BJR

https://doi.org/10.1259/bjr.20190919

Received: 30 October 2019 | Revised: 20 January 2020 | Accepted: 24 January 2020

Cite this article as:
Winterhalter C, Aitkenhead A, Oxley D, Richardson J, Weber DC, MacKay RI, et al. Pitfalls in the beam modelling process of Monte Carlo calculations for proton pencil beam scanning. *Br J Radiol* 2020; **93**: 20190919.

PROTON THERAPY SPECIAL FEATURE: FULL PAPER

Pitfalls in the beam modelling process of Monte Carlo calculations for proton pencil beam scanning

[1,2]CARLA WINTERHALTER, [3,4]ADAM AITKENHEAD, [1]DAVID OXLEY, [3]JENNY RICHARDSON, [1,5,6]DAMIEN C. WEBER, [3,4]RANALD I. MACKAY, [1,2]ANTONY J. LOMAX and [1]SAIROS SAFAI

[1]Centre for Proton Therapy, Paul Scherrer Institute, Villigen, Switzerland
[2]Department of Physics, ETH Zürich, Switzerland
[3]Christie Medical Physics and Engineering, The Christie NHS Foundation Trust, Manchester, UK
[4]Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK
[5]Department of Radiation Oncology, University Hospital of Bern, Bern, Switzerland
[6]Department of Radiation Oncology, University Hospital of Zürich, Zürich, Switzerland

Address correspondence to: Dr Carla Winterhalter
E-mail: *carla.winterhalter@manchester.ac.uk*

Objective: Monte Carlo (MC) simulations substantially improve the accuracy of predicted doses. This study aims to determine and quantify the uncertainties of setting up such a MC system.
Methods: Doses simulated with two Geant4-based MC calculation codes, but *independently* tuned to the *same* beam data, have been compared. Different methods of MC modelling of a pre-absorber have been employed, either modifying the beam source parameters (descriptive) or adding the pre-absorber as a physical component (physical).
Results: After the independent beam modelling of both systems in water (resulting in excellent range agreement) range differences of up to 3.6/4.8 mm (1.5% of total range) in bone/brain-like tissues were found, which resulted from the use of different mean water ionisation potentials during the energy tuning process. When repeating using a common definition of water, ranges in bone/brain agreed within 0.1 mm and gamma-analysis (global 1%,1mm) showed excellent agreement (>93%) for all patient fields. However, due to a lack of modelling of proton fluence loss in the descriptive pre-absorber, differences of 7% in absolute dose between the pre-absorber definitions were found.
Conclusion: This study quantifies the influence of using different water ionisation potentials during the MC beam modelling process. Furthermore, when using a descriptive pre-absorber model, additional Faraday cup or ionisation chamber measurements with pre-absorber are necessary.
Advances in knowledge: This is the first study quantifying the uncertainties caused by the MC beam modelling process for proton pencil beam scanning, and a more detailed beam modelling process for MC simulations is proposed to minimise the influence of critical parameters.

INTRODUCTION

Due to the favourable depth dose characteristics of protons when compared to conventional photon therapy, the number of proton centres is growing globally. Most of the new centres are built for proton pencil beam scanning (PBS), as its intrinsic dose painting potential[1] can be used to maximise the dose to the tumour while sparing nearby organs at risk.

To ensure good treatment outcome however, accurate dose calculations to predict and plan the delivered dose are critical. At treatment sites with complex inhomogeneities, there is an uncertainty in the proton range predicted by analytical algorithms, and Monte Carlo (MC) particle transport simulations can substantially improve the accuracy of the predicted doses.[2] Therefore, many centres are now using MC calculation algorithms for proton PBS (*e.g.* Yepes et al, Tommasino et al, Widesott et al[3–5]).

