**IPEM**
Institute of Physics and
Engineering in Medicine

**PAPER • OPEN ACCESS**

# A systematic approach for calibrating a Monte Carlo code to a treatment planning system for obtaining dose, LET, variable proton RBE and out-of-field dose

To cite this article: Johannes Tjelta *et al* 2023 *Phys. Med. Biol.* **68** 225010

View the article online for updates and enhancements.

You may also like

- Implementation of a double scattering nozzle for Monte Carlo recalculation of proton plans with variable relative biological effectiveness
Lars Fredrik Fjæra, Daniel J Indelicato, Camilla H Stokkevåg et al.

- Dual-layer collimator for improved spatial resolution in SPECT with CZT camera: an analytical and Monte Carlo study
Nasreddine Boutaghane, Michel Hesse, Boualem Bouzid et al.

- The FLUKA Monte Carlo code coupled with the NIRS approach for clinical dose calculations in carbon ion therapy
G Magro, T J Dahle, S Molinelli et al.

# Physics in Medicine & Biology

**IPEM**
Institute of Physics and
Engineering in Medicine

**PAPER**

# A systematic approach for calibrating a Monte Carlo code to a treatment planning system for obtaining dose, LET, variable proton RBE and out-of-field dose

**Johannes Tjelta**[1,2,*] , **Lars Fredrik Fjæra**[2,3] , **Kristian Smeland Ytre-Hauge**[2] ,
**Camilla Grindeland Boer**[1]  and **Camilla Hanquist Stokkevåg**[1,2]

1   Department of Oncology and Medical Physics, Haukeland University Hospital, Bergen, Norway
2   Department of Physics and Technology, University of Bergen, Bergen, Norway
3   Department of Oncology and Medical Physics, Oslo University Hospital, Oslo, Norway
*   Author to whom any correspondence should be addressed.

**E-mail:** johannes.tjelta@uib.no

## Abstract

*Objective.* While integration of variable relative biological effectiveness (RBE) has not reached full clinical implementation, the importance of having the ability to recalculate proton treatment plans in a flexible, dedicated Monte Carlo (MC) code cannot be understated . Here we provide a step-wise method for calibrating dose from a MC code to a treatment planning system (TPS), to obtain required parameters for calculating linear energy transfer (LET), variable RBE and in general enabling clinical realistic research studies beyond the capabilities of a TPS. *Approach.* Initially, Pristine Bragg peaks (PBP) were calculated in both the Eclipse TPS and the FLUKA MC code. A rearranged Bortfeld energy-range relation was applied to the initial energy of the beam to fine-tune the range of the MC code at 80% dose level distal to the PBP. The energy spread was adapted by dividing the TPS range by the MC range for dose level 80%–20% distal to the PBP. Density and relative proton stopping power were adjusted by comparing the TPS and MC for different Hounsfield units. To find the relationship of dose per primary particle from the MC to dose per monitor unit in the TPS, integration was applied to the area of the Bragg curve. The calibration was validated for spread-out Bragg peaks (SOBP) in water and patient treatment plans. Following the validation, variable RBE were calculated using established models. *Main results.* The PBPs ranges were within ±0.3 mm threshold, and a maximum of 5.5% difference for the SOBPs was observed. The patient validation showed excellent dose agreement between the TPS and MC, with the greatest differences for the lung tumor patient. *Significance. A* procedure for calibrating a MC code to a TPS was developed and validated. The procedure enables MC-based calculation of dose, LET, variable RBE, advanced (secondary) particle tracking and more from treatment plans.

## Introduction

Treatment planning systems (TPS) for radiotherapy of cancer enable rapid and detailed optimization and dose calculation for treatment volumes and surrounding organs at risk (OAR). However, TPSs have limitations in their flexibility of calculating proton relative biological effectiveness (RBE) and out-of-field doses. The increasing use of proton therapy has brought on new clinical challenges, and in particular the issue of variable RBE has received considerable attention throughout the past years (Underwood *et al* 2022). Proton RBE is well known to vary with tissue type, fractionation, total dose, dose rate and linear energy transfer (LET) (Willers *et al* 2018, Paganetti 2022), but is currently substituted by a fixed RBE of 1.1 during treatment planning (Dalrymple *et al* 1966, ICRU 2007, Paganetti *et al* 2019). The factor 1.1 represents an increase in the biological effectiveness by 10% compared to conventional photon-based radiotherapy. The clinical implementation of variable RBE is

not yet readily available, much due to uncertainty in the existing flurry of models and clinical impact (Rorvik *et al* 2018). Certain TPSs have introduced graphical processing unit (GPU)-based Monte Carlo (MC) dose and LET calculation software (Schreuder *et al* 2019). Whereas this development represents a step forward, it is not as adaptable and informative as a dedicated MC code, where implementation and calculation of state-of-the-art variable RBE is feasible (Antonie and Schneider 2019) with high flexibility and transparency to the user. A so-called general purpose MC code also allows for inclusion of advanced scoring of neutral particles i.e. neutron and prompt gamma radiation, which is related to out-of-field doses distant from the target (Antonie and Schneider 2019) with implications for late effects of cancer treatment. Furthermore, studies involving advanced particle tracking beyond TPS capabilities are central in the advancement of proton therapy, e.g. in the context of radioprotection (Englbrecht *et al* 2021) range and dose verification systems (Smeets *et al* 2012, Ytre-Hauge *et al* 2019). Also, for such studies/purposes it is of great advantage to be able to perform clinically realistic MC simulations.

While the gold standard MC has been widely used in studies of both variable RBE and out-of-field doses, accurate dose calculation in patients requires calibration to a TPS or dose measurements. The calibration can be performed by synchronizing range and energy-spread, correction for density and relative proton stopping power, and lastly an absolute dose conversion across the two methods. For studies of LET, RBE and out-of-field doses, several groups have calibrated MC to TPSs or measurements, but there is not an established standard, and the literature does not offer a complete and detailed description of the full procedure (Grassberger *et al* 2014, Fracchiolla *et al* 2015, Kozłowska *et al* 2019). In particular for calibration of the energy spread, Fiorini *et al* (2018) and Grevillot *et al* (2011) estimated energy spread by simulating test samples nearby an assumed value in a trial and error approach. These trial and error approaches for energy spread can be time consuming, especially if performed for several energies. In the current work, we provide a mathematical method to directly predict energy and energy-spread needed to calibrate the MC code to the TPS. Moreover, we present a semi-analytical step-wise method of the complete calibrating procedure, with the specific aim of enabling calculation of LET, RBE and out-of-field doses.
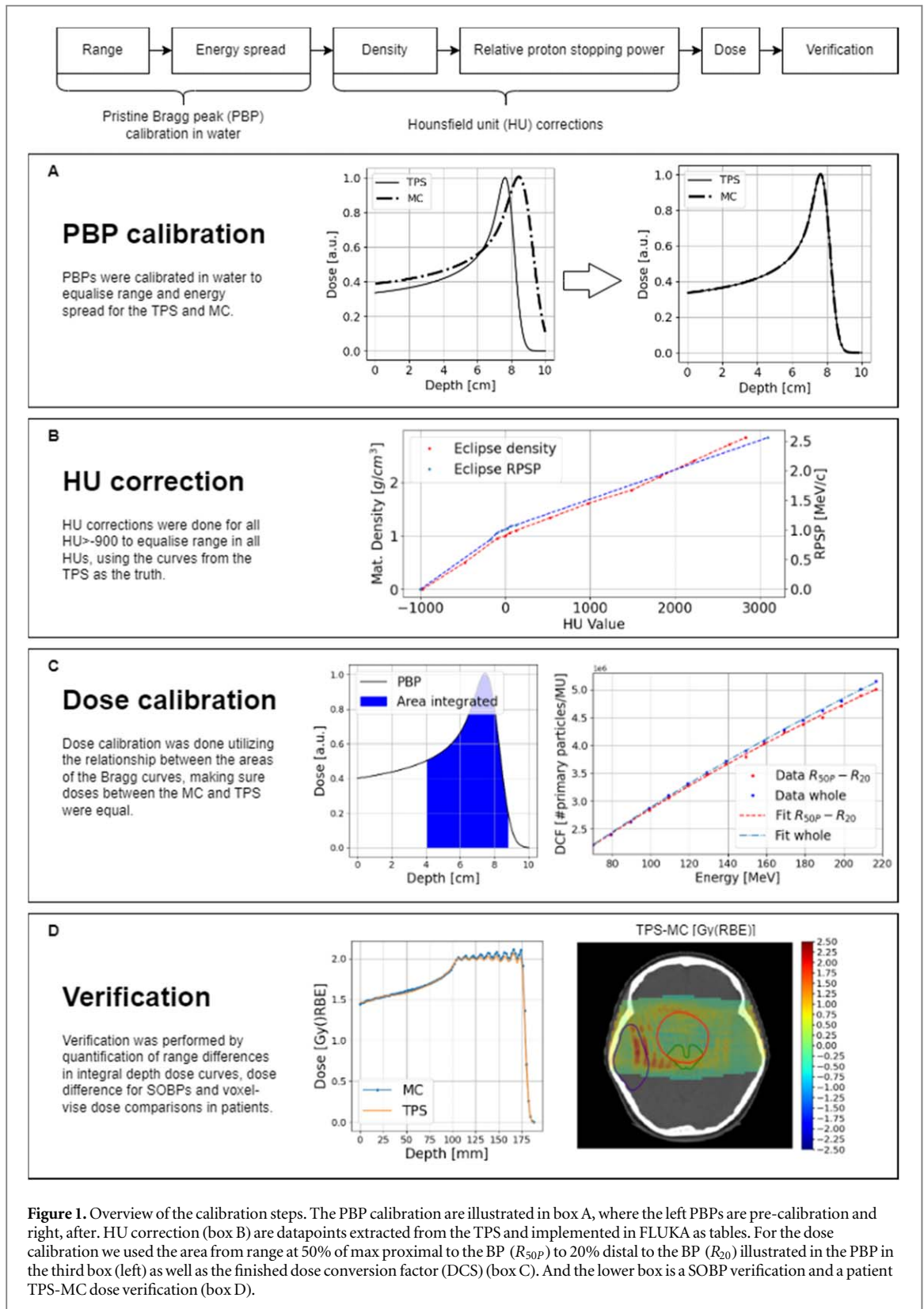
## Materials and methods

The Eclipse TPS (Varian Medical Systems, Palo Alto, California, US/ Siemens Healthineers, artificial ProBeam360 data, pencil beam convolution superposition (PCS) algorithm) and the FLUKA general purpose MC code (Ferrari PRS *et al* 2005, Böhlen *et al* 2014, Battistoni *et al* 2016) were utilized throughout this calibration. An overview of the step-wise calibration procedure is presented in figure 1. Calibrating range and energy spread was initially conducted in water for pristine Bragg peaks (PBP) with 10 MeV energy increments by adjusting the mean energy and the distal dose fall off for each beam, respectively. Range and energy spread were implemented as separate functions in FLUKA. Corrections to density and relative proton stopping power (RPSP) were set for the borders of Hounsfield Unit (HU) intervals. These were both implemented as lookup tables in FLUKA. The calibration phantoms were set at a constant HU, this is done to minimize differences between TPS and MC simulations. An absolute dose calibration from the MC simulated data to files comparable to the TPS was calculated and implemented in our workflow.

The TPS planning files were imported into FLUKA for recalculation using an in-house framework (Fjaera *et al* 2017). To calibrate range, energy spread and RPSP, virtual phantoms of dimensions $100 \times 10 \times 10$ cm$^3$ were generated using pydicom. For range and energy spread calibrations, the HU of the phantom was set to a RPSP $= 1.0001$ in the TPS, while in FLUKA, the phantom was defined as water. Water/water equivalent was used in the calibration by convention (Paganetti *et al* 2008), and all dose calculations were done as dose to water. RPSP calibrations were done for different densities (HUs $\in [-900, 3072]$). For RPSP calibrations at very low densities (HU $< -830$), an extended phantom with a length of 400 cm was used.

