PAPER

Proton therapy delivery method affects dose-averaged linear energy transfer in patients

IOP Publishing Phys. Med. Biol. 66 (2021) 074003 https://doi.org/10.1088/1361-6560/abe835

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

RECEIVED
25 January 2021

ACCEPTED FOR PUBLICATION
19 February 2021

PUBLISHED
1 April 2021

PAPER

Proton therapy delivery method affects dose-averaged linear energy transfer in patients

Lydia J Wilson[1], Fakhriddin Pirlepesov[1], Vadim Moskvin[1], Zuofeng Li[2], Yian Guo[3], Yimei Li[3], Thomas E Merchant[1] and Austin M Faught[1]

[1] St. Jude Children's Research Hospital, Department of Radiation Oncology, Memphis, TN, United States of America
[2] University of Florida Proton Therapy Institute, Department of Radiation Oncology, Jacksonville, FL, United States of America
[3] St. Jude Children's Research Hospital, Department of Biostatistics, Memphis, TN, United States of America

E-mail: Austin.Faught@StJude.org

Keywords: linear energy transfer, proton therapy, passive scattering, pencil-beam scanning, childhood cancer, craniopharyngioma

Abstract

The dosimetric advantages of proton therapy have led to its rapid proliferation in recent decades. This has been accompanied by a shift in technology from older units that deliver protons by passive scattering (PS) to newer units that increasingly use pencil-beam scanning (PBS). The biologic effectiveness of proton physical dose purportedly rises with increasing dose-weighted average linear energy transfer (LET_D). The objective of this study was to determine the extent to which proton delivery methods affect LET_D. We calculated LET_D from simple, dosimetrically matched, and clinical treatment plans with TOPAS Monte-Carlo transport code. Simple treatment plans comprised single fields of PS and PBS protons in a water phantom. We performed simulations of matched and clinical treatment plans by using the treatment and anatomic data obtained from a cohort of children with craniopharyngioma who previously received PS or PBS proton therapy. We compared the distributions of LET_D from PS and PBS delivery methods in clinically relevant ROIs. Wilcoxon signed-rank tests comparing single fields in water revealed that the LET_D values from PBS were significantly greater than those from PS inside and outside the targeted volume ($p < 0.01$). Statistical tests comparing LET_D-volume histograms from matched and clinical treatment plans showed that LET_D was generally greater for PBS treatment plans than for PS treatment plans ($p < 0.05$). In conclusion, the proton delivery method affects LET_D both inside and outside of the target volume. These findings suggest that PBS is more biologically effective than PS. Given the rapid expansion of PBS proton therapy, future studies are needed to confirm the applicability of treatment evaluation methods developed for PS proton therapy to those for modern PBS treatments to ensure their safety and effectiveness for the growing population of patients receiving proton therapy. This study uses data from two clinical trials: NCT01419067 and NCT02792582.

1. Introduction

Proton therapy has rapidly expanded in use and availability since the late 1980s. Worldwide, the number of treatment centers offering proton therapy doubled between 2012 and 2016 (Warwick 2018). As of January 2021, 98 proton therapy centers are operational worldwide, and another 58 are planned or under construction (PTCOG 2020). Early proton therapy systems delivered protons using passive-scattering (PS) methods to broaden the beam, facilitating the treatment of clinical targets. PS methods require beam-shaping apertures to control the lateral extent of treatment fields and compensators to conform the dose to the distal edge of the target. Recently, however, pencil-beam scanning (PBS) has become the standard for new beamlines and facilities (Paganetti 2014, Jermann 2019). PBS methods use magnetically scanned beams to control the lateral size and shape of treatment fields and energy modulation to achieve dose conformity to both the proximal and distal edges of the target.

The increased prevalence of proton therapy is due to its superior pattern of physical dose deposition over that of photon therapy. The unique depth-dose characteristics of proton therapy decrease the integral physical dose deposited outside of target volumes to below that of photon therapy. Radiation dose deposited in healthy tissues outside of target volumes is linked to early and late side effects of treatment; the latter of which may arise several decades after treatment is completed (Newhauser *et al* 2009, Perez-Andujar *et al* 2013, Zhang *et al* 2013). Therefore, the theoretical advantages of proton therapy (i.e. decreased risk of early and late side effects) make it particularly well suited for treating children, whose bodies are inherently more susceptible to radiation effects and for whom survivorship is measured in decades (ASTRO 2014).

The physical dose, however, differs from the biological dose by a factor termed the relative biological effectiveness (RBE). Current clinical practice aims to achieve uniformity across photon and proton therapy clinical regimens by applying a uniform RBE (i.e. 1.1), regardless of the proton therapy treatment-planning and delivery method, and in all regions of dose deposition (Paganetti *et al* 2019). However, this practice does not take into consideration the complex function of several physical (e.g. particle type, delivery method, dose, linear energy transfer), and biological (e.g. cell type and endpoint) characteristics (Paganetti and Schmitz 1996, Wouters *et al* 1996, Paganetti and Goitein 2001, Wilkens and Oelfke 2004). Previous experimental studies revealed that the RBE dependence on linear energy transfer is defined by the dose-weighted average linear energy transfer (LET_D) (Perris *et al* 1986, Belli *et al* 1989, 1992, Folkard *et al* 1996, Coutrakon *et al* 1997). LET_D itself depends on such factors as treatment-planning technique (Grassberger *et al* 2011), particle type (Folkard *et al* 1996), and particle origin (Grassberger and Paganetti 2011). Relatively less is known, however, about the effect of proton delivery methods on LET_D (Grassberger and Paganetti 2011, Grassberger *et al* 2011, Paganetti 2014, Gridley *et al* 2015, Moskvin 2016, 2016a, Paganetti *et al* 2019). Grassberger *et al* hypothesized that the broader energy distributions characteristic of PS methods result in lower LET_D magnitudes at the end of range than do PBS methods (Grassberger *et al* 2011). Although their selected case studies demonstrated the potential differences in LET_D from different delivery methods, the clinical importance in patient cohorts remains unknown. Therefore, the objective of this study was to determine the extent to which proton delivery methods affect LET_D. We used Monte-Carlo methods to calculate LET_D from single- and multi-field treatment plans of PS and PBS proton therapy. We then compared the LET_D magnitudes in clinically relevant regions of interest (ROIs).

2. Methods

2.1. Monte Carlo simulations

We performed Monte Carlo simulations via previously benchmarked models (TOPAS, Geant4 Simulation Toolkit Perl *et al* 2012) for two distinct proton beamlines used to treat pediatric patients. The first model used TOPAS v1/1.3 and was based on design information for the beamline at the University of Florida Health Proton Therapy Institute (Shin *et al* 2015). This system (IBA, Louvain La Neuve, Belgium) delivered proton therapy by the PS method. These simulations use patient-specific brass apertures and polyethylene compensators. The model implementation and validation are previously described (Shin *et al* 2015). We transported primary protons as well as secondary neutrons and electrons. Calculations using this model simulated 250×10^6 proton histories in 50 parallel jobs, requiring approximately 500 000 CPU-s, to achieve a maximum standard error of the mean dose of 3.5% in voxels within the spread-out Bragg peak.