Multiple MC codes are available for medical physics applications, as for example Geant4[6] (and Geant4 wrappers GATE,[7] TOPAS[8] and GAMOS[9]), PHITS,[10] MCNP[11] and FLUKA.[12] To investigate the differences between MC algorithms in a clinical setting, in[13] Geant4 and a MCNPX-based MC calculation algorithms were compared. Since these two models had been set up for two different proton therapy treatment heads however, only monoenergetic beams, and not patient dose distributions were compared in the patient CT. For Geant4 applications, physics settings for proton therapy have been thoroughly analysed in

water and in PMMA,[14] as well as in water and for a Faraday cup model,[15] with both studies resulting in optimised physics and parameter settings. Additionally, the influence of ionisation values on proton ranges has been evaluated for monoenergetic proton beams in water/tissue cylinders.[16]

Multipurpose MC codes such as Geant4, however, are highly flexible, and are not trivial to set-up and configure for non-expert users in the clinical environment. For ease of configuration of Geant4[6] for medical physics applications therefore, two toolkits are available—GATE (Geant4 Application for Tomographic Emission[7]) and TOPAS (TOol for PArticle Simulations[8]). Both allow the user to define complex simulation setups in text based parameter files, and MC calculations for proton PBS using these toolkits have been set up at multiple institutes.[17–20] Indeed, with steadily increasing numbers of proton treatment centres, more and more will set up institute specific MC calculations. As such, GATE and TOPAS facilitate the use of Geant4 by *users* (such as Medical Physicists) without requiring them to have *developer* level experience of Geant4. Even using such toolkits however, MC calculations for proton PBS are not yet off-the-shelf tools, and there is a lack of comparison of MC descriptions of proton PBnoS Gantries for clinical proton dose calculations which take into account the whole setup and beam modelling process.

In this study, at two different institutions, similar MC calculation systems, one based on GATE and one TOPAS, have been *independently* developed and tuned using the *same* measured data from a proton PBS Gantry. In this collaborative study, dose distributions calculated in both systems were then compared in simple geometric setups and in clinical patient calculations. This study aimed to identify how choices during the configuration of MC simulation systems influence and potentially bias the MC simulation results and how this can be prevented. It therefore highlights potential pitfalls during the MC beam modelling process and concludes with an updated beam modelling procedure to prevent these.

The study is structured as follows: first, the setup of the MC systems is presented. Next, the system validation and the comparison of dose results in simple geometric and clinical patient cases are shown. Based on these results, critical factors when setting up a MC system for proton PBS are identified and the dose differences due to the different beam modelling discussed. Finally, an updated beam modelling procedure is introduced to mitigate these factors and lead to more consistent MC results.

METHODS AND MATERIALS

Setting up a PBS MC system

To set up a PBS MC system, the underlying MC code needs to be chosen and physics models defined. Additionally, institute-specific geometry and beam models need to be implemented. In the following, the beam modelling process is explained (see "Beam modelling for a PBS MC system") and the two setups are explained ("PSI and The Christie MC Systems") and compared in more detail ("System comparison").

Beam modelling for a PBS MC system

For passive scattering proton therapy, a detailed model of the treatment head needs to be included in the MC simulation (see, *e.g.* Verburg et al[21]). In contrast however, for PBS proton therapy, even though there is a low-dose spray due to interactions within the beam line and Gantry,[19,22] MC simulations can be started at a defined point before the patient, and not all hardware needs to be explicitly modelled.[17–21] At this starting point, each pencil beam is described by its number of protons, energy, energy spread, and initial optical phase space (beam size, angular spread and correlation). In this study, proton numbers per monitor units have been defined based on Faraday cup measurements (see Winterhalter et al[17] for more details on the absolute dose validation of this approach), nominal energy and energy spreads have been iteratively adjusted to reproduce depth–dose curves in water and the initial optical phase space characterised by subtracting the air scattering contribution from the measured beam data (see Grevillot et al[20] for more detail). Additionally, to deliver pencil beams with energies below the lowest range that can be transported through the beam line (70 MeV for the PSI Gantry 2), a pre-absorber (range shifter) is typically required. This pre-absorber can either be modelled as a *physical* component, whereby all protons are tracked through the pre-absorber as part of the MC simulations,[17,19,23] or as a *descriptive* component by adjusting the beam energy, energy spread and phase space after the pre-absorber accordingly.[18] These data are then used to perform a pre-absorber specific beam modelling of the MC.

In summary, based on previous literature, the beam modelling process contains the following steps:

(1) Without pre-absorber:
- Characterise the beam optics (initial phase space: beam size, angular spread and correlation) of pencil beams in air (for each nozzle extension).
- Tune energy and energy spread to reproduce measured depth–dose curves in water.