The highest achievable dose scoring resolution from the TPS is 1 mm, and this was replicated for the MC calculations.

Each patient field was initialized in a dedicated simulation using a phase space compiled from the DICOM plan file (RT). The phase space contained each spot with its energy, position, spot size, spot direction and spot weight, where the spot weight was used to determine the fractional number of primaries each spot exhibits/ contains based on the total amount primaries. The orientation of the spots was calculated with the gantry angle exported from the RT file and transformed to be within the xy-plane in FLUKA. Spot positions were extracted in the same manner as direction, but in addition the distance to the source distance was also extracted (usually 1.5 m) and transformed to be within the same *xy*-plane. Spot sizes were extracted and given the same orientation as the direction in the phase space file; however, these were divided by a factor $2\sqrt{2\ln 2}$ to obtain the standard deviation and multiplied with a normal distribution function in FLUKA.

**Figure 1.** Overview of the calibration steps. The PBP calibration are illustrated in box A, where the left PBPs are pre-calibration and right, after. HU correction (box B) are datapoints extracted from the TPS and implemented in FLUKA as tables. For the dose calibration we used the area from range at 50% of max proximal to the BP ($R_{50P}$) to 20% distal to the BP ($R_{20}$) illustrated in the PBP in the third box (left) as well as the finished dose conversion factor (DCS) (box C). And the lower box is a SOBP verification and a patient TPS-MC dose verification (box D).

**Range**

The ranges obtained from the TPS and the MC were calibrated to coincide by changing the initial beam energy for the MC. MC PBPs were individually adjusted to the TPS PBPs in water. At 80% of maximum dose distal to the PBP ($R_{80}$), the fluence of initial protons is ideally 50% (Gottschalk 2012), making it independent of the energy spread and therefore a suitable parameter to use for range calibration. PBP calibration ranges are illustrated in appendix A, figure A.1. Since the PBP data are discrete, a MatLab (Gajewski 2023) curve fitting code using the Bortfeld approximation (Bortfeld 1997) was applied, and the $R_{80}$ distal to the PBP was determined

from the curve fit. The difference in $R_{80}$ between the MC and the TPS, denoted $\Delta R_{80}$, was then used to derive the new required energy ($E_{MC}$) for the MC such that $\Delta R_{80}$ would be less than 0.3 mm. This was derived through a rearranged energy-range relation (Bortfeld 1997, Fjaera *et al* 2020) as given by:

$$E_{MC} = \left( \frac{\Delta R_{80}}{\alpha} - E_{TPS}^{p} \right)^{\frac{1}{p}}. \tag{1}$$

The constants $\alpha = 0.0022$ cm/MeV, and $p = 1.77$ were used throughout this calibration but are estimates given by the Bortfeld approximation (Bortfeld 1997). $E_{TPS}$ is the nominal beam energy from the TPS. The difference between PBP ranges calculated for the TPS and MC are denoted $\Delta R_d$, where $d$ is the depth at a certain percentage of the max. $E_{MC}$ was calculated for each PBP in 10 MeV intervals. Once $E_{MC}$ was obtained for the energy intervals, a curve fit of the data was calculated, modeling $E_{TPS}$ to $E_{MC}$ ($E_{TPS} \in [70 \text{ MeV}, 218 \text{ MeV}]$). A 2nd-degree polynomial was used to obtain the fit and was implemented in our in-house software for automatically converting a RT plan to files executable by FLUKA such that the nominal beam energy in FLUKA give equal range outcome to the TPS.

### Energy spread

The energy spread was approximated with a Gaussian distribution in FLUKA, with a mean at approximately zero and a standard deviation as a percentage of the beam energy for each initial energy. For each simulated initial proton, a random number from the Gaussian distribution was generated and added to the initial proton energy closely resembling the energy spread of the TPS. The energy spread was calibrated for PBPs in the aforementioned phantom. Distal dose fall-off of 80%–20% in a PBP ($R_{8020}$) i.e. the distance where the dose decreases from 80% to 20% can be described by range straggling and energy spread as described by B. Gottschalk was used (Gottschalk 2012):

$$R_{8020} = 1.3(\sigma_{range}^2 + \sigma_{beam}^2)^{1/2}. \tag{2}$$

Range straggling is fluctuations in path lengths; hence, the range has a statistical error, $\sigma_{range}$ accounting for this. $\sigma_{beam}$ is the additional contribution from the beam energy spread. The range straggling is assumed to be equivalent for the TPS and MC for equal energies ($\sigma_{range} \approx 0.012 \times$ range). For beams of equal energies, we can approximate $\sigma^2 = \sigma_{range}^2 + \sigma_{beam}^2$. By dividing $R_{8020}$ calculated with PBPs from the TPS (DDF$_{TPS}$) by the distal dose fall-off calculated in FLUKA (DDF$_{MC}$), the difference in energy spread can be calculated with equation (3):

$$\sigma_{MC} \frac{DDF_{TPS}}{DDF_{MC}} = \sigma_{TPS}. \tag{3}$$

$\sigma_{MC}$ in equation (3) is the assumed energy spread implemented at the start of the calibration in the MC. The energy spread was initially projected to be linearly decreasing with increasing beam energy, as an initial energy spread were needed. The assumed energy spread started at 1.7% of the initial energy at 70 MeV, to 0.2% at 218 MeV, since a decreasing energy spread was similar to results presented by Grassberger *et al* (2015). The $\sigma_{TPS}$ is the energy spread equal to the TPSs. Similarly, to the energy adjustments, a 2nd-degree polynomial was used to obtain the fit and implemented in FLUKA.

### Density and stopping power corrections

The graphical user interface FLAIR (Vlachoudis 2009) was used with the FLUKA MC tool. FLAIR implements several features including a DICOM module which translates voxel-based computer tomography (CT) files i.e. DICOM, into voxel files readable by FLUKA based on HUs. The CT-based voxel file requires a conversion for each individual voxel with its HU value to a tissue equivalent material as well as specific density and RPSP. It would be computationally intensive to assign a material for each HU, hence common practice is to approximate with HU intervals with the Schneider *et al* (2000) parametrization, which segments the CT into 24 materials of defined elemental compositions. The density is set for the midpoint of the interval, using a density-to-HU curve extracted from the TPS (figure A.2). Densities for the outer edges of the interval were calculated by using the same density-to-HU curve extracted from the TPS. FLUKA then calculates a gradient within the interval, assigning densities for individual HU values.

Additionally, the TPS has a HU to RPSP curve (figure A.2) that were used to calculate ranges for the TPS for different HU. With the range in water ($R_{80}$) obtained from the range calibrations, we calculated the TPS range for HUs for 150 MeV by subsequently divided $R_{80}$ in water by the RPSP from the HU to RPSP curve, resulting in projected range at a certain HU (PR$_{TPS}$), displayed in figure A.3. In FLUKA, we simulated a 150 MeV (corrected for energy) beam in phantoms with equal HU to the TPS calculations (figure A.3). The projected range for the MC (PR$_{MC}$) for the specific HU was then estimated with the Bortfeld approximation (Bortfeld 1997) in the same manner as with the range. This was done for both border HUs in each interval to match range between TPS and MC for most HUs. The RPSP-corrections i.e. relative range correction (*RRC*), which are implemented in the

MC, as the relationship between the projected ranges (equation (4)), the MC and TPS are assumed to have equal HU

$$\frac{\mathrm{PR_{MC}}}{\mathrm{PR_{TPS}}} = \mathrm{RRC}. \tag{4}$$

**Absolute dose MU conversion**

The method implemented in FLUKA calculates an average dose per primary particle, whereas the TPS calculates dose to water for each monitor unit (MU). Hence, to obtain absolute doses in FLUKA, we performed the conversion by integrating the PBP depth dose curves. The depth at 50% proximal (p in the subscript) to the PBP ($R_{50P}$) to the depth at 20% distal to the PBP ($R_{20}$) and the whole Bragg curve was numerically integrated for the PBPs calculated in FLUKA and the TPS using the trapezoidal rule for each energy. Dividing the TPS area ($A_{\mathrm{TPS}}$) by the MC area ($A_{\mathrm{MC}}$) divided by MUs for each energy interval yields the dose conversion factor (DCF).

$$\left(\frac{A_{\mathrm{TPS}}}{A_{\mathrm{MC}}}\right)/\#\mathrm{MU} = \mathrm{DCF}. \tag{5}$$
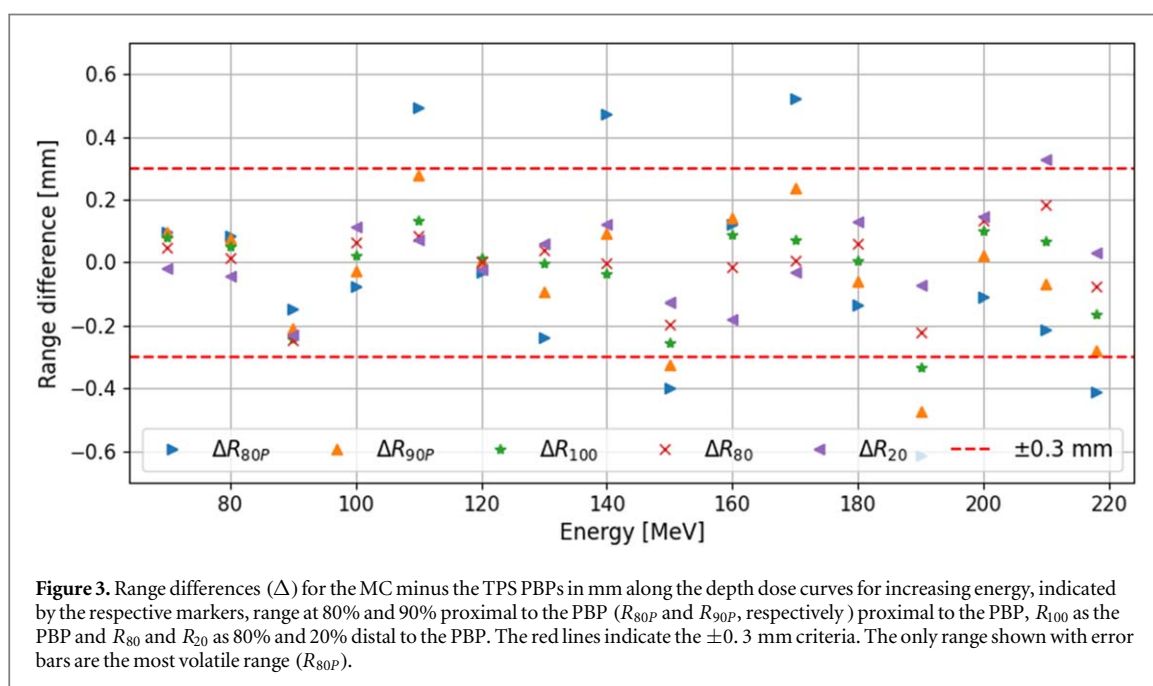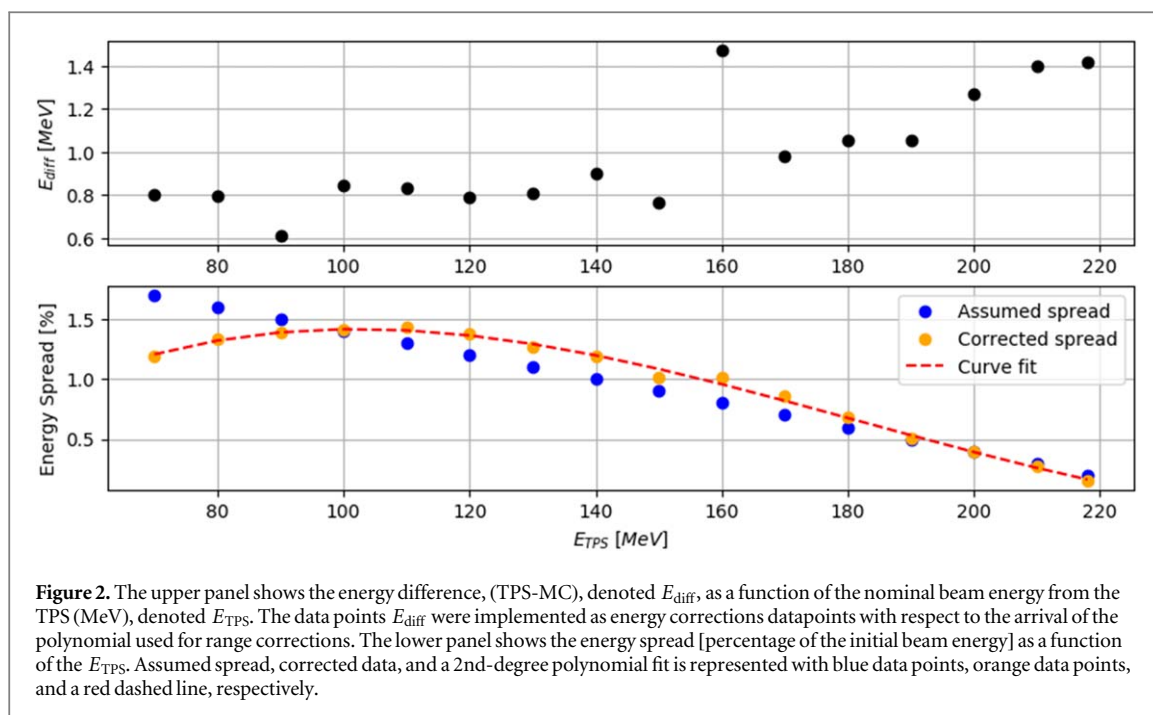
This relation was only energy-dependent and was approximated with a 2nd-degree polynomial and implemented in our internal framework for converting FLUKA output files to DICOM. Several segment integrations were tested and currently the best dose conformity in the clinical target volume (CTV) was achieved using the $R_{50P}$ to $R_{20}$ area.