The second model we used was based on vendor specifications for the beamline at St. Jude Children's Research Hospital (Moskvin *et al* 2016b, Farr *et al* 2018). The PROBEAT-V (Hitachi, Tokyo, Japan) system has a scanning nozzle and gantry system that delivers protons by the PBS method with a spot full width at half maximum at isocenter as small as 4.8 mm in air for the highest energy. The Monte Carlo model for this system performs simulations in two stages. Stage 1 simulates the beamline and scanning nozzle with TOPAS v2, and Stage 2 transports particles through computed tomography (CT)-based geometries with TOPAS v3.1.3 (Moskvin *et al* 2017). The job submission and execution processes are previously described (Kaluarachchi *et al* 2020). We transported primary protons, as well as secondary helium ions, neutrons, and electrons. Calculations using this model simulated 180×10^6 proton histories in eight parallel jobs, requiring approximately 5×10^6 CPU-s and achieving a maximum standard error of the mean dose of 2.0% in voxels within the target volume.

The scored resolution of simulations using both models matched that of the imaging dataset (i.e. $0.98 \times 0.98 \times 1$ mm). The calculation runs were submitted in batches to an institutional high-performance computing cluster running Red Hat Enterprise Linux (Red Hat). All simulations scored absorbed dose and LET_D. We used the TOPAS DoseToWater scorer for simulations in water and DoseToMaterial for simulations in patient CT data. The TOPAS ProtonLET scorer calculated LET_D values for all simulations in consideration of primary and secondary protons, including energy deposited by secondary electrons. The TOPAS built-in LET_D scoring technique adopted the methods described by Granville and Sawakuchi (Granville and Sawakuchi 2015).

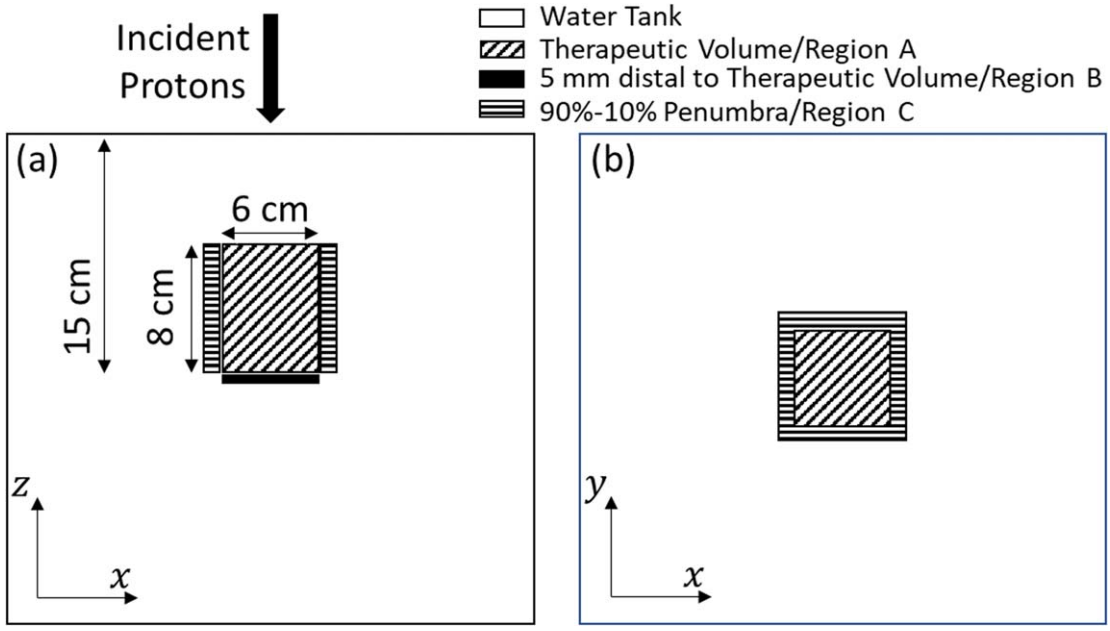

Figure 1. (a) Coronal and (b) axial views of the geometry for simulations of single proton fields in a computational water phantom. Patterned regions correspond to clinically relevant regions of interest for dose-weighted average linear energy transfer (LET_D) analyses.

As with our previous implementations of these models (Kaluarachchi *et al* 2020), we performed all simulations by using the TOPAS Geant4 standard option 4 modules (TOPAS 2020) and default thresholds recommended by the Geant4 electromagnetic physics group (i.e. a 0.5 mm range cut for all particles; minimum and maximum energies of 100 eV and 500 MeV, respectively, for electromagnetic tables; and a step-by-step upper cutoff of 100 MeV mm^{-1} (g^{-1} cm^3) for the ProtonLET scorer) (TOPAS 2016).

2.2. Single fields in water

We used the Monte Carlo models described above to calculate the absorbed dose and LET_D for simple treatment plans using PS and PBS methods. The treatment geometry comprised a single proton beam direction and a simulated, homogeneous water phantom (50 × 50 × 50 cm), as shown in figure 1. A commercial treatment-planning system (Eclipse v13.7, Varian Medical Systems, Palo Alto, CA) generated PS and PBS treatment plans to irradiate a 6 × 6 cm square, with a range of 15 cm and a modulation of 8 cm with 2 cobalt-gray equivalents (CGE). We tested whether the PS and PBS treatment plans were dosimetrically equivalent by comparing the field sizes (i.e. lateral distance between the 50% isodose lines in the *X* and *Y* directions), ranges (i.e. depth of the distal 90% isodose line), and modulation widths (i.e. distance in depth between the distal and proximal 90% isodose lines) of PS and PBS dose distributions.

We assessed the LET_D differences between the two delivery methods in three clinically relevant ROIs (figure 1). Region *A* comprised the therapeutic volume (i.e. 6 × 6 × 8 cm cube), which was equivalent to the target volume. Region *B* consisted of the volume extending 5 mm in depth from the distal edge of the therapeutic volume. The clinical equivalence of this volume was an organ at risk (OAR) immediately distal to the target volume. Region *C* comprised the volume lateral to the therapeutic volume in which the absorbed dose was between 10% and 90% relative to the therapeutic dose, known as the 90%–10% penumbra, and extended for the 8-cm modulation width of the therapeutic volume. The clinical equivalence was an OAR lateral to the target volume.

We assessed the statistical significance of differences between the LET_D distributions from the PS and PBS simulations with paired voxel-wise comparisons. We calculated the LET_D deviation between corresponding voxels, $\Delta LET_{D,i}$ as

$$\Delta LET_{D,i} = LET_{D,PBS,i} - LET_{D,PS,i}, \quad (1)$$

where $LET_{D,PBS,i}$ and $LET_{D,PS,i}$ are the LET_D values in voxel *i* from the PBS and PS simulations, respectively. We performed a one-sided Wilcoxon signed-rank test of whether the LET_D from PBS is greater than that from PS, with an $\alpha = 0.05$ significance level (i.e. H_0: $\widetilde{\Delta LET_{D,i}} \leqslant 0$; H_1: $\widetilde{\Delta LET_{D,i}} > 0$) in each ROI. We corrected the resulting *p*-values for multiple comparisons (i.e. three tests performed in related ROIs, $n = 3$) with the

Table 1. Demographic and treatment details of patients enrolled in the RT2CR and RT3CR clinical trials.