(2) With a *physical* pre-absorber
- Tune material properties of the pre-absorber (*e.g.* density within manufacturer specifications) to match depth–dose curves in water measured with the pre-absorber

OR With a *descriptive* pre-absorber
- Characterise initial phase space of pencil beams in air (for each nozzle extension) to match measurements of beam sizes in air with the pre-absorber.
- Tune energy and energy spread to reproduce depth–dose curves in water measured with the pre-absorber.

(3) Convert CT voxels to material and density.

It is worth mentioning that the PSI Gantry 2 has a telescopic nozzle, which moves the pre-absorber and all nozzle monitors in relation to the isocentre.[24] The phase space parametrisation (for the descriptive pre-absorber and also without pre-absorber due to the changing positions of nozzle monitors), therefore needs to be repeated for each nozzle extension.

PSI and The Christie MC Systems

Two MC models, tuned to the same PSI Gantry 2 commissioning measured beam data, have been independently developed at two different institutes (PSI and The Christie). These beam data consisted of integral depth–dose curves measured with a Bragg peak chamber (diameter 8 cm) and spot sizes in air measured with a scintillating screen.

Both systems are based on the Geant4 10.02.p01 MC code, with the PSI system being modelled with the TOPAS 3.0.p1 and The Christie system with the GATE 7.2 toolkit. Physics have been defined as the TOPAS default physics list[8,15] (consisting of the G4EMStandardPhysics_option4, HadronPhysicsQGSP_BIC_HP, G4DecayPhysics, G4IonBinaryCascadePhysics, G4HadronElasticPhysicsHP and G4StoppingPhysics Geant4 classes) in the PSI system and the QGSP_BIC Geant4 reference list (chosen based on[25] and[26]) in The Christie model. Range cuts of 0.05 mm (PSI system) and 0.1 mm (The Christie system) have been used. No variance reduction techniques were applied.

For the PSI model, simulations started at the upstream side of the pre-absorber with the nozzle fully retracted (47.8 cm upstream of isocenter) and for The Christie at the primary dose monitor (74.1 cm upstream of isocenter). In addition, the pre-absorber has been modelled as a physical component at PSI and as a descriptive component (as used in Fracchiolla et al,[18] see "Setting up a PBS MC system" above) by The Christie. Both systems apply the same number of protons per pencil beam and the two beam models have been tuned independently (Beam modelling for a PBS MC system). Supplementary Table 1 summarises the settings of both systems.

Finally, for both systems, the same CT calibration has been applied, derived using the calibration described by Schneider et al.[27] In this process, the CT scanner is first characterised by scanning a set of materials of known composition and density, and for which the proton stopping power can be theoretically calculated. The scanner response is then characterised by parametrising the relationship between the theoretical stopping powers and the measured Hounsfield unit (HU) values for each material.[28] Using this parametrisation, predicted HU values are then calculated for a set of 71 reference human tissues. This provides a table which acts as a calibration for the CT scanner, and which can be applied to any patient CT image to provide a mapping between a voxel's measured HU value and a corresponding reference tissue definition (in terms of material composition and density).

System comparison

Water and air

For three example pencil beams without pre-absorber (energies 71, 163 and 229 MeV) and three with pre-absorber (86, 163 and 229 MeV), simulations in air and water have been performed, using the same geometry and scoring setup for both MC models. Doses were scored and compared in water/air cuboids with 10 x 10 cm cross-sections and 36 cm (water)/50 cm (air) depths. Scoring dimensions were 0.5 mm (in depth) x 1 x 1mm and 10^7 protons were simulated for each pencil beam, and beam sizes (σ of the Gaussian beam) in air, depth–dose curves and absolute doses scored in the whole water cuboid compared.

The initial phase space and the position of the pre-absorber corresponds to a fully extracted nozzle, and, in accordance with the commissioning measurements, the surface of the water phantom has been positioned at the isocenter/13.1 cm upstream of isocenter (without/with pre-absorber). With pre-absorber, this leads to an air-gap of 6.2 cm between the downstream surface of the pre-absorber and the upstream surface of the water phantom.

Material blocks

For three pencil beams (71 MeV (no pre-absorber)/86 MeV (with pre-absorber), 163 MeV, 229 MeV), simulated doses for both MC setups have been compared in brain and in bone like cuboids with the same physical and scoring dimensions as the water cuboid. Brain (density = 1.04 g/cm^3) and bone (density = 1.42 g/cm^3) are defined according to ICRU_46[29] (Table A2).