**Verification**

Verification of the calibration in water was done for ranges at specific percentages of the PBPs, their range differences needs to be within the standard deviation of a uniform distribution with a scoring size of 1 mm as this is the best a calibration can be. A standard deviation of a uniform distribution is defined as $\pm\Delta z/\sqrt{12} = \pm 1/\sqrt{12}$ mm $\approx \pm 0.3$ mm (Illowsky 2020) and are the limit for range implemented. The assumption was a uniform hit registration across the scoring voxel. The difference in range at 80% and 90% proximal to the PBP ($\Delta R_{80P}$ and $\Delta R_{90P}$), the PBP ($\Delta R_{100}$) and at 80% and 20% distal to the peak ($\Delta R_{80}$ and $\Delta R_{20}$) were calculated (ranges illustrated in figure A.1). Spot sizes were also verified at the isocenter in air for equal energies as range and energy-spread, with the size measured as full width half maximum (FWHM). Differences in spot size was represented as a percentage of the original (TPSs) FWHM, for comparison a 10% spot difference is utilized in current clinical treatment planning (Schwarz *et al* 2016). The calibration was verified for spread-out Bragg peaks (SOBPs) in water for low (70–120 MeV), medium (120–160 MeV) and high (160–210 MeV) energies. Dose verification was also performed in three patient cases: a pediatric brain tumor (ependymoma), a lung cancer patient and a prostate cancer patient. The lung plan includes a posterior–anterior field, which is not the most common practice, but are included for the purpose of stretching the calibration to its limits. For the patient treatment plans, dose-volume histograms (DVH), gamma test(s) (Low and Dempsey 2003) and visual dose inspection layer by layer were performed to confirm the equivalence between the MC and the TPS. The gamma test was performed with distance to agreement (DTA) equal to 1, 2 and 3 mm, a dose difference (DD) of 1%, 2% and 3% and a dose threshold of 10% of the prescribed dose for each patient. The gamma test was performed with Slicer (Fedorov *et al* 2012) and a sub package, Slicer RT (Csaba Pinter and Slicer 2021). Variable RBE, LET and weighted neutron dose calculations were performed to demonstrate the utility of the MC implementation.

**RBE**

We used our implemented framework to calculate variable RBE with the McNamara linear model (MCN) (McNamara *et al* 2015) and Rørvik non-linear weighted model (RORW) (Rorvik *et al* 2018). The dose averaged LET ($\mathrm{LET_d}$) was calculated to water as described by Fjæra *et al* (2020) with only the primary and secondary protons as per Fjæra *et al* (2017). For the RBE-weighted dose models in this work all $\alpha/\beta$ values used were generic values. $\alpha/\beta = 10$ Gy was used for the tumor in the pediatric (Paganetti 2017) and lung patient (Klement *et al* 2020) plan, while an $\alpha/\beta = 1.5$ Gy was used for the tumor for the prostate cancer (Brenner and Hall 1999) plan. $\alpha/\beta = 2.1$ Gy was used for all organs not mentioned (McNamara *et al* 2015). For the brainstem and right lung $\alpha/\beta = 3.0$ Gy was used (Borst *et al* 2010, Kondziolka *et al* 2015), while an $\alpha/\beta = 2.0$ Gy (Kehwar 2005) was used for the heart. Finally the neutron dose to the pediatric patient was calculated using ICRP103 (ICRP 2007) weighting factors

**Figure 2.** The upper panel shows the energy difference, (TPS-MC), denoted $E_{\text{diff}}$, as a function of the nominal beam energy from the TPS (MeV), denoted $E_{\text{TPS}}$. The data points $E_{\text{diff}}$ were implemented as energy corrections datapoints with respect to the arrival of the polynomial used for range corrections. The lower panel shows the energy spread [percentage of the initial beam energy] as a function of the $E_{\text{TPS}}$. Assumed spread, corrected data, and a 2nd-degree polynomial fit is represented with blue data points, orange data points, and a red dashed line, respectively.



**Figure 3.** Range differences ($\Delta$) for the MC minus the TPS PBPs in mm along the depth dose curves for increasing energy, indicated by the respective markers, range at 80% and 90% proximal to the PBP ($R_{80P}$ and $R_{90P}$, respectively ) proximal to the PBP, $R_{100}$ as the PBP and $R_{80}$ and $R_{20}$ as 80% and 20% distal to the PBP. The red lines indicate the $\pm 0.3$ mm criteria. The only range shown with error bars are the most volatile range ($R_{80P}$).

## Results

The difference in required energy adjustment systematically increased by 0.6–1.4 MeV when compared to the nominal energy. The ranges at 160 MeV showed greatest deviation and had the largest differences, although all ranges were within the defined threshold (figure 2). The adjusted energy spread function showed an overall similar trend to the initially assumed trend but with less energy spread at low energies (figure 2). The energy spread ranged from approximately 0.2%–1.4% of the beam energy, depending on the proton energy. In general, the PBP depth difference across the TPS and MC increased with energy (figure 3). Only a single data point, 210 MeV, used in the calibration of range and energy spread failed to meet the $\pm 0.3$ mm criteria ($\Delta R_{20}$). Distal to the BP 110, 140, 150, 170, 190 and 218 MeV failed to meet the $\pm 0.3$ mm for $R_{80P}$ and 150 and 190 MeV for $R_{90P}$.