Characteristics	RT2CR	RT3CR
Number of patients	94	72
Number of males (%)	45 (48)	38 (34)
Mean age $\pm$ 1 σ (year)	9.74 $\pm$ 4.63	9.25 $\pm$ 4.40
Mean target volume $\pm$ 1 σ (cc)	28.32 $\pm$ 19.01	23.71 $\pm$ 19.12
Proton delivery method	PS	PBS
Prescribed dose (CGE)	54	54
Prescribed number of fractions	30	30
Chemotherapy status ($\pm$)	—	—

Bonferroni method. We used nonparametric statistical tests for these analyses because the assumption of normality could not be reasonably asserted (Shapiro–Wilks $p < 0.01$ in each ROI).

2.3. Multiple fields in patients

We assessed the effects of the beam delivery methods on the LET_D distributions in patients by using the treatment and CT imaging data obtained from a cohort of children with craniopharyngioma. Craniopharyngioma is a brain tumor that arises in the suprasellar region of the intracranial compartment and is intimately associated with the optic nerves and chiasm, hypothalamic–pituitary axis, and diencephalic structures. Patients in this cohort were enrolled in one of two consecutive phase 2 clinical trials (table 1). Those enrolled in the RT2CR clinical trial (NCT01419067) received PS proton therapy at the University of Florida Health Proton Therapy Institute between 2011 and 2016. Those enrolled in the RT3CR clinical trial (NCT02792582) received PBS proton therapy at the St. Jude Red Frog Events Proton Therapy Center from 2016 to 2020. The clinical trials had identical inclusion and exclusion criteria. The distribution of age, target volume, and sex did not significantly differ between the two cohorts (Student t-tests: $p = 0.49$ and 0.12 for age and target volume, respectively, and Pearson's chi-squared test $p = 0.53$ for sex).

2.3.1. Matched treatment plans

We designed PBS treatment plans for a sub-sample of 10 children who received PS proton therapy in order to minimize the confounding factors when comparing LET_D distributions. Specifically, we randomly selected 10 children from those who were enrolled in RT2CR and used their CT image sets, structure sets contoured by a board-certified radiation oncologist, and radiation therapy plan data to generate matching PBS treatment plans.

Most children enrolled in RT2CR received a three-beam treatment consisting of two lateral beams and one vertex beam. The treatments were conventionally planned to cover a planning target volume (PTV), determined by expanding the clinical target volume (CTV) by a 5-mm margin. We used a commercial treatment planning system (Eclipse v15.1, Varian Medical Systems, Palo Alto, CA) to re-plan the treatments for each of the 10 sampled patients with the number, direction, and relative weighting of PBS proton beams equal to those in the PS treatment plans. We conventionally optimized the PBS treatment plans to deliver single-field uniform doses to cover the same PTV that was used for the corresponding PS treatment plans and specifically sought to match the PS dose-volume histogram (DVH) in the PTV.

We used the Monte Carlo models described above to calculate LET_D in each voxel of the CT data for both PS and PBS treatment plans for each child. We evaluated the differences in LET_D between the pairs of PS and PBS treatment plans in the target volume (i.e. PTV) and five OARs: brainstem, left cochlea, right cochlea, left optic nerve, and right optic nerve (henceforth referred, together, as ROIs).

We used commercial software (MIM v7.0.6, MIM Software Inc., Cleveland, OH) to calculate $\Delta LET_{D,i}$ as in equation (1) and the voxel-wise dose difference (ΔD_i) as

$$\Delta D_i = D_{PBS,i} - D_{PS,i}, \tag{2}$$

where $D_{PBS,i}$ and $D_{PS,i}$ represent the dose in voxel i from PBS and PS treatment plans, respectively. We used the same software to generate ΔD- and ΔLET_D-volume histograms (ΔDVH's and ΔLVH's, respectively) for each patient and each ROI. Finally, we wrote an in-house script in MATLAB (vR2019a, Mathworks, Natick, MA) to extract 5 ΔDVH and ΔLVH metrics representing the lowest dose and LET_D differences among the highest 2%, 10%, 50%, 90%, and 98% of the ROI volume. These metrics are denoted as ΔD_{xx} and $\Delta LET_{D,xx}$ where 'xx' indicates the percentile level (i.e. 2%, 10%, 50%, 90%, or 98%). We selected these percentile levels based on similar analyses in the literature (Paelinck *et al* 2006).

We tested for the difference of dose and LET_D distributions between PS and PBS treatment plans with Wilcoxon signed rank tests. Tests were performed for each metric in each ROI with an $\alpha = 0.05$ significance level and included a Bonferroni correction for multiple comparisons (i.e. five tests in six related ROIs, $n = 30$). Wilcoxon signed rank tests for ΔDVH metrics were two tailed to test whether the PBS dose distribution differed from that from PS treatment plans (i.e. H_0: $\widetilde{\Delta D_{xx}} = 0$; H_1: $\widetilde{\Delta D_{xx}} \neq 0$). Wilcoxon signed rank tests for ΔLVH metrics were one tailed to test whether LET_D from PBS is greater than that from PS (i.e. H_0: $\widetilde{\Delta LET_{D,xx}} \leqslant 0$; H_1: $\widetilde{\Delta LET_{D,xx}} > 0$).

2.3.2. Clinical treatment plans

Next, we compared LET_D calculated for clinically delivered treatment plans from children who received PS or PBS proton therapy. The treatment data comprised the structure set contoured by a board-certified radiation oncologist and the radiation therapy plan data exported from the commercial treatment-planning system (Eclipse v13.7, Varian Medical Systems, Palo Alto, CA). In contrast to the RT2CR treatment plans (described above), most children enrolled in RT3CR received a two-beam treatment comprising two lateral beams. This was done in an effort to shift the increased LET_D expected at the end of range out of the brainstem. These treatments were planned using the clinical standard of care for PBS proton therapy: robust optimization to ensure coverage of the CTV with 95% of the prescription dose with 3%/3 mm robustness.

We used the Monte Carlo models described above to calculate LET_D in each voxel of the CT data. We evaluated the differences in LET_D between patients who received PS and those who received PBS in the six ROIs described above. In order to minimize the effects of differences in treatment-planning method on this study, we defined the target volume for this comparison as the clinical target volume, which was independent of treatment-planning method. The big-data tools of the software package ProKnow DS (Elekta, Stockholm, Sweden) calculated LET_D-volume histograms (LVH's) for each ROI and each patient.

We performed a permutation test for the population LVH's in each ROI to test whether the LVH's of patients who received PBS proton therapy were shifted to the right (i.e. exhibited higher LET_D magnitudes) of those of patients who received PS proton therapy with an $\alpha = 0.05$ significance level (i.e. H_0: $\widetilde{LET_{D,PBS}} \leqslant \widetilde{LET_{D,PS}}$; H_1: $\widetilde{LET_{D,PBS}} > \widetilde{LET_{D,PS}}$). Permutation tests provide a nonparametric method of estimating statistical significance (Camargo *et al* 2008) and are widely accepted for performing multiple comparisons without the need for additional correction (Belmonte and Yurgelun-Todd 2001, Dudoit *et al* 2003, Chen *et al* 2013). Our permutation tests comprised 1000 permutations (Edgington 1969) and compared the test statistic, T,

$$T = \sum_{LET_D=0}^{LET_{D,max}} \left(\widetilde{V_{PBS}}(LET_D) - \widetilde{V_{PS}}(LET_D) \right), \tag{3}$$

calculated in increments of 0.1 keV μm^{-1}, where $LET_{D,max}$ was the maximum LET_D and $\widetilde{V_{PBS}}(LET_D)$ and $\widetilde{V_{PS}}(LET_D)$ represented the median cumulative volume magnitudes at $x = LET_D$ for the PBS and PS cohorts of each permutation, respectively.