Patient CT

Dose distributions have also been compared in a patient CT data (CT voxel dimensions 1.7 × 1.7 × 2.0 mm; dose grid voxel dimensions 2.0 × 2.0 × 2.0 mm) for a three field SFUD (single field uniform dose) plan without pre-absorber and a four-field IMPT (intensity modulated proton therapy) plan with pre-absorber (CT voxel dimensions 1.2 × 1.2 × 2.0 mm; dose grid voxel dimensions 2.0 × 2.0 × 2.0 mm). Proton numbers for each spot are based on Faraday cup measurements.[17]. For all simulations, a sufficiently large number of primary protons have been simulated in order to ensure that the statistical uncertainty in the calculated dose was acceptably low, such that any observed differences within the study were due to the configuration of the simulations, rather than due to statistical artefacts. Therefore, all fields have been calculated by scaling the number of protons per field to 1.7–4.0 × 10^8 protons per field, providing a statistical uncertainty at the 90–100% dose level of ≤0.35% (as calculated in the GATE simulations according to Chetty et al[30]).

Differences are quantitatively analysed by examining dose difference distributions and by using the global γ analysis,[31] in which the data simulated with the PSI model is the evaluated, and the data simulated with the The Christie model the reference data set.

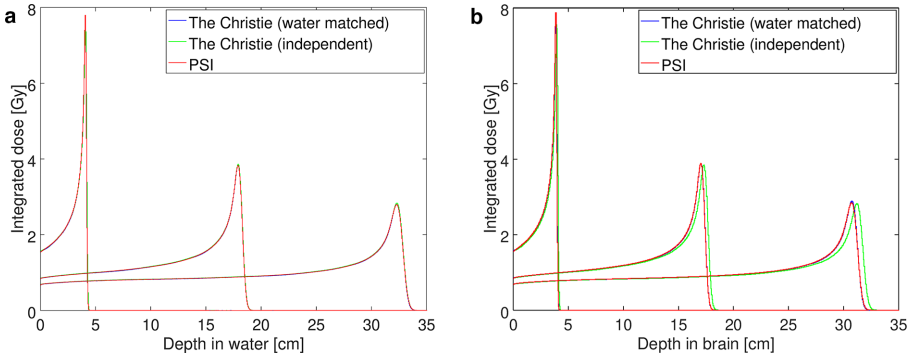
RESULTS

Comparison without pre-absorber

Comparison in water

For simulations in water without the pre-absorber (Figure 1a) ranges for both models agreed to within 0.25 mm, which is a good agreement taking into account that the beam modelling processes have been performed independently, with tolerances of 0.1 mm and independent scoring settings. In addition,

Figure 1. Depth–dose curves in water (a) and brain (b) without pre-absorber, plotted as step plots. Statistical uncertainty does not influence the shape of the curves and resulting range and absolute dose agreement.



absolute doses agreed to within 1.01% and beam sizes in air (σ of the Gaussian beam) to within 0.34 mm.

Comparison in non-water tissues (material blocks and patient CTs)

In contrast, in the simulated material blocks, ranges were found to differ by 0.7/2.1/3.6 mm (2.2/1.5/1.5% of total range) for energies of 71/163/229 MeV transported through bone and by 0.9/2.8/4.8 mm (2.2/1.6/1.5% of total range) through brain tissue (Figure 1b) with the The Christie system predicting deeper ranges than the PSI model.

When comparing clinical dose distributions in the patient CT (Figure 2), The The Christie model again predicted higher ranges in the patient and, although γ analysis showed relatively good agreement for a (2%, 2 mm) criterion (>95% agreement for all fields), a significant amount of voxels failed for a stricter (1%, 1 mm) comparison (≤75% pass rate for all fields).

Further investigation of range discrepancies

Based on these results, the substantial range discrepancies found in both material blocks and the patient CT between the two systems were further investigated. Consequently, it was

Figure 2. For a clinical proton field without pre-absorber, doses calculated with the PSI MC algorithm (a), the independently developed The Christie model (b) and the dose difference PSI – The Christie on one example slice through the high dose region (c) and as a histogram for the whole distribution (d). MC, Monte Carlo.