Spot sizes were verified for energies from 70 MeV to 218 MeV. Without any change in spot size between the TPS and MC a maximum difference of 5.3% was achieved (figure A.4). Additionally, most of the difference was
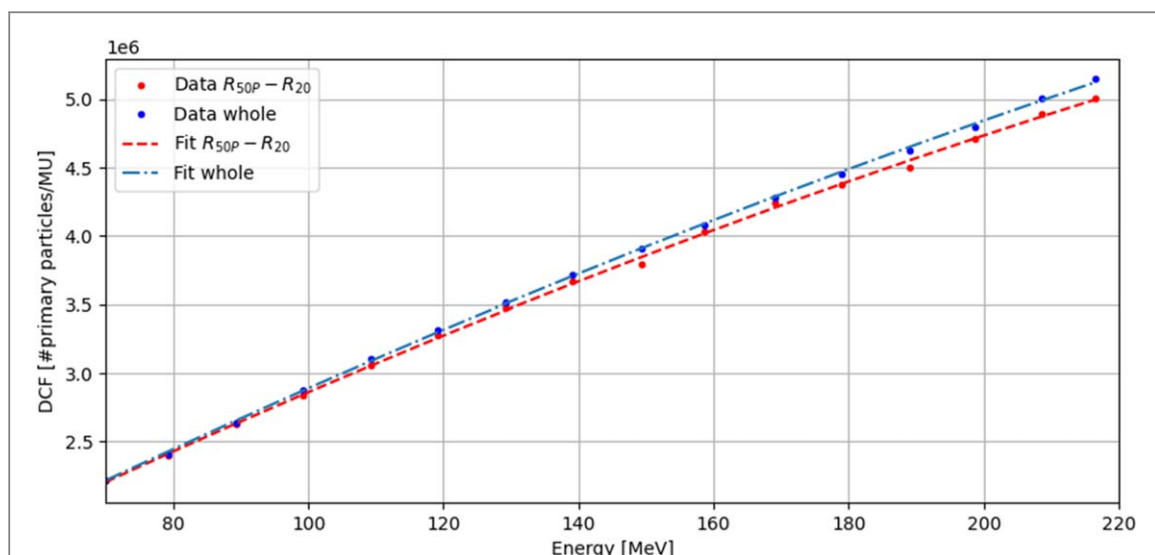
**Figure 4.** Number of primary particles needed to achieve one MU plotted against specific energies. The two different data sets with fits represent the area of the integrated PBPs, where the blue data points and fit provide the integration of the whole track, while the red data points and fit contain integration from $R_{50P}$ to $R_{20}$, only.
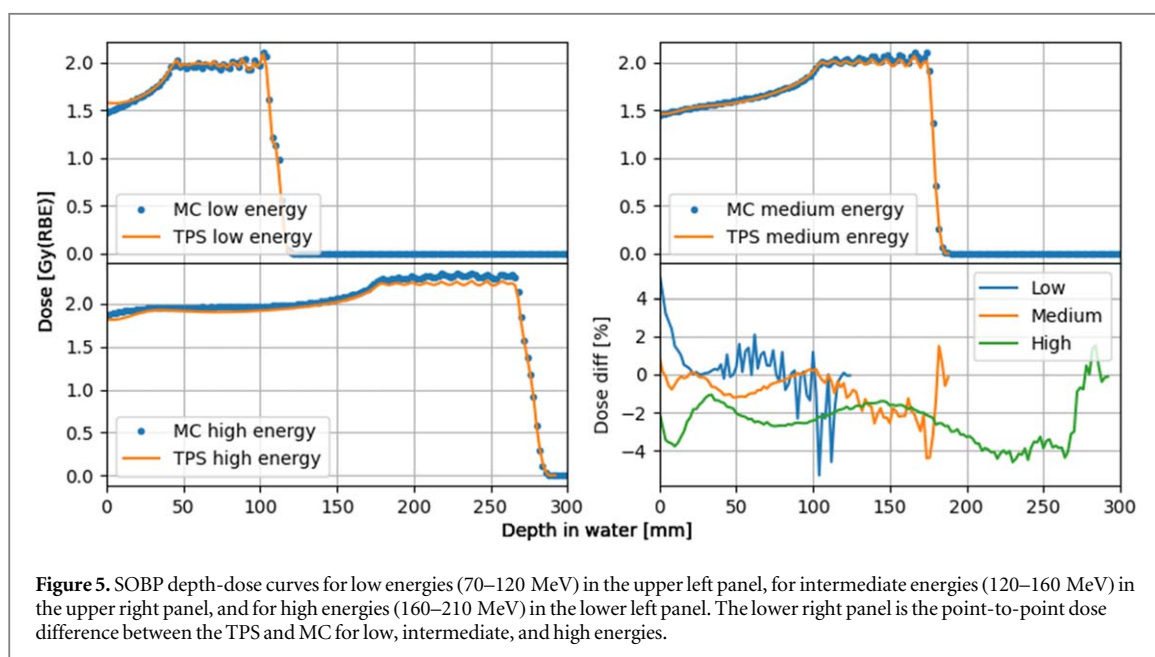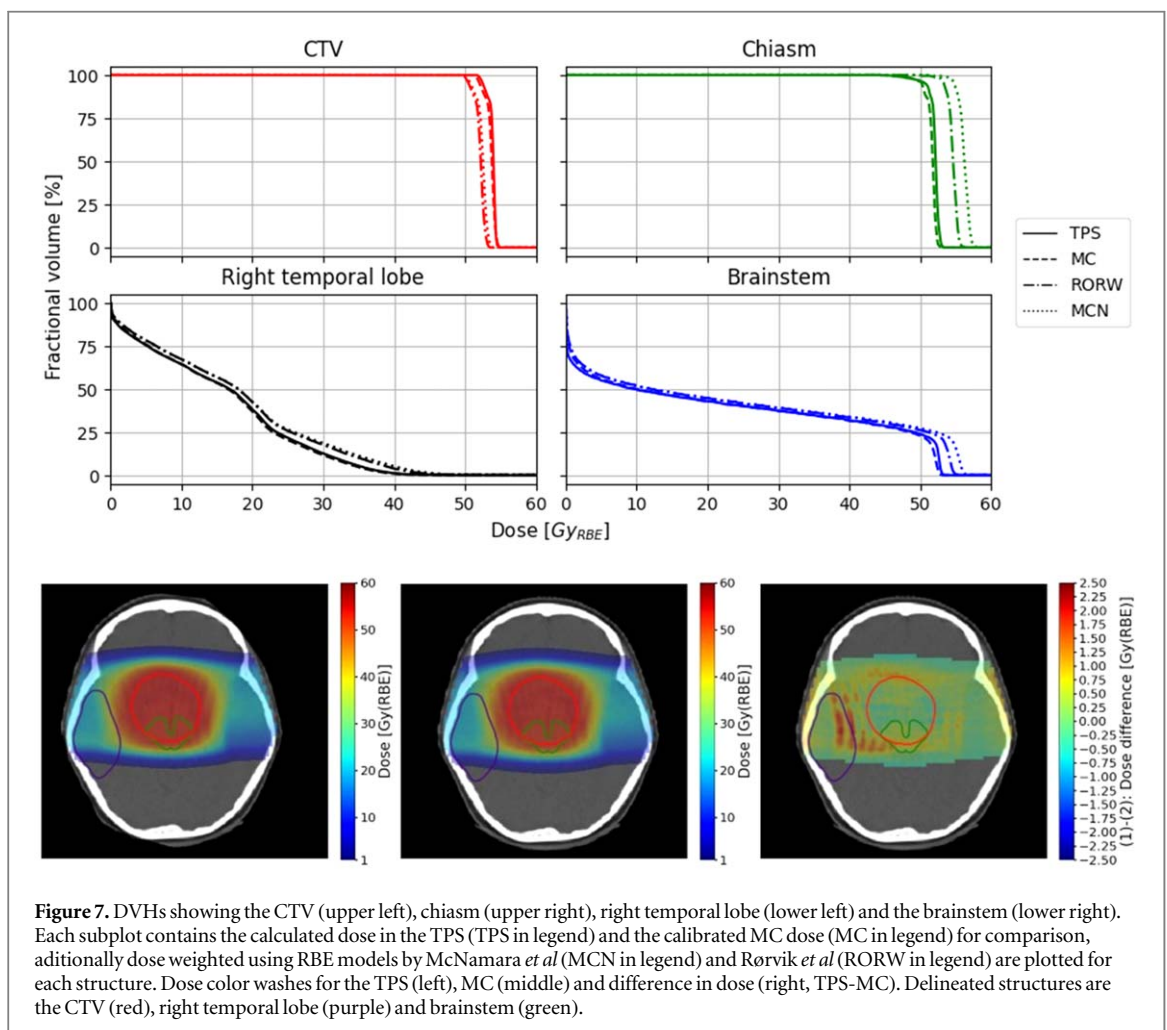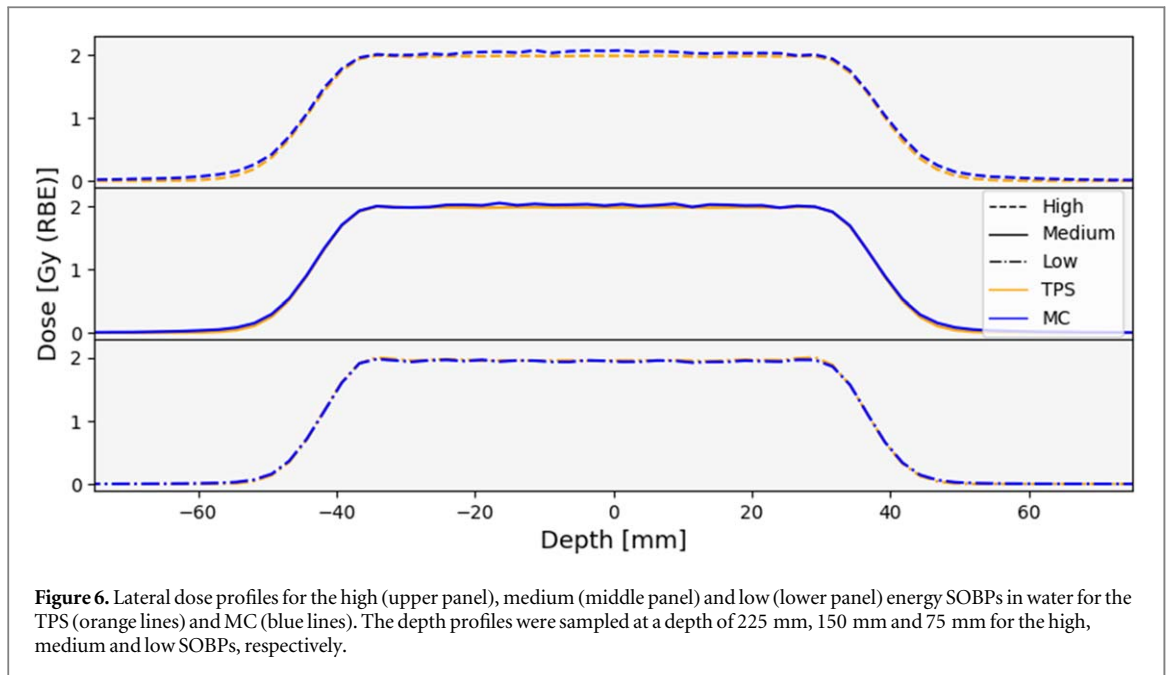


**Figure 5.** SOBP depth-dose curves for low energies (70–120 MeV) in the upper left panel, for intermediate energies (120–160 MeV) in the upper right panel, and for high energies (160–210 MeV) in the lower left panel. The lower right panel is the point-to-point dose difference between the TPS and MC for low, intermediate, and high energies.

under 4% and the average FWHM difference were 0.27 mm. The estimated number of primary particles per MU was dependent on the integrated area of the PBPs (figure 4). For higher energies, a lower number of primaries was estimated when calibrating with $R_{50P} - R_{20}$. As a measure of agreement between the MC and TPS, both models were tested with SOBPs and gamma test pass rates for the different patients. The integration $R_{50P}-R_{20}$ showed the best agreement, hence, this method was utilized further throughout this work.

The SOBPs (figure 5), resulted in a moderate dose difference at the beam entry for low energy SOBPs. The largest dose difference was 5.5% located at the distal end of the low energy SOBP. The differences were 4.4% in dose at the depth of largest difference, for both the medium and high energy SOBPs.

The lateral dose profiles of the low, medium and high SOBPs had a maximum dose difference of 1.3%, −2.8% and −4.1% for TPS-MC at the center of the SOBPs. Furthermore, a deviance of 0.5 mm in the lateral spread was observed for the high energy SOBP at 10% of the field edge (figure 6). The lateral dose profiles were also examined (figure A.5) at the entrance of each SOBP, with a maximum dose difference of −0.4%, −1.5% and −1.9%.

The dose comparison for the pediatric brain tumor case showed a minimal dose difference in the CTV and low to negligible dose difference in OARs (figure 7). $D_{50\%}$ values were 53.9 Gy for the TPS and 53.8 Gy for the

**Figure 6.** Lateral dose profiles for the high (upper panel), medium (middle panel) and low (lower panel) energy SOBPs in water for the TPS (orange lines) and MC (blue lines). The depth profiles were sampled at a depth of 225 mm, 150 mm and 75 mm for the high, medium and low SOBPs, respectively.



**Figure 7.** DVHs showing the CTV (upper left), chiasm (upper right), right temporal lobe (lower left) and the brainstem (lower right). Each subplot contains the calculated dose in the TPS (TPS in legend) and the calibrated MC dose (MC in legend) for comparison, aditionally dose weighted using RBE models by McNamara *et al* (MCN in legend) and Rørvik *et al* (RORW in legend) are plotted for each structure. Dose color washes for the TPS (left), MC (middle) and difference in dose (right, TPS-MC). Delineated structures are the CTV (red), right temporal lobe (purple) and brainstem (green).

MC in the CTV. The right temporal lobe had a $D_{50\%}$ dose difference of only 0.2 Gy. When observing the DVHs in figure 7, a lower effective dose for the target volume (high $\alpha/\beta$) and a higher effective dose for the OARs (low $\alpha/\beta$) can be seen for the variable RBE models.

**Figure 8.** Weighted neutron dose for the pediatric patient. Deliniated in the dose distribution are the CTV (red), right temporal lobe (blue) and brainstem (green). The DVH is calculated for the CTV and OARs.
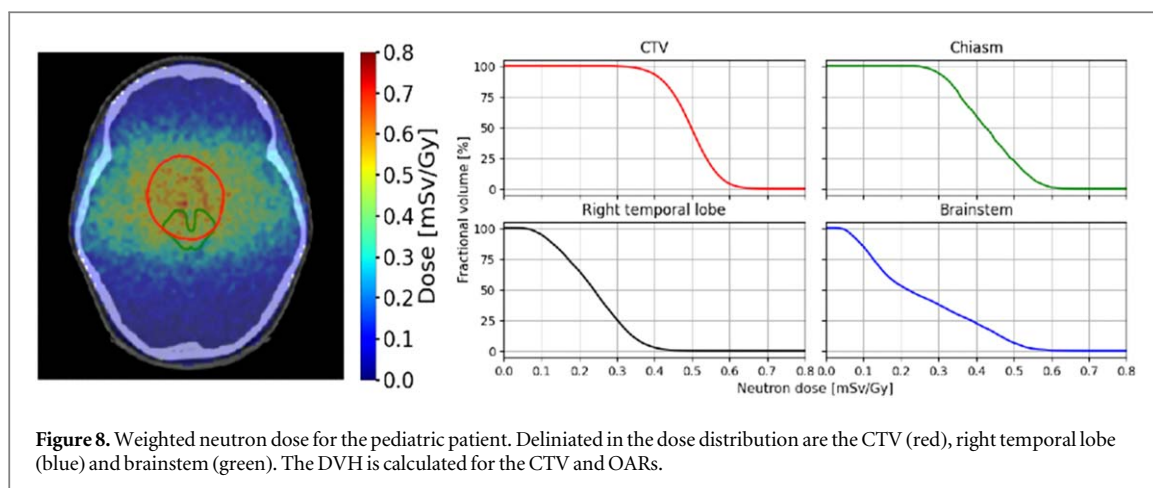
**Table 1.** Gamma test pass rate in percentage, performed on the TPS and MC calculations, for different volumes of interest and different criteria.

|                | 3 mm, 3% | 2 mm, 2% | 1 mm, 1% |
|----------------|----------|----------|----------|
| Brain          | 100.0%   | 99.9%    | 95.5%    |
| CTV            | 100.0%   | 99.9%    | 86.3%    |
| Brainstem      | 100.0%   | 100.0%   | 87.2%    |
| Chiasma        | 100.0%   | 98.1%    | 80.7%    |
| Right temp Lobe| 100.0%   | 100.0%   | 99.1%    |

While there were indications of hotspots for the neutron dose, the doses rarely exceeded 0.6 mSv Gy$^{-1}$ (figure 8). Calculated median dose for volumes in the brain also utilized in dose verification showed a low 0.2–0.3 m Sv Gy$^{-1}$.

