should however be emphasized that the merit of implementing DECT imaging in proton preclinical practice cannot be evaluated just with a single case study. The investigation was limited to the head, which can be considered relatively homogeneous with only the brain and the skull along the beam direction. The framework can also be used to assess other treatment regions particularly those with more complex geometry and increased heterogeneity to test the robustness of the calibration methods. Additionally, DECT may offer advantages in enhancing contrast in the abdomen, where tissues have similar densities, but may vary in effective atomic number due to the presence of adipose tissue or fat.

5. Conclusion

In this study, we proposed a framework that enables *in silico* modelling of preclinical proton irradiations. The framework will be helpful for the development and optimization of irradiation setups, assessment of the quality of small animal irradiations, and quantification of associated uncertainties in preclinical proton dose delivery. As an example, we have demonstrated how the framework can be used to assess the impact of SECT and DECT calibration methods on proton dose distributions in small animals. Calculations on a mouse brain revealed that treatment planning based on DECT offered no added benefit to the accuracy of the dose.

Data availability statement

The data cannot be made publicly available upon publication because some data are subject to licensing restrictions. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

This work was partially financially supported by a grant from the Dutch Cancer Society (KWF research project 12092).

ORCID iDs

Justin Malimban https://orcid.org/0000-0001-7783-2647
Felix Ludwig https://orcid.org/0009-0004-0803-7761
Danny Lathouwers https://orcid.org/0000-0003-3810-1926
Marius Staring https://orcid.org/0000-0003-2885-5812
Frank Verhaegen https://orcid.org/0000-0001-8470-386X
Sytze Brandenburg https://orcid.org/0000-0002-9887-1642

References

Bär E, Andreo P, Lalonde A, Royle G and Bouchard H 2018a Optimized I-values for use with the Bragg additivity rule and their impact on proton stopping power and range uncertainty *Phys. Med. Biol.* **63** 165007