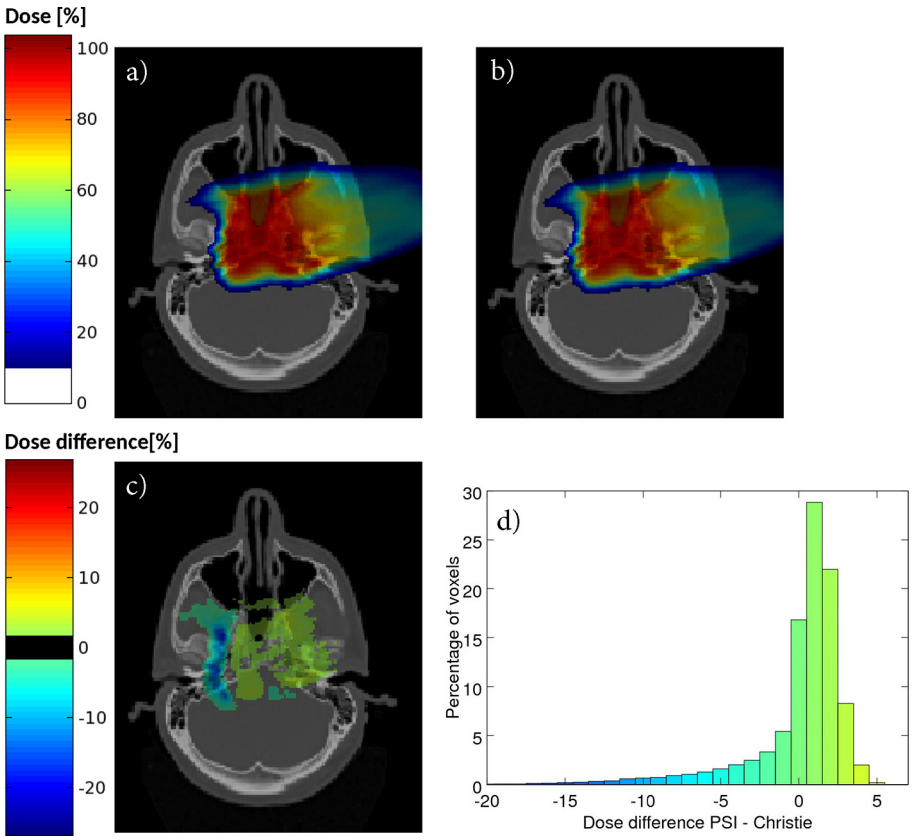
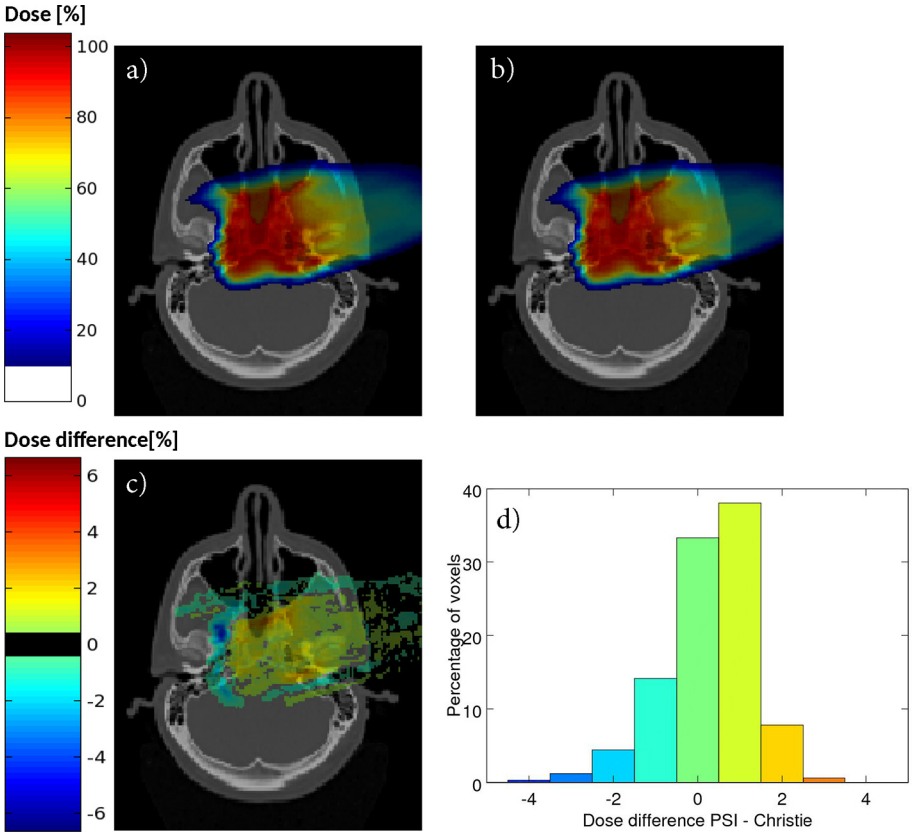