The gamma test, performed on the TPS and MC proton dose calculations, for RBE = 1.1, showed solid 95.5% pass rates for the total brain, and an acceptable pass rate for the CTV, brainstem, and chiasm for the extreme requirements 1 mm distance to agreement and 1% dose difference (table 1).

The results from the pediatric patient can be utilized to demonstrate the difference when integrating the whole Bragg curve compared to $R_{50p}$–$R_{20}$. When integrating over the whole Bragg curve, the MC mean dose increased by 0.7 Gy for the CTV (53.6 Gy for $R_{50p} - R_{20}$, compared to 54.3 Gy for the whole PBP integrated), where the mean dose calculated by the TPS was 53.8 Gy. The exception were volumes preceding the PBP such as the right temporal lobe (figure 9). The dose difference between the two conversion factors favored integrating the whole Bragg curve. The gamma test preformed for the two dose DCF models favored the integration from $R_{50P}$–$R_{20}$ (table 2). In this comparison, the gamma test was only performed for the extreme scenario 1 mm distance to agreement, 1% dose difference since this difference is easier to determine. When the whole PBP was integrated over and utilized in the absolute dose conversion, a lower pass rate was observed for the target region (PBP area). In this example, a notable, more significant pass rate was found for regions proximal to the PBP such as the right temporal lobe for whole Bragg curve integration (table 2).

Compared to the brain tumor patient, the dose difference in the CTV was larger for the lung patient ($D_{50\%}$ 1.5 Gy), while the dose difference for the right lung ($D_{20\%}$ of 0.4 Gy) and heart was minimal. Dose difference calculated for the right lung was larger at low doses and lower at high doses when compared to the MC (figure 10). The beam entering posterior to the patient had a lower calculated dose in the TPS. The CTV received a higher dose (figure 10). A gamma test was also performed for the lung plan, where we found an overall lower pass rate compared to the brain tumor plan (table 3).

When excluding the posterior field for the lung patient, a significantly larger pass rate was observed for all structures except for the heart with the criteria 1 mm distance to agreement and 1% dose difference (table 4).

The dose difference between the TPS and MC was larger for the prostate patient compared to the pediatric case for high doses in the CTV (figure 11). The CTV had a $D_{50\%}$ difference of 0.3 Gy (69.9 Gy and 70.2 Gy for the TPS and MC, respectively. The 0.3 Gy difference is a rather small difference. The dose distribution (figure 11) showed good agreement between the TPS and MC except for distal to bony structures. The variable RBE models (Rørvik and McNamara) had a moderate increase in RBE-weighted dose, a consequence of the low $\alpha/\beta$ of the
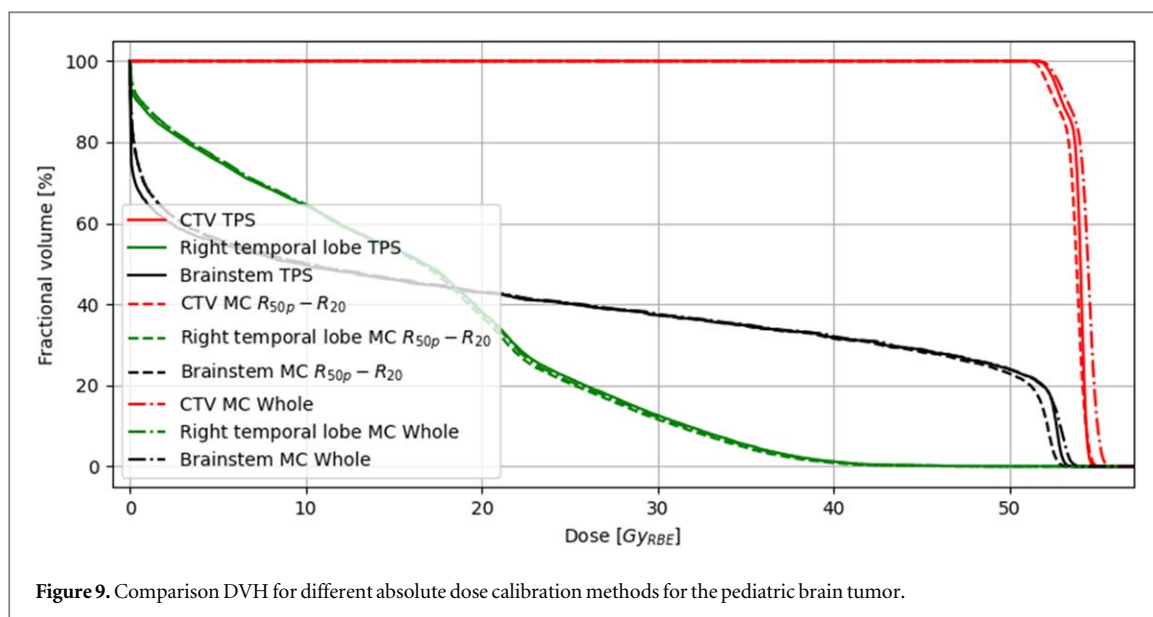
**Figure 9.** Comparison DVH for different absolute dose calibration methods for the pediatric brain tumor.

**Table 2.** Gamma test performed for the pediatric brain tumor plan comparing the $R_{50p} - R_{20}$ integration to integration of the whole Bragg curve. Where $R_{50p} - R_{20}$ area of the Bragg curve is currently employed in FLUKA. A DTA of 1 mm and a DD of 1% was employed as this is the only criteria displaying the difference to a meaningful degree.

| | 1 mm, 1%, $R_{50p} - R_{20}$ | 1 mm, 1%, whole |
|---|---|---|
| Brain | 95.5% | 46.9% |
| CTV | 86.3% | 67.1% |
| Brainstem | 87.2% | 95.0% |
| Chiasm | 80.7% | 78.8% |
| Right temporal lobe | 99.1% | 99.9% |

tumor. The gamma test showed good agreement for the prostate case between the TPS and MC except for the 1% dose difference,1 mm distance to agreement for the CTV (table 5).

## Discussion

The motivation for presenting this calibration procedure is to increase availability of MC tools in proton therapy research and for independent dose verification. Despite the introduction of MC particle transportation calculations in some TPS software, it is not as adaptable as a dedicated MC code, where implementation and calculation of state-of-the-art variable RBE is feasible with high flexibility and transparency to the user. The use of general purpose MC codes also enables deeper insights into mechanisms and interactions which is essential to support development work (Schneider *et al* 2016, Ytre-Hauge *et al* 2019, Winterhalter *et al* 2020). In this work, we present a novel systematic method of calibrating a MC to a TPS. In particular, we introduce a new mathematical method for estimating energy spread only using a few obtainable parameters and a first of its kind dose conversion method only requiring integral depth dose curves. Range differences for pristine Bragg curves was measured along with verifications of SOBPs for a broad selection of energies. Verification and variable RBE calculations were performed in patient cases and included a wide range of densities and cellular $\alpha/\beta$ to cover relevant clinical scenarios.

A range uncertainty of 0.5 mm has commonly been applied during previous range calibrations (Espana and Paganetti 2010, Schuemann *et al* 2014, Grassberger *et al* 2015). Grevillot *et al* (2011) used a percentage range difference, which at low energies could misrepresent the uncertainty. In our calibration, we assumed a uniform hit registration across scoring volumes and applied a standard deviation of a uniform distribution (Illowsky 2020). With this assumption, $\pm 0.3$ mm range difference uncertainty is within what can be obtained from experimental measurements. A few measurements proximal to the BP itself failed to meet this uncertainty, however that is to be expected with differences in algorithms. In comparison, prior to the calibration of PBPs the
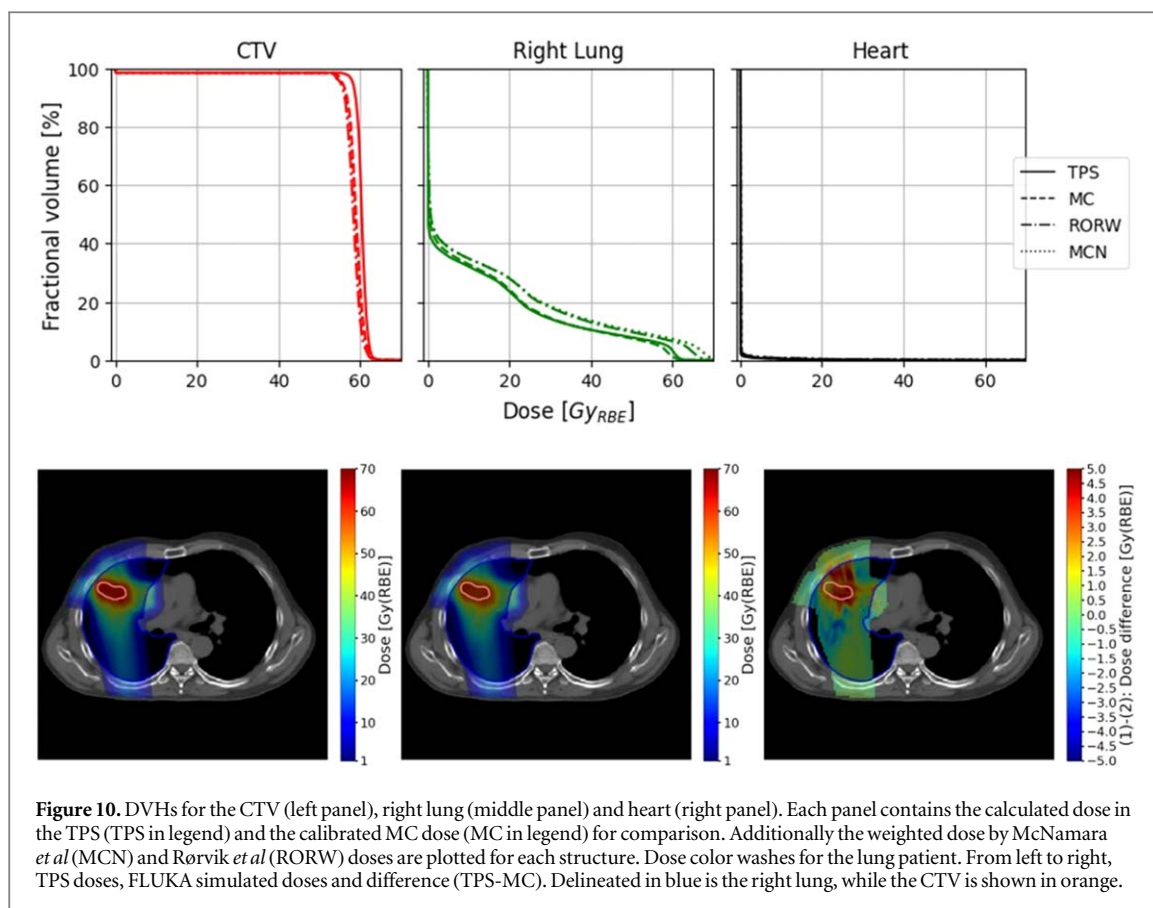
**Figure 10.** DVHs for the CTV (left panel), right lung (middle panel) and heart (right panel). Each panel contains the calculated dose in the TPS (TPS in legend) and the calibrated MC dose (MC in legend) for comparison. Additionally the weighted dose by McNamara *et al* (MCN) and Rørvik *et al* (RORW) doses are plotted for each structure. Dose color washes for the lung patient. From left to right, TPS doses, FLUKA simulated doses and difference (TPS-MC). Delineated in blue is the right lung, while the CTV is shown in orange.