3. Results

3.1. Single fields in water

Comparing the PS and PBS dose distributions revealed minimal differences: the proton range agreed within 2.7 mm; modulation width agreed within 3.4 mm; and field size agreed within 0.2 mm and 0.1 mm in the X and Y directions, respectively. Figure 2 depicts plots of representative profiles of the relative absorbed dose and LET_D distributions calculated by simulations of single PS and PBS fields in water. These plots show that although the absorbed doses from PS and PBS were comparable, the calculated LET_D values were greater for PBS. Specifically, LET_D from PBS was on average 1.56-fold (95% confidence interval 0.96–2.16) greater than that from PS in the region where the absorbed dose was $\geqslant$1%. The magnitude of LET_D deviation, however, varied with location relative to the target volume.

Figure 3 shows violin plots of the distribution of LET_D deviations (ΔLET_D) in each ROI of the computational water phantom. The width of each violin plot represents the probability density of the corresponding ΔLET_D magnitude. These plots show that ΔLET_D was consistently greater than zero (i.e. LET_D from PBS $>$ LET_D from PS) in all three ROIs (figure 1). The largest deviations were observed in Region *B*, the volume distal to the target volume, with a mean ΔLET_D of 2.86 (95% confidence interval 1.50–4.21). These results were significant (Wilcoxon signed-rank test: $p < 0.01$) for each ROI.

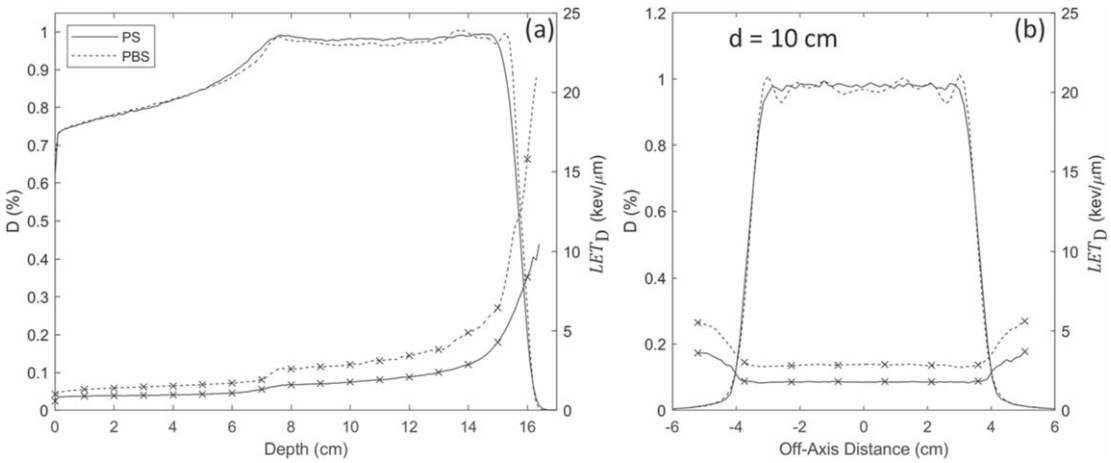

Figure 2. Profiles of relative absorbed dose (D) and dose-weighted average linear energy transfer (LET_D) from simulations of single passively scattered (PS) and pencil-beam-scanned (PBS) proton-therapy beams in a computational water phantom. LET_D is shown by the x-marked lines, and the magnitude is indicated on the right axis. (a) Profile in depth at the center of the treatment field. (b) Profile in the X-plane at a depth of 10 cm. The profile in the Y-plane was comparable.

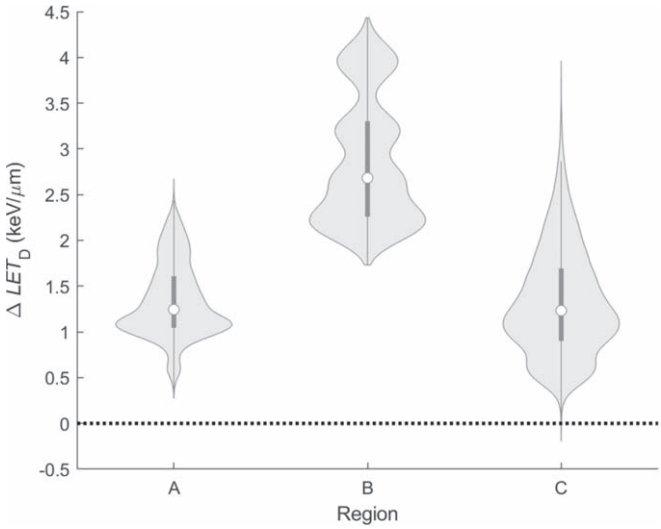

Figure 3. Violin plots of the voxel-wise deviations of the dose-weighted average linear energy transfer (ΔLET_D) in each of the three regions of interest in a water phantom (see figure 1). The width of each violin plot represents the probability density of each ΔLET_D magnitude. White circles represent the median of each ΔLET_D distribution; thick vertical bars show the inter-quartile range, and thin bars represent 1.5 × inter-quartile range.

3.2. Multiple fields in patients

3.2.1. Matched treatment plans

Figure 4 plots the DVH and LVH for a representative pair of matched treatment plans of both delivery methods. Figure 4(a) shows that although the dose distributions from the two delivery methods were equivalent in the PTV, PBS tended to reduce the minimum dose in the surrounding OAR's compared to PS. Despite this, figure 4(b) shows that PBS produced higher LET_D magnitudes in all ROI's (i.e. both inside and outside the target volume) than did PS.

Figure 5 plots the distributions of ΔDVH metrics in each ROI among the sub-sample of 10 children (see section 2.3.1). ΔDVH metrics were most uniform in the PTV, where we specifically aimed to match the PBS and PS dose distributions. This figure also shows that ΔDVH metrics tended to be below null in the surrounding OAR's, indicating that PBS tended to reduce the doses in OAR's compared to PS. Interestingly, the lower-percentile metrics (e.g. ΔD_{02} and ΔD_{10}) tended to be similar between the delivery methods across all OAR's. These metrics are closely related to the maximum dose in an organ. The largest dose differences between the delivery methods appeared in the higher-percentile metrics (ΔD_{90}, ΔD_{98}). These metrics relate to the minimum organ dose and indicate that, as expected, PBS decreases minimum organ doses below the levels