Figure 3. For a clinical proton field without pre-absorber, doses calculated with the PSI MC algorithm (a), doses calculated with the water matched The Christie MC model (b) and the dose difference PSI – Christie on one example slice through the high dose region (c) and as a histogram for the whole distribution (d). MC, Monte Carlo.



found that, during the independent beam modelling of the two systems, the definition of water for the two systems differed. Water had not been defined using the HU conversion, but, in the PSI system defined using the Geant4 default water description (fractions: H: 0.111894 & O:0.888106, density 1.0 g/cm^3, I = 78 eV), and in The Christie system using its elemental composition (H_2O, density 1.0 g/cm^3, I = 69 eV calculated internally in Geant4). Consequently, the Christie system was then re-tuned using the same water definition as the PSI. After this re-tuning, ranges between both systems agreed to within 0.1 mm for water and all material blocks (Figure 1, PSI and water matched The Christie system), beam sizes in air agreed within 0.33 mm and absolute doses in water within 0.24%. In addition, agreement substantially improved for the comparison in the patient CT (Figure 3), with the γ analysis agreement at the 1%, 1 mm level increasing to >93% for all fields, and most voxels (>97% for all fields) agreeing to within ±2.5% of prescription dose (Table 1) for doses > 10% in the patient after re-tuning of the Christie system.

Comparison with pre-absorber

Comparison in air/water with pre-absorber

With the pre-absorber (Figure 4), ranges between the water matched systems also agree to within 0.25 mm. Additionally, beam sizes are reproduced well downstream of the pre-absorber for higher energies, with a slightly higher difference of 0.68 mm in air for the lowest energy. Most importantly however, absolute doses in water simulated with the descriptive pre-absorber (The Christie system) were found to be 6.9%, 4.0%. 4.2% higher (86, 163, 229 MeV) than for the physical pre-absorber (PSI setup).

Figure 5 shows the lateral beam profiles for the 163 MeV beam in water with (a–c) and without pre-absorber (d–f) at the entrance of the water phantom (a, d), at 5 cm depth (b, e) and at 10 cm depth (c,f) on a logarithmic scale to further investigate the origin of these absolute dose differences.

Comparison in non-water tissues with pre-absorber

Ranges in bone and brain agreed to within 0.27 mm after water matching. Absolute dose differences however were found to be comparable to the water simulations (6.9/3.8/4.0% bone, 6.9/4.0/4.2% brain). Similarly, although relative doses agree well between the two pre-absorber descriptions in the patient CT (Figure 6 and Table 1), the absolute doses predicted using the descriptive pre-absorber (The Christie system) are also 7% higher than those predicted when using the physical pre-absorber (PSI model).

When, however, a physical instead of a descriptive pre-absorber is included in the Christie setup (using the same definitions as in the PSI system), relative and absolute doses agree well between the two systems with >98% γ analysis agreement at the 1%, 1 mm level without any additional scaling (Figure 7 and Table 1).