**Table 3.** Gamma test results performed on the TPS and MC dose for the CTV, right lung, and heart of the lung tumor patient.

|            | 3 mm, 3% | 2 mm, 2% | 1 mm, 1% |
|------------|----------|----------|----------|
| CTV        | 83.4%    | 65.4%    | 36.5%    |
| Right lung | 96.9%    | 90.5%    | 67.0%    |
| Heart      | 99.6%    | 99.3%    | 71.6%    |

**Table 4.** Gamma test comparison for the original lung patient treatment plan and a plan where the posterior–anterior field is excluded. The comparison was done for the extreme scenario with a DTA of 1 mm, and a DD of 1% as these criteria highlights the difference in dose.

| lung field comparison | 1 mm, 1% 3 fields | 1 mm, 1% 2 fields |
|-----------------------|-------------------|-------------------|
| CTV                   | 36.5%             | 72.7%             |
| Right lung            | 67.0%             | 66.2%             |
| Heart                 | 71.6%             | 61.5%             |

values were almost completely outside the accepted uncertainty (figure A.6). If only range calibrations would be done, $R_{20}$ could shift significantly compared to the TPS, which could also lead to systematic $LET_d$ shift. Nevertheless, there has been a lack of reporting proximal calculated ranges for PBPs (Grevillot *et al* 2011, Fracchiolla *et al* 2015, Grassberger *et al* 2015, Kozłowska *et al* 2019), and for our method for absolute dose MU conversion will be affected by a skewed range difference in the proximal PBP. We therefore suggest reporting ranges proximal to the peak, the peak itself and distal to the peak. Moreover, the American association for physics in medicine (AAPM) recommends the usage of ±1 mm uncertainty for clinical measurements (Arjomandy *et al* 2019) which is well within the ±0.3 mm utilized here. Another task group from AAPM, Farr *et al* (2021) recommended a 95% pass rate for a 3 mm/3% gamma-test for a 2 dimensional dose distribution,
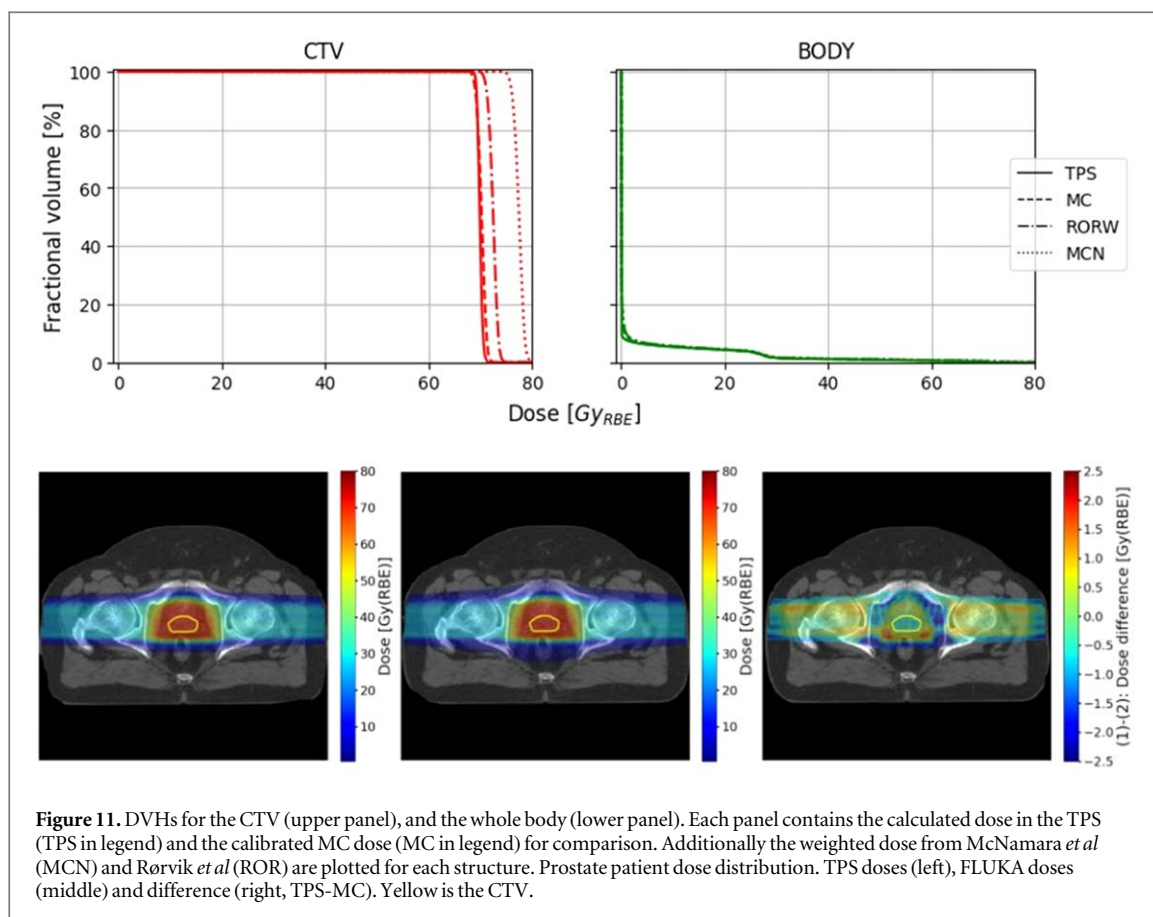
**Figure 11.** DVHs for the CTV (upper panel), and the whole body (lower panel). Each panel contains the calculated dose in the TPS (TPS in legend) and the calibrated MC dose (MC in legend) for comparison. Additionally the weighted dose from McNamara *et al* (MCN) and Rørvik *et al* (ROR) are plotted for each structure. Prostate patient dose distribution. TPS doses (left), FLUKA doses (middle) and difference (right, TPS-MC). Yellow is the CTV.

**Table 5.** Gamma test results for TPS and MC doses for different thresholds for CTV and body for the prostate patient, using the criteria DTA = 3, 2 and 1 mm and DD = 3, 2 and 1%.

|       | 3 mm, 3% | 2 mm, 2% | 1 mm, 1% |
|-------|----------|----------|----------|
| CTV   | 100.0%   | 97.8%    | 60.6%    |
| Body  | 99.7%    | 98.9%    | 88.5%    |

where all of our cases passed with the exception for the CTV in the extreme lung case scenario with 3 fields. Based on the results from our MC calibration and verification in the different patient cases, we recommend reporting gamma scores for at least 3 mm/3% as a measure of general accuracy and 1 mm/1% as a measure of how precise the calibration is for volumes both proximal to the BP and the target volume.

For the absolute dose calibration, the integrated area of the Bragg curves will affect the required dose conversion factor (DCF) (figure 4). This, in turn, will determine whether an average over the whole curve or the BP area exclusively, is favored. Kelleter *et al* (2019) found that by excluding proton and electronic buildup, the dose could be reduced by 16% at the entrance and affect the dose accuracy up to 150 mm in depth. The deterministic TPS algorithms does to some degree account for this entrance buildup but have limitations in the transition across different densities. Also, PCS algorithm can moderately account for non-elastic scattering through its double Gaussian fluence model (Shen *et al* 2016), but will deviate in the plateau region of the Bragg curve. Furthermore, non-elastic scattering will be more prominent for higher energies (figure 4). As the beam energy increases, the difference between the full Bragg curve area and only the $R_{50P} - R_{20}$ DCF, will also increase. The PCS algorithm is less accurate in inhomogeneous regions (De Martino *et al* 2021), such as in our applied phantoms with an intersection of vacuum and water. Due to subsequently inconsistency between the TPS and MC in these regions, the entrance as well as parts of the plateau of the Bragg curve was omitted when calculating the DCF. A DCF derived from integrating whole Bragg curves may for the reasons mentioned above lead to an excessive high dose for the high dose regions, for example the CTV in the MC recalculations. Moreover, a DCF utilizing the area $R_{50P} - R_{20}$ will result in a more similar dose in the CTV for the two calculation methods.

Differences across dose calculation methods/algorithms are quantified by calculating dose differences and gamma tests. Paganetti *et al* (2008) calibrated the Geant4 MC code to measurements, they examined the difference in a MC algorithm and pencil beam algorithm, evaluating the dose differences by gamma tests for the criteria 2 mm, 2%. Whereas tabulated values were not provided, gamma pass scores cannot be directly compared to this work and for similar patients. Furthermore, Kozlowska *et al* (2019) compared FLUKA to the TPSs from two different institutions. The simulation preformed with the Trento commissioned beam line included beam characteristics only, i.e. no geometry involved in the MC simulation, as has been done for this calibration as well. Distinguishing the Trento and the calibration done in this work a PTV dose difference up to 1.4% (0.8 Gy) in a chordoma case (Trento-FLUKA), compared to 0.2% (0.1 Gy) in our calibration (Eclipse-FLUKA) was noted. The difference in PTV dose from the CNAO commissioning, was however, similar to our work (0.1 Gy).

The corrections done to RPSP and density within intervals were performed for HUs $\geqslant -900$, as most HU values were above this limit (lung cancer patient, figure A.7). Dose calculation algorithms (pencil beam) tend to underestimate dose to normal lung tissue (Grassberger *et al* 2014), as shown in the dose distributions for the lung patient (figure 10). For the RPSP calibrations of lower HUs ($< -830$) a phantom with extended depth (4 m) was utilized. A lower HU phantom could result in a larger lateral deviation between the TPS and MC, though, this would not be of concern unless a beam traverses through a large HU region, as with the posterior anterior lung field. This tendency can also be seen for low dose regions for the lung DVH (figure 10). In another study by the same group (Grassberger *et al* 2015), TOPAS was commissioned for different patient scenarios and compared to a pencil beam algorithm. They found up to 30% dose difference between the two modalities for a lung cancer patient plan. In our lung patient case, we obtained a maximum dose difference of 15%. A pencil beam algorithm will only account for material along its central axis (Petti 1992) and thus multiple coulomb scattering is not modeled as well as in a MC, furthermore, in inhomogeneous material the pencil beam algorithm will be less accurate (De Martino *et al* 2021). To investigate the calibration limits, we used a posterior–anterior field in a lung cancer case with a pass rate of 36.5% (1 mm, 1%), where the beam traverses through a large low HU-value organ (Grassberger *et al* 2014). As expected, a higher pass rate (72.7%) for the CTV was achieved when the posterior–anterior field was excluded from the treatment plan (table 4). As modern TPSs move towards an integrated MC, it might be reasonable to it in cases with low densities regions.

Grevillot *et al* (2012) adapted a TPS from Electa (XiO) to the MC, though without implementing nozzle components, similarly to our current work. This implementation allowed for the user to adjust the precision by tuning the number of sub-spots in each calculation (using 1, 49 and 121 sub-spots), 1 sub-spot would be comparable to the PCS algorithm. They further compared to a calibrated MC toolkit (GATE). With a medium degree of precision for XiO and dose to water for GATE, they achieved similar dose differences (for $D_{50\%}$) at 0.4%, whereas with our set-up we achieved 0.2% in the GTV of a prostate patient.

For recalculation in this work, we utilized an Intel Xeon Gold 6248 with 80 cores and 128 Gb of RAM and achieved sufficient statistical accuracy within 4–12 h per patient. *Commercial treatment planning systems with integrated GPU* (Schreuder *et al* 2019) *are faster, but with a different scope of utility. With the flexibility and extra data opportunity, a dedicated MC is a valuable addition to the clinic as it can provide selectivity related to RBE models, out-of-field doses, and support experimental measurements.*

## Conclusion

In this study we calibrated a MC to a TPS for pencil beam scanning proton therapy and presented a novel, easy-to-reproduce, stepwise method. We achieved excellent accuracy in both water and patient cases, with the exception for extreme fields in low density tissue. The method of calibrating range and energy spread will work for any modern TPS and general purpose MC code. Having a dedicated MC enables reproducing proton treatment plans calculated with clinical TPSs, paving the way for studies including LET and out-of-field doses.

## Acknowledgments

## Data availability statement

The data cannot be made publicly available upon publication because the cost of preparing, depositing and hosting the data would be prohibitive within the terms of this research project. The data that support the findings of this study are available upon reasonable request from the authors.