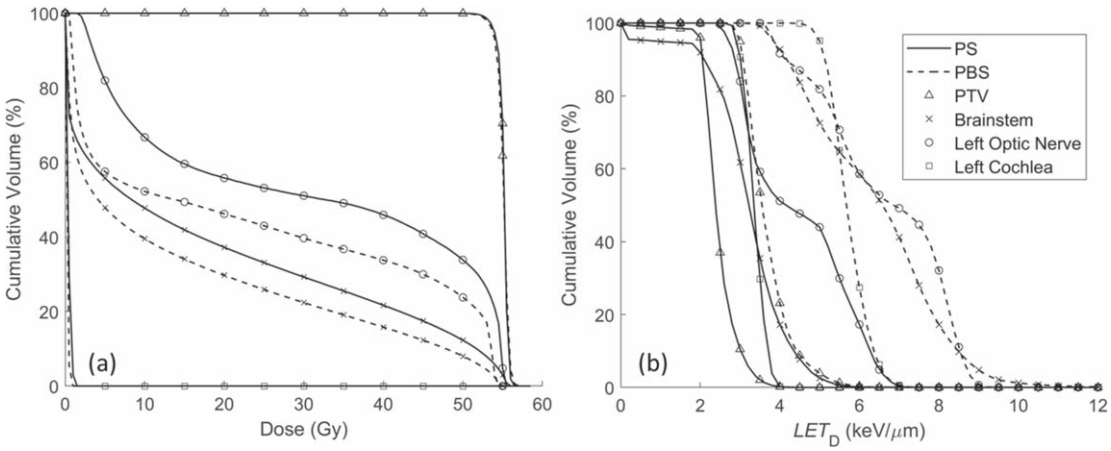

Figure 4. Representative (a) dose- and (b) LET_D-volume histograms for matched passively scattered (PS) and pencil-beam scanned (PBS) treatment plans. Curves for left-sided regions of interest are representative of those from the contralateral side. (PTV: planning target volume).

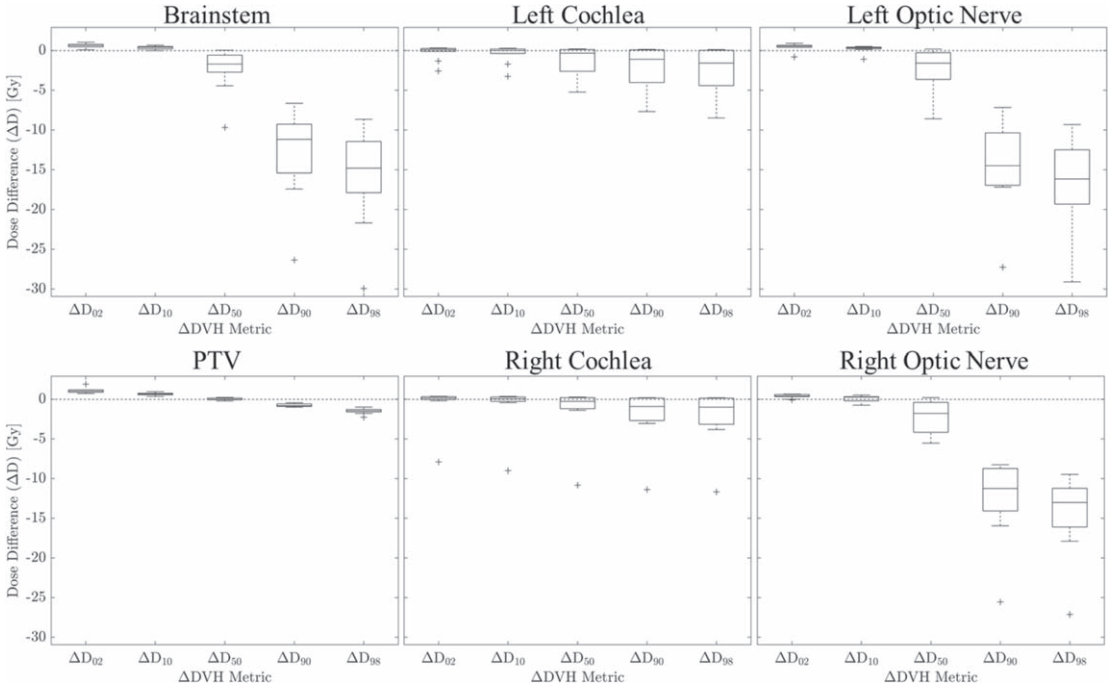

Figure 5. Box plots of dose-difference volume histogram (ΔDVH) metrics in each region of interest among 10 representative children. Dose difference was calculated as the dose from pencil-beam scanning (PBS) minus that from passive scattering (PS) (see equation (2)). The dashed line placed at null separates the region indicating the dose from PBS was higher (above the line) from the region where the dose from PS was higher (below the line). ΔD_{XX} indicates the percentile level of the dose in each region of interest (i.e. 2%, 10%, 50%, 90%, or 98%).

typical of PS treatments. Wilcoxon signed rank tests revealed no ΔDVH metrics differed from zero at the 95% confidence level (all p-values > 0.05). This indicates PS and PBS treatment plans produced statistically equivalent dose distributions.

Figure 6 plots the distributions of ΔLVH metrics in each ROI among the sub-sample of 10 children (see section 2.3.1). This figure shows that ΔLVH metrics tended to be higher than null in all ROI's, indicating that PBS tended to produce higher LET_D magnitudes than PS both inside and outside the target volume. The metric that registered the largest variation with beam delivery method across all ROI's was $\Delta LET_{D,02}$. This metric roughly represents the maximum ΔLET_D and was an average of 5.51 keV μm^{-1} higher (range [1.41, 19.36] keV μm^{-1}) for PBS treatment plans than for PS plans across all patients and all ROI's. The largest ΔLVH-metric magnitudes across all metrics were observed in the brainstem, with an average ΔLVH-metric magnitude of 4.51 keV μm^{-1} (range [0.69, 19.36] keV μm^{-1}) across all patients and all metrics. Wilcoxon signed rank tests