Bär E, Lalonde A, Royle G, Lu H-M and Bouchard H 2017 The potential of dual-energy CT to reduce proton beam range uncertainties *Med. Phys.* **44** 2332–44
Bär E, Lalonde A, Zhang R, Jee K W, Yang K, Sharp G, Liu B, Royle G, Bouchard H and Lu H M 2018b Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples *Med. Phys.* **45** 48–59
Beyreuther E, Brand M, Hans S, Hideghéty K, Karsch L, Leßmann E, Schürer M, Szabó E R and Pawelke J 2019 Feasibility of proton FLASH effect tested by zebrafish embryo irradiation *Radiother. Oncol.* **139** 46–50
Biguri A, Dosanjh M, Hancock S and Soleimani M 2016 TIGRE: a MATLAB-GPU toolbox for CBCT image reconstruction *Biomed. Phys. Eng. Express* **2** 055010
Bijl H P, van Luijk P, Coppes R P, Schippers J M, Konings A W T and van der Kogel A J 2006 Influence of adjacent low-dose fields on tolerance to high doses of protons in rat cervical spinal cord *Int. J. Radiat. Oncol. Biol. Phys.* **64** 1204–10
Denbeigh J M, Howard M E, Garcia D A, Debrot E K, Cole K C, Remmes N B and Beltran C J 2024 Characterizing proton-induced biological effects in a mouse spinal cord model: a comparison of Bragg peak and entrance beam response in single and fractionated exposures *Int. J. Radiat. Oncol. Biol. Phys.* **119** 924–35
Ford E *et al* 2017 An image-guided precision proton radiation platform for preclinical in vivo research *Phys. Med. Biol.* **62** 43–58
Ghobadi G *et al* 2012 Physiological interaction of heart and lung in thoracic irradiation *Int. J. Radiat. Oncol. Biol. Phys.* **84** e639–46
Gomá C, Almeida I P and Verhaegen F 2018 Revisiting the single-energy CT calibration for proton therapy treatment planning: a critical look at the stoichiometric method *Phys. Med. Biol.* **63** 235011
Howard M E, Denbeigh J M, Debrot E K, Garcia D A, Remmes N B, Herman M G and Beltran C J 2021 Dosimetric assessment of a high precision system for mouse proton irradiation to assess spinal cord toxicity *Radiat. Res.* **195** 541–8
Hudobivnik N *et al* 2016 Comparison of proton therapy treatment planning for head tumors with a pencil beam algorithm on dual and single energy CT images *Med. Phys.* **43** 495
Hünemohr N, Krauss B, Tremmel C, Ackermann B, Jäkel O and Greilich S 2014 Experimental verification of ion stopping power prediction from dual energy CT data in tissue surrogates *Phys. Med. Biol.* **59** 83–96
Kim M M *et al* 2021 Comparison of FLASH proton entrance and the spread-out Bragg peak dose regions in the sparing of mouse intestinal crypts and in a pancreatic tumor model *Cancers* **13** 4244
Kim M M, Irmen P, Shoniyozov K, Verginadis I I, Cengel K A, Koumenis C, Metz J M, Dong L and Diffenderfer E S 2019 Design and commissioning of an image-guided small animal radiation platform and quality assurance protocol for integrated proton and x-ray radiobiology research *Phys. Med. Biol.* **64** 135013
Krauss B, Grant K L, Schmidt B T and Flohr T G 2015 The importance of spectral separation: an assessment of dual-energy spectral separation for quantitative ability and dose efficiency *Invest. Radiol.* **50** 114–8
Landry G *et al* 2019 Technical note: relative proton stopping power estimation from virtual monoenergetic images reconstructed from dual-layer computed tomography *Med. Phys.* **46** 1821–8
Landry G, Seco J, Gaudreault M and Verhaegen F 2013 Deriving effective atomic numbers from DECT based on a parameterization of the ratio of high and low linear attenuation coefficients *Phys. Med. Biol.* **58** 6851–66
Li B, Lee H C, Duan X, Shen C, Zhou L, Jia X and Yang M 2017 Comprehensive analysis of proton range uncertainties related to stopping-power-ratio estimation using dual-energy CT imaging *Phys. Med. Biol.* **62** 7056–74
Menzel H G, Clement C and DeLuca P 2009 ICRP publication 110. Realistic reference phantoms: an ICRP/ICRU joint effort. A report of adult reference computational phantoms *Ann. ICRP* **39** 1–164
Nevay L J *et al* 2020 BDSIM: an accelerator tracking code with particle–matter interactions *Comput. Phys. Commun.* **252** 107200
Niepel K B *et al* 2021 Animal tissue-based quantitative comparison of dual-energy CT to SPR conversion methods using high-resolution gel dosimetry *Phys. Med. Biol.* **66** 075009
O'Connell J and Bazalova-Carter M 2021 fastCAT: fast cone beam CT (CBCT) simulation *Med. Phys.* **48** 4448–58
Parodi K *et al* 2019 Towards a novel small animal proton irradiation platform: the SIRMIO project *Acta Oncol.* **58** 1470–5
Perl J, Shin J, Schumann J, Faddegon B and Paganetti H 2012 TOPAS: an innovative proton Monte Carlo platform for research and clinical applications *Med. Phys.* **39** 6818–37
Permatasari F F, Eulitz J, Richter C, Wohlfahrt P and Lühr A 2020 Material assignment for proton range prediction in Monte Carlo patient simulations using stopping-power datasets *Phys. Med. Biol.* **65** 185004
Poludniowski G, Omar A, Bujila R and Andreo P 2021 Technical note: spekPy v2.0—a software toolkit for modeling x-ray tube spectra *Med. Phys.* **48** 3630–7
Prezado Y *et al* 2017 Proton minibeam radiation therapy spares normal rat brain: long-term clinical, radiological and histopathological analysis *Sci. Rep.* **7** 14403
Prezado Y *et al* 2019 Tumor control in RG2 glioma-bearing rats: a comparison between proton minibeam therapy and standard proton therapy *Int. J. Radiat. Oncol. Biol. Phys.* **104** 266–71
PTCOG 2024 Particle Therapy Co-operative Group: *Particle Therapy Facilities in Clinical Operation* (available at: https://www.ptcog.site/index.php/facilities-in-operation-public)
Rohrer U 2007 PSI graphic transport framework *based on a CERN-SLAC-FERMILAB Version by KL Brown et al* (available at: https://aea.web.psi.ch/Urs_Rohrer/MyWeb/trans.htm)
Saito M 2012 Potential of dual-energy subtraction for converting CT numbers to electron density based on a single linear relationship *Med. Phys.* **39** 2021–30
Saito M and Sagara S 2017a A simple formulation for deriving effective atomic numbers via electron density calibration from dual-energy CT data in the human body *Med. Phys.* **44** 442293–303
Saito M and Sagara S 2017b Simplified derivation of stopping power ratio in the human body from dual-energy CT data *Med. Phys.* **44** 4179–87
Schneider M, Schilz J D, Schürer M, Gantz S, Dreyer A, Rothe G, Tillner F, Bodenstein E, Horst F and Beyreuther E 2024 SAPPHIRE—establishment of small animal proton and photon image-guided radiation experiments *Phys. Med. Biol.* **69** 095020
Schneider W, Bortfeld T and Schlegel W 2000 Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions *Phys. Med. Biol.* **45** 459
Schyns L E, Almeida I P, van Hoof S J, Descamps B, Vanhove C, Landry G, Granton P V and Verhaegen F 2017 Optimizing dual energy cone beam CT protocols for preclinical imaging and radiation research *Br. J. Radiol.* **90** 20160480
Schyns L E, Eekers D B, van der Heyden B, Almeida I P, Vaniqui A and Verhaegen F 2019 Murine vs human tissue compositions: implications of using human tissue compositions for photon energy absorption in mice *Br. J. Radiol.* **92** 20180454
Segars W P, Tsui B M, Frey E C, Johnson G A and Berr S S 2004 Development of a 4-D digital mouse phantom for molecular imaging research *Mol. Imaging Biol.* **6** 149–59