## Ethical statement

All DICOM data used was applied in an anonymized fashion and approved by the Regional Committees for Medical and Health Research Ethics, REK No. 2016/1636. This study is exempt from the requirement of consent from the participants. The research was also conducted in accordance with the principles embodied in the Declaration of Helsinki as well as in accordance with local statutory requirements.
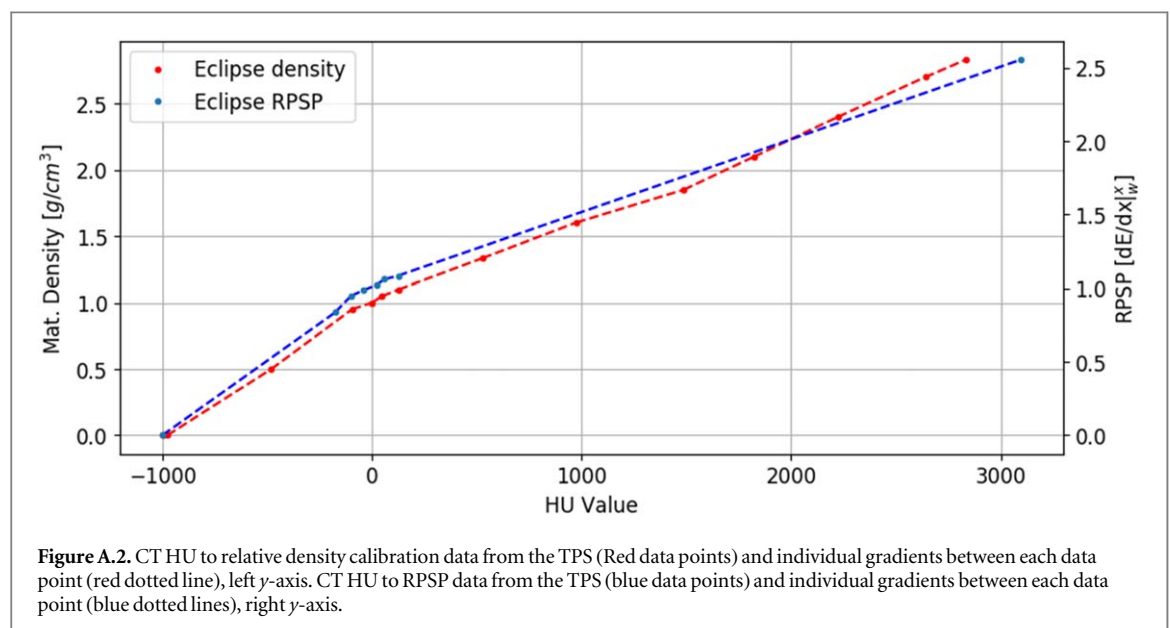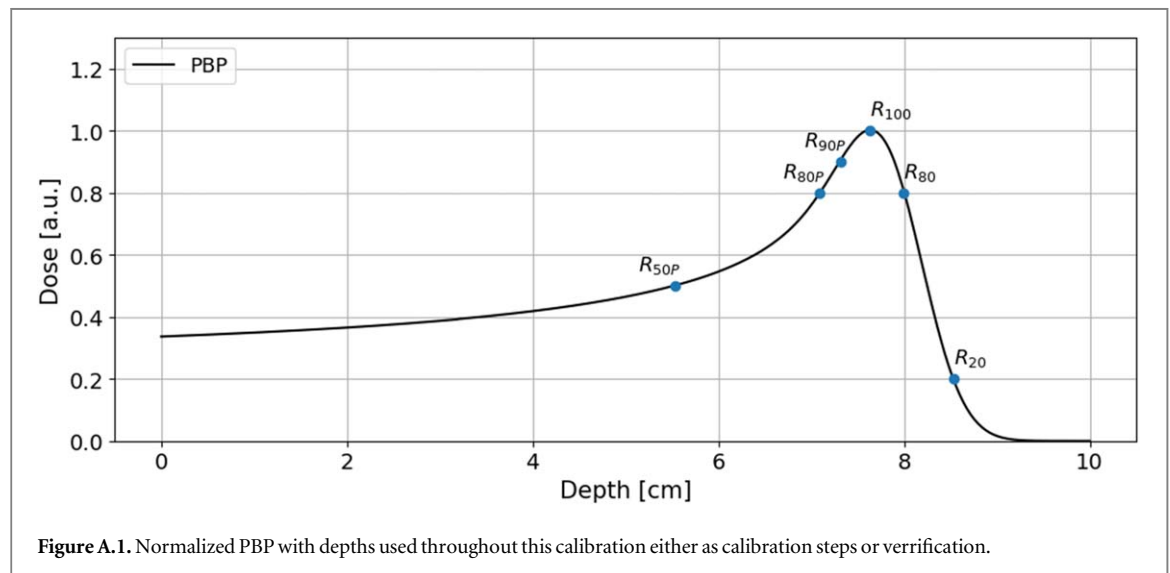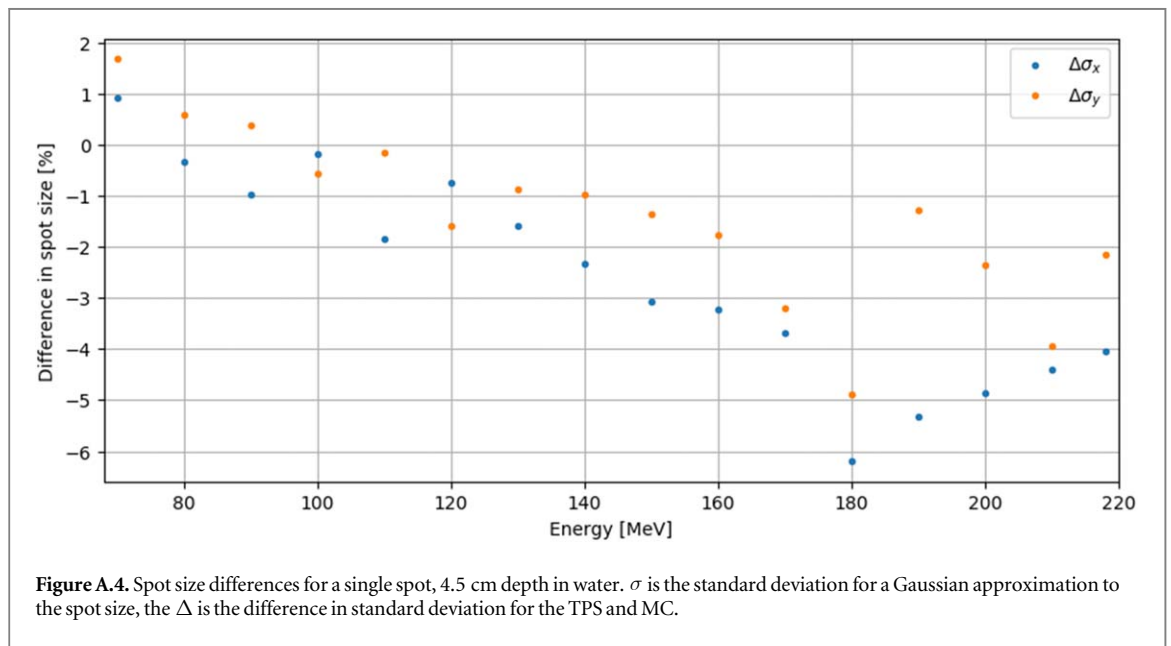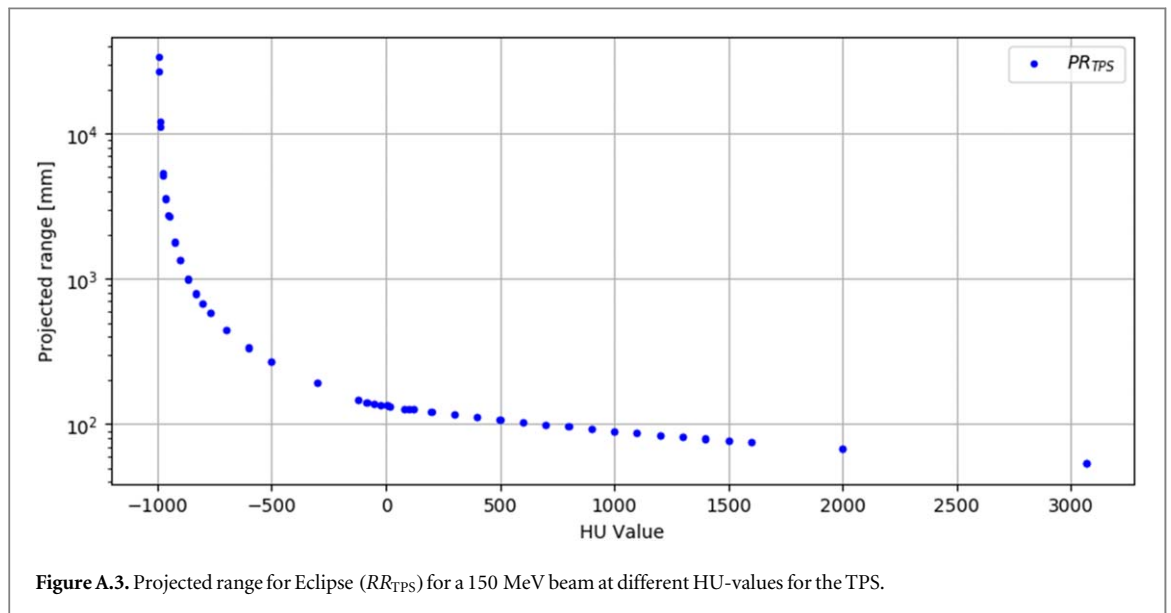
## Appendix A

### Calibration data

The energy spread for the TPS, and MC should be equal after corrections are made. The ratio between $R_{8020TPS}$ and $R_{8020MC}$ should then be 1 for all energies, as this indicates a perfect calibration. The ratios were all found to be within 10% of 1 (figure A.8).

Figure A.8). The CT calibration curves for density and RPSP implemented in FLUKA can be seen in figure A.2.

Spot sizes were only verified (figure A.4) since they were within the clinical threshold (10% FWHM) between the two modalities.

In figure A.7, the HU-values has been summarized for the lung patient and it can be determined that approximately none will be lower than $-900$.



**Figure A.1.** Normalized PBP with depths used throughout this calibration either as calibration steps or verrification.



**Figure A.2.** CT HU to relative density calibration data from the TPS (Red data points) and individual gradients between each data point (red dotted line), left *y*-axis. CT HU to RPSP data from the TPS (blue data points) and individual gradients between each data point (blue dotted lines), right *y*-axis.

**Figure A.3.** Projected range for Eclipse ($RR_{TPS}$) for a 150 MeV beam at different HU-values for the TPS.



**Figure A.4.** Spot size differences for a single spot, 4.5 cm depth in water. $\sigma$ is the standard deviation for a Gaussian approximation to the spot size, the $\Delta$ is the difference in standard deviation for the TPS and MC.

**Figure A.5.** Lateral dose profiles for the high (upper panel), medium (middle panel) and low (lower panel) energy SOBPs in water for the TPS (orange lines) and MC (blue lines). The depth profiles were sampled at a depth of 175 mm, 105 mm and 45 mm for the high, medium and low SOBPs, respectively.



**Figure A.6.** Range differences ($\Delta$) prior to the calibration of the PBPs for the MC minus the TPS PBPs in mm along the depth dose curves for increasing energy, indicated by the respective markers, range at 80% and 90% proximal to the PBP ($R_{80P}$ and $R_{90P}$, respectively ) proximal to the PBP, $R_{100}$ as the PBP and $R_{80}$ and $R_{20}$ as 80% and 20% distal to the PBP. The red lines indicate the $\pm 0.3$ mm criteria. The only range shown with error bars are the most volatile range ($R_{80P}$).

**Figure A.7.** HU- value count for the lung patient for the right lung volume. The *y*-axis is normalized so that max is one.



**Figure A.8.** $R_{8020}$ for the MC devided by the $R_{8020}$ for the TPS after the calibration is done for all energy intervalls used for the PBPs.