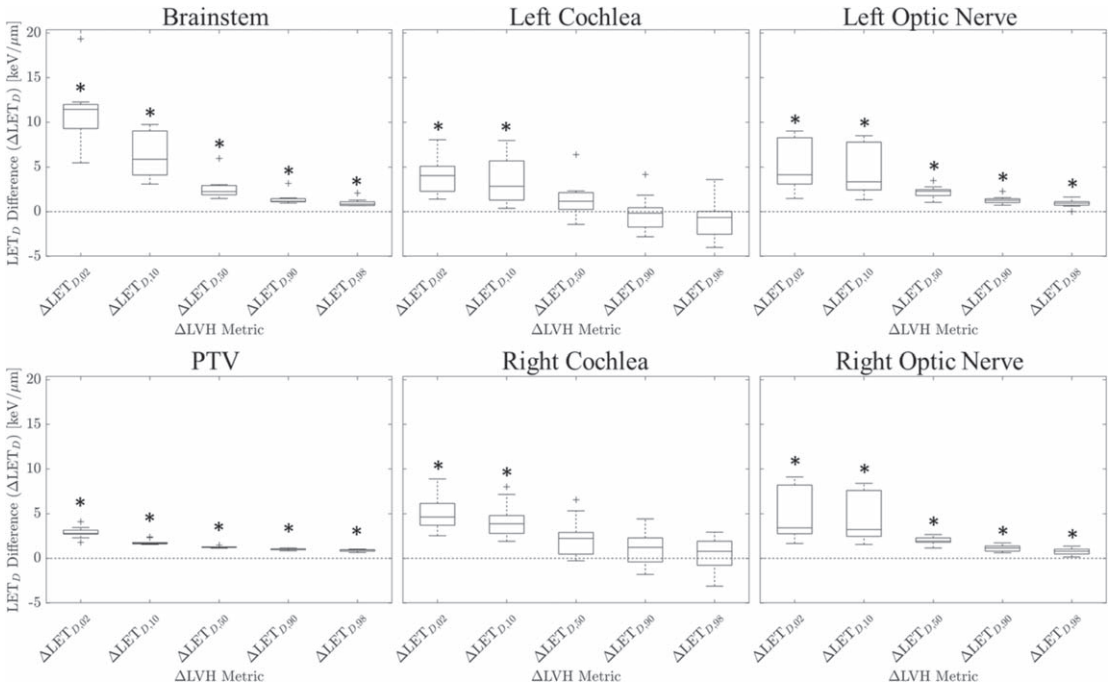

Figure 6. Box plots of dose-weighted average linear energy transfer (LET_D)-difference volume histogram (ΔLVH) metrics in each region of interest among 10 representative children. LET_D difference was calculated as the LET_D from pencil-beam scanning (PBS) minus that from passive scattering (PS) (see equation (1)). The dashed line placed at null separates the region indicating the LET_D from PBS was higher (above the line) from the region where the LET_D from PS was higher (below the line). $\Delta LET_{D,XX}$ indicates the percentile level of the LET_D in each region of interest (i.e. 2%, 10%, 50%, 90%, or 98%). * indicates statistical significance as determined by Wilcoxon signed rank tests at the $\alpha = 0.05$ significance level.

revealed most ΔLVH metrics were greater than zero at the 95% confidence level (p-values $\leqslant 0.05$ for 26 of 30 ΔLVH metrics tested, indicated by asterisks in figure 6).

3.2.2. Clinical treatment plans

Figure 7 shows population LVH's from the clinical treatment plans of patients who received PS or PBS proton therapy (table 1) along with the corresponding p-values of the permutation tests. LVH comparisons in the target volume, brainstem, and left and right optic nerves significantly differed ($p < 0.05$), supporting the alternate hypothesis that PBS produces significantly higher LET_D values than does PS. The largest differences between median LVH curves were observed at $LET_{D,min}$, which was an average of 1.98 keV μm^{-1} greater in patients treated with PBS than in those who received PS across all of the ROIs considered. The brainstem exhibited the largest LVH differences across all LET_D magnitudes, with an average absolute LVH separation of 4.27% in cumulative volume.

4. Discussion

In this study, we assessed the extent to which proton delivery methods affect LET_D by using two cohorts of children who were treated with successive proton therapy protocols using PS or PBS methods. We calculated LET_D from simple, matched, and clinical proton therapy plans with Monte Carlo methods and assessed the differences between that produced by PS and PBS methods. The major finding of this work is that although both delivery methods produced comparable physical dose distributions within the target and in immediately adjacent critical structures, PBS produced higher LET_D values than did PS inside and outside the targeted volume.

A major implication of these findings is that the biologic effectiveness of the physical dose, which purportedly depends on LET_D (Perris *et al* 1986, Belli *et al* 1989, 1992, Folkard *et al* 1996, Paganetti *et al* 2019), most likely differs between the delivery methods. Therefore, dose-based methods to evaluate PS treatment plans and outcomes may not directly apply to patients receiving PBS. Dose-volume constraints are an integral component of conventional treatment-planning methods. Recent studies attempted to directly optimize the clinical goals of treatment by using dose-response models (Wilkens and Oelfke 2005, Allen *et al* 2012, Rechner *et al* 2015, Modiri *et al* 2018). Both of these methods, however, rely on scaling proton doses by a delivery method-invariant factor of 1.1. This factor aims to account for the biologic effectiveness of proton therapy relative to that

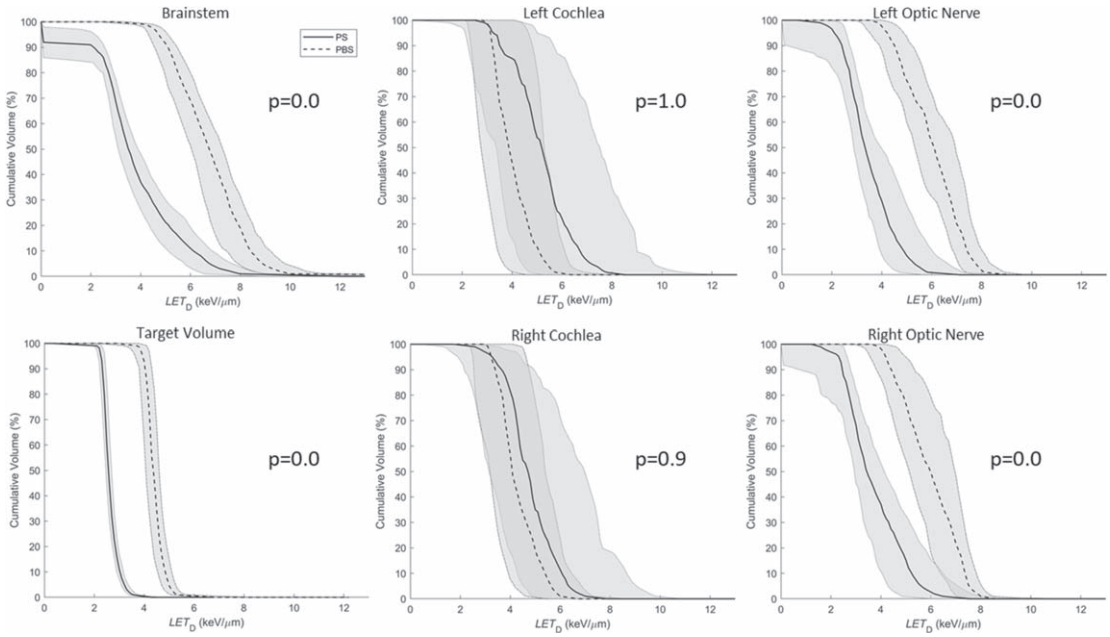

Figure 7. Population LET_D-volume histograms (LVH's) for the target and organs at risk from delivered passively scattered (PS) and pencil beam-scanned (PBS) proton therapy plans for a cohort of children with craniopharyngioma (see table 1). The solid and dashed lines represent the median LVH's from PS and PBS cohorts, respectively. The shaded regions around each LVH line represent the interquartile range. The *p*-values indicate the results of permutation tests as described in the Methods.

of photon therapy. Dose-volume metrics and dose-response models determined over a century of photon therapy experience are then applied to proton therapy treatment planning. This scaling factor, however, was largely determined from studies using the PS delivery method (Paganetti *et al* 2002). The results of our study suggest that a unique scaling factor may be required for each delivery method to ensure the safety and effectiveness of the treatments, whether they are planned by conventional or advanced treatment-planning methods.