Sørensen B S, Bassler N, Nielsen S, Horsman M R, Grzanka L, Spejlborg H, Swakoń J, Olko P and Overgaard J 2017 Relative biological effectiveness (RBE) and distal edge effects of proton radiation on early damage in vivo *Acta Oncol.* **56** 1387–91

Sørensen B S, Sitarz M K, Ankjærgaard C, Johansen J G, Andersen C E, Kanouta E, Grau C and Poulsen P 2022 Pencil beam scanning proton FLASH maintains tumor control while normal tissue damage is reduced in a mouse model *Radiother. Oncol.* **175** 178–84

Suckert T *et al* 2021 Late side effects in normal mouse brain tissue after proton irradiation *Front. Oncol.* **10** 598360

Taasti V T, Petersen J B, Muren L P, Thygesen J and Hansen D C 2016 A robust empirical parametrization of proton stopping power using dual energy CT *Med. Phys.* **43** 5547

van Luijk P *et al* 2015 Sparing the region of the salivary gland containing stem cells preserves saliva production after radiotherapy for head and neck cancer *Sci. Transl. Med.* **7** 305ra147

Vaniqui A, Schyns L, Almeida I P, van der Heyden B, van Hoof S J and Verhaegen F 2017 The impact of dual energy CT imaging on dose calculations for pre-clinical studies *Radiat. Oncol.* **12** 181

Verhaegen F, Granton P and Tryggestad E 2011 Small animal radiotherapy research platforms *Phys. Med. Biol.* **56** R55

White D R, Woodard H Q and Hammond S M 1987 Average soft-tissue and bone models for use in radiation dosimetry *Br. J. Radiol.* **60** 907–13

Wieser H-P *et al* 2017 Development of the open-source dose calculation and optimization toolkit matRad *Med. Phys.* **44** 2556–68

Woodard H Q and White D R 1986 The composition of body tissues *Br. J. Radiol.* **59** 1209–18

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

Yang M, Virshup G, Clayton J, Zhu X R, Mohan R and Dong L 2011 Does kV-MV dual-energy computed tomography have an advantage in determining proton stopping power ratios in patients? *Phys. Med. Biol.* **56** 4499

Yang M, Zhu X R, Park P C, Titt U, Mohan R, Virshup G, Clayton J E and Dong L 2012 Comprehensive analysis of proton range uncertainties related to patient stopping-power-ratio estimation using the stoichiometric calibration *Phys. Med. Biol.* **57** 4095