## Appendix B

**Energy spread**

Steps for arrival to equation (3):

$$R_{8020} = 1.3(\sigma_{\text{range}}^2 + \sigma_{\text{beam}}^2)^{1/2}$$

$$\frac{R_{8020}}{1.3(\sigma_{\text{range}}^2 + \sigma_{\text{beam}}^2)^{\frac{1}{2}}} = 1.$$

To simplify, we write the $R_{8020}$ as distal dose fall-off i.e. DDF and use subscript TPS and MC to clarify which DDF we are using

$$\frac{\text{DDF}_{\text{TPS}}}{1.3(\sigma_{\text{range}T}^2 + \sigma_{\text{beam}T}^2)^{\frac{1}{2}}} = \frac{\text{DDF}_{\text{MC}}}{1.3(\sigma_{\text{range}M}^2 + \sigma_{\text{beam}M}^2)^{\frac{1}{2}}}.$$

By arranging the equations with the values obtained through PBP analysis (DDF$_{\text{TPS}}$, DDF$_{\text{MC}}$ and assumed energy spread) we get:

$$\frac{\mathrm{DDF}_{\mathrm{TPS}}}{\mathrm{DDF}_{\mathrm{MC}}}(\sigma^2_{\mathrm{rangeMC}} + \sigma^2_{\mathrm{beamMC}})^{\frac{1}{2}} = (\sigma^2_{\mathrm{rangeTPS}} + \sigma^2_{\mathrm{beamTPS}})^{\frac{1}{2}}.$$

If we assume that the contribution to range straggling and beam spread can be included as one parameter, $\sigma^2 = \sigma^2_{\mathrm{range}} + \sigma^2_{\mathrm{beam}}$ we arrive at equation (3):

$$\sigma_{\mathrm{MC}}\frac{\mathrm{DDF}_{\mathrm{TPS}}}{\mathrm{DDF}_{\mathrm{MC}}} = \sigma_{\mathrm{TPS}}.$$

## ORCID iDs

Johannes Tjelta ⓘ https://orcid.org/0009-0005-4368-9551
Lars Fredrik Fjæra ⓘ https://orcid.org/0000-0001-6521-506X
Kristian Smeland Ytre-Hauge ⓘ https://orcid.org/0000-0002-3597-1611
Camilla Hanquist Stokkevåg ⓘ https://orcid.org/0000-0001-6352-4571

## References

Antonie R H and Schneider U 2019 Neutron dose and its measurement in proton therapy—current state of knowledge *BJR* **93** 1107–19
Arjomandy B *et al* 2019 AAPM task group 224: comprehensive proton therapy machine quality assurance *Med. Phys.* **46** e678–705
Battistoni G *et al* 2016 The FLUKA code: an accurate simulation tool for particle therapy *Front. Oncol.* **6** 116–39
Böhlen T T *et al* 2014 The FLUKA code: developments and challenges for high energy and medical applications *Nucl. Data Sheets* **120** 211–4
Borst G R *et al* 2010 Radiation pneumonitis after hypofractionated radiotherapy: evaluation of the LQ(L) model and different dose parameters *Int. J. Radiat. Oncol. Biol. Phys.* **77** 1596–603
Bortfeld T 1997 An analytical approximation of the bragg curve for therapeutic proton beams *Med. Phys.* **24** 2024–33
Brenner D J and Hall E J 1999 Fractionation and protraction for radiotherapy of prostate carcinoma *Int. J. Radiat. Oncol. Biol. Phys.* **43** 1095–102
Pinter C, Lasso A, Wang A, Jaffray D and Fichtinger G 2021 SlicerRT: radiation therapy research toolkit for 3D Slicer *Medical Physics* **39** 6332–8 (http://slicerrt.github.io/)
Dalrymple G V *et al* 1966 Some effects of 138 Mev protons on primates *Radiat. Res.* **28** 471–88
Englbrecht F S *et al* 2021 A comprehensive Monte Carlo study of out-of-field secondary neutron spectra in a scanned-beam proton therapy gantry room *Med. Phys.* **31** 215–28
Espana S and Paganetti H 2010 The impact of uncertainties in the CT conversion algorithm when predicting proton beam ranges in patients from dose and PET-activity distributions *Phys. Med. Biol.* **55** 7557–71
Farr J B *et al* 2021 Clinical commissioning of intensity-modulated proton therapy systems: Report of AAPM Task Group 185 *Med. Phys.* **48** e1–30
Fedorov A *et al* 2012 3D slicer as an image computing platform for the quantitative imaging network *Magnetic Resonance Imaging* 4.11.20210226. https://slicer.org/ 30.9 1323–41
Ferrari A, Ranft J, Sala P R and Fasso' A 2005 FLUKA: a multi-particle transport code (Program version 2005) (No. CERN-2005-10). Cern
Fiorini F, Schreuder N and Van den Heuvel F 2018 Technical note: defining cyclotron-based clinical scanning proton machines in a fluka Monte Carlo system *Med. Phys.* **45** 963–70
Fjaera L F *et al* 2017 Linear energy transfer distributions in the brainstem depending on tumour location in intensity-modulated proton therapy of paediatric cancer *Acta Oncol.* **56** 763–8
Fjaera L F, Indelicato D J, Stokkevag C H, Muren L P, Hsi W C and Ytre-Hauge K S 2020 Implementation of a double scattering nozzle for Monte Carlo recalculation of proton plans with variable relative biological effectiveness *Phys. Med. Biol.* **65** 2250–70
Fracchiolla F, Lorentini S, Widesott L and Schwarz M 2015 Characterization and validation of a Monte Carlo code for independent dose calculation in proton therapy treatments with pencil beam scanning *Phys. Med. Biol.* **60** 8601–19
Gajewski J 2023 *Bragg Peak Analysis.* (Version 1.1.0.0). MATLAB Central File Exchange, https://www.mathworks.com/matlabcentral/fileexchange/63405-bragg-peak-analysis. Retrieved August 25, 2021
Gottschalk B 2012 Physics of proton interaction in matter *Proton Therapy Physics* ed H Paganetti (Taylor & Francis Group)
Grassberger C, Daartz J, Dowdell S, Ruggieri T, Sharp G and Paganetti H 2014 Quantification of proton dose calculation accuracy in the lung *Int. J. Radiat. Oncol. Biol. Phys.* **89** 424–30
Grassberger C, Lomax A and Paganetti H 2015 Characterizing a proton beam scanning system for Monte Carlo dose calculation in patients *Phys. Med. Biol.* **60** 633–45
Grevillot L, Bertrand D, Dessy F, Freud N and Sarrut D 2011 A Monte Carlo pencil beam scanning model for proton treatment plan simulation using GATE/GEANT4 *Phys. Med. Biol.* **56** 5203–19
Grevillot L, Bertrand D, Dessy F, Freud N and Sarrut D 2012 GATE as a GEANT4-based Monte Carlo platform for the evaluation of proton pencil beam scanning treatment plans *Phys. Med. Biol.* **57** 4223–44
ICRP, Protection R 2007 The 2007 Recommendations of the international commission on radiological protection *Ann ICRP* publication 103 **37** 2
ICRU 2007 Prescribing, recording, and reporting proton-beam therapy *J. ICRU* **7** 2
Illowsky B and Dean S 2020 S D. Statistics OpenStax: OpenStax; [Available from: https://openstax.org/books/statistics/pages/5-2-the-uniform-distribution
Kehwar T S 2005 Analytical approach to estimate normal tissue complication probability using best fit of normal tissue tolerance doses into the NTCP equation of the linear quadratic model *J. Cancer Res. Ther.* **1** 168–79
Kelleter L *et al* 2019 Technical note: simulation of dose buildup in proton pencil beams *Med. Phys.* **46** 3734–8
Klement R J *et al* 2020 Estimation of the alpha/beta ratio of non-small cell lung cancer treated with stereotactic body radiotherapy *Radiother. Oncol.* **142** 210–6
Kondziolka D, Shin S M, Brunswick A, Kim I and Silverman J S 2015 The biology of radiosurgery and its clinical applications for brain tumors *Neuro Oncol.* **17** 29–44

Kozłowska W S *et al* 2019 FLUKA particle therapy tool for Monte Carlo independent calculation of scanned proton and carbon ion beam therapy *Phys. Med. Biol.* **64** 75–87

Low D A and Dempsey J F 2003 Evaluation of the gamma dose distribution comparison method *Med. Phys.* **30** 2455–64

De Martino F, Clemente S, Graeff C, Palma G and Cella L 2021 Dose calculation algori thms for external radiation therapy: an overview for practitioners *Appl. Sci.* **11** 15

McNamara A L, Schuemann J and Paganetti H 2015 A phenomenological relative biological effectiveness (RBE) model for proton therapy based on all published *in vitro* cell survival data *Phys. Med. Biol.* **60** 8399–416

Paganetti H 2017 Relating the proton relative biological effectiveness to tumor control and normal tissue complication probabilities assuming interpatient variability in alpha/beta *Acta Oncol.* **56** 1379–86

Paganetti H *et al* 2019 Report of the AAPM TG-256 on the relative biological effectiveness of proton beams in radiation therapy *Med. Phys.* **46** e53–78

Paganetti H 2022 Mechanisms and review of clinical evidence of variations in relative biological effectiveness in proton therapy *Int. J. Radiat. Oncol. Biol. Phys.* **112** 222–36

Paganetti H, Jiang H, Parodi K, Slopsema R and Engelsman M 2008 Clinical implementation of full Monte Carlo dose calculation in proton beam therapy *Phys. Med. Biol.* **53** 4825–53

Petti P L 1992 Differential-pencil-beam dose calculations for charged particles *Med. Phys.* **19** 137–49

Rorvik E *et al* 2018 Exploration and application of phenomenological RBE models for proton therapy *Phys. Med. Biol.* **63** 185–204

Schneider U, Halg R A, Baiocco G and Lomax T 2016 Neutrons in proton pencil beam scanning: parameterization of energy, quality factors and RBE *Phys. Med. Biol.* **61** 6231–42

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** 20

Schreuder A N *et al* 2019 Validation of the raystation Monte Carlo dose calculation algorithm using realistic animal tissue phantoms *J. Appl. Clin. Med. Phys.* **20** 160–71

Schuemann J, Dowdell S, Grassberger C, Min C H and Paganetti H 2014 Site-specific range uncertainties caused by dose calculation algorithms for proton therapy *Phys. Med. Biol.* **59** 4007–31

Schwarz M *et al* 2016 Clinical pencil beam scanning: present and future practices *Particle Radiotherapy Emerging Technology for Treatment of Cancer* ed K A Rath and N Sahoo (Springer India) p 169

Shen J *et al* 2016 An efficient method to determine double Gaussian fluence parameters in the eclipse proton pencil beam model *Med. Phys.* **43** 6544

Smeets J *et al* 2012 Prompt gamma imaging with a slit camera for real-time range control in proton therapy *Phys. Med. Biol.* **57** 3371–405

Underwood T S A, McNamara A L, Appelt A, Haviland J S, Sorensen B S and Troost E G C 2022 A systematic review of clinical studies on variable proton Relative Biological Effectiveness (RBE) *Radiother. Oncol.* **175** 79–92

Vlachoudis V 2009 Flair: a powerful but user friendly graphical interface for FLUKA *In Proc. Int. Conf. on Mathematics* (New York: Computational Methods & Reactor Physics (M&C 2009), Saratoga Springs) 176 (http://inis.iaea.org/search/search.aspx?orig_q=RN:42064858)

Willers H *et al* 2018 Toward A variable RBE for proton beam therapy *Radiother. Oncol.* **128** 68–75

Winterhalter C *et al* 2020 Pitfalls in the beam modelling process of Monte Carlo calculations for proton pencil beam scanning *BJR* **93** 919–30

Ytre-Hauge K S, Skjerdal K, Mattingly J and Meric I 2019 A Monte Carlo feasibility study for neutron based real-time range verification in proton therapy *Sci. Rep.* **9** 2011