Our findings are consistent with those of previously published studies. Most of these studies focused on the maximum LET_D, $LET_{D,max}$, in various ROIs. Paganetti (2014) reviewed variations in LET_D with range and modulation width of PS proton therapy beams in water. Among all of the range and modulation-width combinations considered, the $LET_{D,max}$ was ~8 keV μm^{-1} and 12 keV μm^{-1} within the volumes receiving 90% and 2% of the therapeutic dose, respectively. Our data from a PS beam in water revealed similar $LET_{D,max}$ values of 5 keV μm^{-1} and 10 keV μm^{-1} within the 90% and 2% relative isodose lines, respectively. The lower $LET_{D,max}$ magnitudes in our study can be explained by the larger modulation width in our treatment plan, which decreased LET_D (Paganetti 2014). Grassberger *et al* (2011) reported that the $LET_{D,max}$ was ~12 keV μm^{-1} within a single PBS treatment field and ~8 keV μm^{-1} within a single PS treatment field. Likewise, our data from single fields in a water phantom revealed that the $LET_{D,max}$ within the 50% isodose line was 12.8 keV μm^{-1} and 7.5 keV μm^{-1} from the PBS and PS beams, respectively. In this same publication, however, Grassberger *et al* stated that the large LET_D discrepancies at the end of range of single PS and PBS beams were 'significantly reduced by using multiple fields.' Conversely, our findings from clinical treatment plans consisting of two or three treatment beams indicated that using multiple fields does not eliminate the increased LET_D with PBS, as compared with the LET_D of PS proton therapy. Limited information regarding the treatment plans compared by Grassberger *et al* hinders identifying the source of this discrepancy. Finally, Gridley *et al* (2015) studied variations in RBE according to the delivery method and reported enhanced biologic responses to PBS over that of PS in lung epithelial cells. They attributed this phenomenon to the increased instantaneous dose rate of PBS compared to PS. Our findings suggest that elevated LET_D, on which the RBE depends and which is dose-rate independent, may be partly responsible.

This study has several notable strengths. First, including simple treatment plans comprising single fields in a computational water phantom minimized any potential confounding factors in our analysis and increased the reproducibility of our findings in future studies. We also minimized confounding factors by using two populations of patients treated in successive clinical trials and creating dosimetrically matched treatment plans for a sub-sample of those patients. Pediatric patients in these trials received identical treatment prescriptions and were treated over a brief time. Finally, our results represent the largest dataset depicting LET_D characteristics in a homogeneous population. The large sample size of each clinical trial population provided a comprehensive view

of LET_D in clinical settings, which is strengthened by the statistical significance of our comparisons between the two delivery methods.

This study also has some notable limitations. First, the Monte Carlo model of the PS beamline used a different version of TOPAS than did that of the PBS beamline. We did not consider this a substantial limitation, however, because both versions used the same proton LET_D scorer. Furthermore, the initial version of the PBS model was developed for TOPAS v1.3 and later converted to TOPAS v2. We observed no notable discrepancies in the dose or LET_D calculated with either version of TOPAS (Moskvin *et al* 2016b). The difference in the beam arrangements for the two cohorts is another limitation of our study. Most patients in the RT2CR trial were treated with a three-beam arrangement, comprising two lateral beams and a third apex beam. Most RT3CR treatment plans, however, were limited to two lateral beams. Although some of the LET_D differences we observed in the analysis of clinical treatment plans may be attributable to differences in the beam arrangement rather than the delivery method, we do not consider this a meaningful limitation. Specifically, we considered one patient in the RT3CR trial whose treatment was planned with both the two- and three-beam arrangements of the same delivery method (i.e. PBS). Directly comparing the LET_D from these plans revealed only minor differences in the LET_D in each ROI according to beam arrangement. Furthermore, in all but one ROI, the two-beam arrangement led to lower LET_D than did the three-beam arrangement. Therefore, variations in the beam arrangement between treatment plans of differing delivery methods most likely attenuated the delivery method-based LET_D differences. Furthermore, the results from our analyses of clinical treatment plans were consistent with those of single fields in water and matched treatment plans in patients, which were free from this limitation. Finally, although we found that LET_D differs between delivery methods, we did not assess the clinical significance of these differences. The degree to which LET_D affected RBE, and in turn clinical outcomes, remains unknown.

Importantly, the increased biological effectiveness of PBS over that of PS indicated by our results does not imply that PS is safer than PBS. These results merely suggest that the relation between physical dose and biologic effect differs between the delivery methods. Therefore, each method requires unique considerations (e.g. scaling factors) to ensure safe usage. Furthermore, our findings do not reflect differences in the proximal conformality of the dose distributions nor in the dose deposited by secondary particles with delivery method, two areas in which PBS is known to provide improvements over PS (Newhauser and Durante 2011).

5. Conclusion

This study is the first to systematically compare LET_D differences on the basis of proton delivery method for single fields in water, and dosimetrically matched and clinical treatment plans in a cohort of patients with a similar tumor type. These comparisons revealed significantly increased LET_D from PBS over that of PS inside and outside of the targeted volume in a water phantom. This enhanced LET_D also occurred in the ROIs of a cohort of pediatric patients with craniopharyngioma. These findings are important because early studies seeking to establish the relation of LET_D to health outcomes were performed primarily with PS systems. Our findings suggest that such studies, as well as other dosimetric proton therapy studies, may not be directly translatable to modern PBS treatments. Given the rapid expansion of PBS proton therapy availability and use, future studies are needed to confirm the applicability of PS-derived scaling factors (i.e. uniform RBE of 1.1) to modern PBS treatments to ensure their safety and effectiveness for the growing population of proton therapy patients.

Acknowledgments

The authors would like to thank Jared Becksfort, PhD, for his helpful conversations regarding statistical analyses via permutation testing. This work was funded by American Lebanese Syrian Associated Charities. Portions of this research were conducted with high-performance computing resources provided by the Department of Information Services at St Jude Children's Research Hospital.

Ethical statement

This was a retrospective study involving data from two clinical trials (RT2CR: NCT01419067; RT3CR: NCT02792582). Both clinical trials were carried out in accordance with the principles outlined in the IOPScience ethical policy and were reviewed and approved by the St. Jude Institutional Review Board (Institutional Review Board #00000029 FWA00004775; approval numbers: RT2CR IRB# PRO00002248, RT3CR IRB# PRO00006397).

References

Allen L X *et al* 2012 The use and QA of biologically related models for treatment planning: short report of the TG-166 of the therapy physics committee of the AAPM *Med. Phys.* **39** 1386–409

ASTRO 2014 *Model Policies: Proton Beam Therapy* American Society for Radiation Oncology

Belli M, Cera F, Cherubini R, Ianzini F, Moschini G, Sapora O, Simone G, Tabocchini M A and Tiveron P 1992 RBE-LET relationship for survival and mutation induction of V79 cells irradiated with low-energy protons: re-evaluation of the LET values at the LNL facility *Int. J. Radiat. Biol.* **61** 145–6

Belli M, Cherubini R, Finotto S, Moschini G, Sapora O, Simone G and Tabocchini M A 1989 RBE-LET relationship for the survival of V79 cells irradiated with low energy protons *Int. J. Radiat. Biol.* **55** 93–104

Belmonte M and Yurgelun-Todd D 2001 Permutation testing made practical for functional magnetic resonance image analysis *IEEE Trans. Med. Imaging* **20** 243–8

Camargo A, Azuaje F, Wang H and Zheng H 2008 Permutation—based statistical tests for multiple hypotheses *Source Code Biol. Med.* **3** 15

Chen C, Witte M, Heemsbergen W and van Herk M 2013 Multiple comparisons permutation test for image based data mining in radiotherapy *Radiat. Oncol.* **8** 293

Coutrakon G, Cortese J, Ghebremedhin A, Hubbard J, Johanning J, Koss P, Maudsley G, Slater C R and Zuccarelli C 1997 Microdosimetry spectra of the Loma Linda proton beam and relative biological effectiveness comparisons *Med. Phys.* **24** 1499–506

Dudoit S, Shaffer J P and Boldrick J C 2003 Multiple hypothesis testing in microarray experiments *Statist. Sci.* **18** 71–103

Edgington E S 1969 Approximate randomization tests *J. Psychol.* **72** 143–9

Farr J B, Moskvin V, Lukose R C, Tuomanen S, Tsiamas P and Yao W 2018 Development, commissioning, and evaluation of a new intensity modulated minibeam proton therapy system *Med. Phys.* **45** 4227–37

Folkard M, Prise K M, Vojnovic B, Newman H C, Roper M J and Michael B D 1996 Inactivation of V79 cells by low-energy protons, deuterons and helium-3 ions *Int. J. Radiat. Biol.* **69** 729–38

Granville D A and Sawakuchi G O 2015 Comparison of linear energy transfer scoring techniques in Monte Carlo simulations of proton beams *Phys. Med. Biol.* **60** N283–91

Grassberger C and Paganetti H 2011 Elevated LET components in clinical proton beams *Phys. Med. Biol.* **56** 6677–91

Grassberger C, Trofimov A, Lomax A and Paganetti H 2011 Variations in linear energy transfer within clinical proton therapy fields and the potential for biological treatment planning *Int. J. Radiat. Oncol. Biol. Phys.* **80** 1559–66

Gridley D S, Pecaut M J, Mao X W, Wroe A J and Luo-Owen X 2015 Biological effects of passive versus active scanning proton beams on human lung epithelial cells *Technol. Cancer Res. Treat.* **14** 81–98

Jermann M 2019 *Particle Therapy Facilities in Clinical Operation* Particle Therapy Co-Operative Group (PTCOG)

Kaluarachchi M, Moskvin V, Pirlepesov F, Wilson L J, Xie F and Faught A M 2020 Automation of Monte Carlo-based treatment plan verification for proton therapy *J. Appl. Clin. Med. Phys.* **21** 131–8

Modiri A, Stick L B, Rice S R, Rechner L A, Vogelius I R, Bentzen S M and Sawant A 2018 Individualized estimates of overall survival in radiation therapy plan optimization—a concept study *Med. Phys.* **45** 5332–42

Moskvin V 2016 The dose weighted LET and relative biological effectiveness (RBE) for DNA double strand break (DSB) induction for IMPT with spot scanning pencil beam and for conventional technique using broad proton beams with patient specific aperture and compensator *Int. J. Part. Ther.* **3** 338

Moskvin V, Pirlepesov F, Stewart R and Farr J 2016a MO-FG-CAMPUS-TeP3-05: limitations of the dose weighted LET concept for intensity modulated proton therapy in the distal falloff region and beyond *Med. Phys.* **43** 3728

Moskvin V, Pirlepesov F, Tsiamas P, Axente M, Lukose R, Zhao L, Shin J and Farr J 2016b SU-F-T-156: Monte Carlo simulation using TOPAS for synchrotron based proton discrete spot scanning system *Med. Phys.* **43** 3497–8

Moskvin V, Pourmoghaddas A, Pirlepesov F, Yao W and Farr J 2017 Multistage Monte Carlo treatment simulation framework for Hitachi Probeat-V intensity modulated proton primary and micro-beam system *Med. Phys.* **44** 2773

Newhauser W D *et al* 2009 The risk of developing a second cancer after receiving craniospinal proton irradiation *Phys. Med. Biol.* **54** 2277–91

Newhauser W D and Durante M 2011 Assessing the risk of second malignancies after modern radiotherapy *Nat. Rev. Cancer* **11** 438–48

Paelinck L, Smedt B D, Reynaert N, Coghe M, Gersem W D, Wagter C D, Vanderstraeten B, Thierens H and Neve W D 2006 Comparison of dose-volume histograms of IMRT treatment plans for ethmoid sinus cancer computed by advanced treatment planning systems including Monte Carlo *Radiother. Oncol.* **81** 250–6

Paganetti H 2014 Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer *Phys. Med. Biol.* **59** R419–72

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 and Goitein M 2001 Biophysical modelling of proton radiation effects based on amorphous track models *Int. J. Radiat. Biol.* **77** 911–28

Paganetti H, Niemierko A, Ancukiewicz M, Gerweck L E, Goitein M, Loeffler J S and Suit H D 2002 Relative biological effectiveness (RBE) values for proton beam therapy *Int. J. Radiat. Oncol. Biol. Phys.* **53** 407–21

Paganetti H and Schmitz T 1996 The influence of the beam modulation technique on dose and RBE in proton radiation therapy *Phys. Med. Biol.* **41** 1649–63

Perez-Andujar A, Newhauser W D, Taddei P J, Mahajan A and Howell R M 2013 Predicted relative risk of follicular depletion for three radiotherapy modalities in a girl receiving craniospinal irradiation *Phys. Med. Biol.* **58** 3107–23

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

Perris A, Pialoglou P, Katsanos A A and Sideris E G 1986 Biological effectiveness of low energy protons. I. Survival of Chinese hamster cells *Int. J. Radiat. Biol. Relat. Stud. Phys., Chem., Med.* **50** 1093–101

PTCOG 2020 Particle Therapy Co-Operative Group Particle Therapy Centers (https://ptcog.ch/index.php/facilities-in-operation). ed PTCOG

Rechner L A, Eley J G, Howell R M, Zhang R, Mirkovic D and Newhauser W D 2015 Risk-optimized proton therapy to minimize radiogenic second cancers *Phys. Med. Biol.* **60** 3999–4013

Shin J, Merchant T E, Lee S, Li Z, Shin D and Farr J B 2015 SU-E-T-08: a beam source model for Monte Carlo Simulations of a double-scattering proton beam delivery system using beam current modulation *Med. Phys.* **42** 3332

TOPAS 2016 TOPAS User Manual: Modular Physics Lists (https://topas.readthedocs.io/en/latest/parameters/physics/modular.html). (GitHub: TOPAS MC Inc.)

TOPAS 2020 Physics Lists EM constructors in Geant4 10.4 (https://geant4.web.cern.ch/node/1731#opt4)
Warwick S 2018 *Annual Survey Shows Surge in Cancer Patients Treated with Proton Therapy & in Number of Indications* (Scottsdale, AZ: National Association for Proton Therapy)
Wilkens J J and Oelfke U 2004 A phenomenological model for the relative biological effectiveness in therapeutic proton beams *Phys. Med. Biol.* **49** 2811–25
Wilkens J J and Oelfke U 2005 Optimization of radiobiological effects in intensity modulated proton therapy *Med. Phys.* **32** 455–65
Wouters B G, Lam G K, Oelfke U, Gardey K, Durand R E and Skarsgard L D 1996 Measurements of relative biological effectiveness of the 70 MeV proton beam at TRIUMF using Chinese hamster V79 cells and the high-precision cell sorter assay *Radiat. Res.* **146** 159–70
Zhang R, Howell R M, Homann K, Giebeler A, Taddei P J, Mahajan A and Newhauser W D 2013 Predicted risks of radiogenic cardiac toxicity in two pediatric patients undergoing photon or proton radiotherapy *Radiat. Oncol.* **8** 184