Table 1. Dose differences and γ analysis for the comparison of clinical plans between The Christie models and the PSI.model

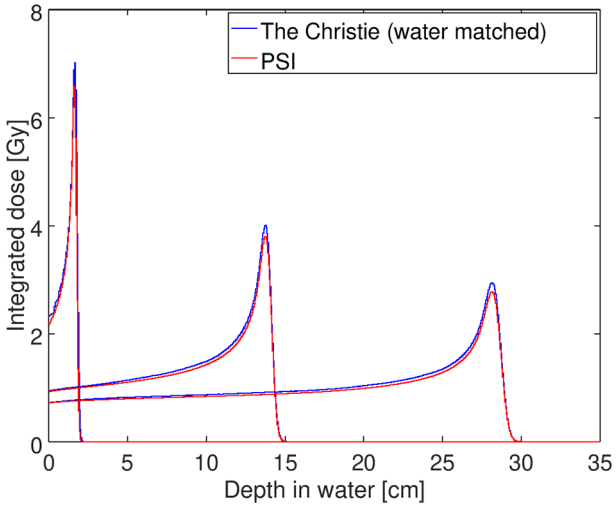
Water matched systems	**F0**	**F1**	**F2**	
Voxels within ±2.5%	97.8 %	98.8%	99.9 %	
γ analysis: 1%, 1 mm	93.8 %	95.5 %	98.6 %	
γ analysis: 2%, 2 mm	99.9 %	99.9 %	100 %	
Pre-absorber models	**F0**	**F1**	**F2**	**F3**
Scaling	0.926	0.933	0.925	0.927
Voxels within ±2.5%	90.6 %	95.2%	87.6 %	93.6 %
γ analysis: 1%, 1 mm	93.3 %	99.0 %	94.6 %	94.9%
γ analysis: 2%, 2 mm	99.5 %	100 %	99.9 %	99.6%
Physical pre-absorber in both systems	**F0**	**F1**	**F2**	**F3**
Voxels within ±2.5%	98.7 %	99.8%	98.8 %	99.0 %
γ analysis: 1%, 1 mm	99.0 %	100 %	99.8 %	98.9 %
γ analysis: 2%, 2 mm	100 %	100 %	100 %	100 %

For the pre-absorber plan, doses predicted by the Christie system have been scaled to account for an absolute dose offset (Scaling factor).

DISCUSSION

In this study, two MC systems based on the same MC code have been independently tuned and compared using identical proton PBS beam data. It is the first study to quantify the precision of such MC setups and to demonstrate potential differences resulting from the setup processes alone from the perspective of a clinical user. Even though these differences are small when compared to other uncertainties, as for example patient setup and CT conversion (see Paganetti[2] for more detail), it is important to note that there is also an uncertainty correlated to the setup process of the MC algorithm itself. Indeed, as it is impossible to measure the dose in the patient, there is no easy way of determining to what extent any MC system represents the ground truth in the patient.

Figure 4. Depth–dose curves in water with pre-absorber, plotted as step plots. Statistical uncertainty does not influence the shape of the curves and resulting range and absolute dose agreement.



During the beam modelling processes performed here, different values for the ionisation potential of water were inadvertently defined in the two systems. In the PSI system, the Geant4 default value for water of 78 eV was used, whereas in the independently tuned The Christie model, this was calculated internally using the Bragg additivity rule,[32] leading to an ionisation potential of 69 eV. Only after retuning of one system to be based on the same ionisation potential of water did both simulation algorithms agree well (1%, 1 mm γ agreement >93% for all fields).

The authors of Andreo[16] have previously demonstrated the differences of proton ranges in water for monoenergetic beams when varying the ionisation potential of water. Despite this however, and even before the re-tuning of the Christie system, ranges in water agreed well, since the energy of the pencil beams were individually tuned to reproduce measurements in water. Consequently, the differences due to the chosen ionisation potentials were not visible when comparing simulations in water and were only revealed when comparing dose distributions in other materials (material blocks or patient CT). In fact, if the authors had not done this comparison study, the two institutions may have used two MC calculation systems predicting substantially different ranges in non-water materials for the same beam data.

Unfortunately, it is not possible to say which system (or choice of ionisation potentials) is more accurate in the patient, as direct validation of calculated range *in vivo* at the resolutions required is extremely challenging. Validation measurements need therefore to be performed in anthropomorphic and

Figure 5. Lateral beam profiles plotted using a logarithmic scale for the 163 MeV beam in water with (a–c) and without pre-absorber (d–f) at the entrance of the water phantom (a, d), at 5 cm depth (b, e) and at 10 cm depth (c,f). Statistical fluctuations lead to variations of the individual low dose points, but do not influence the conclusion drawn from this figure.

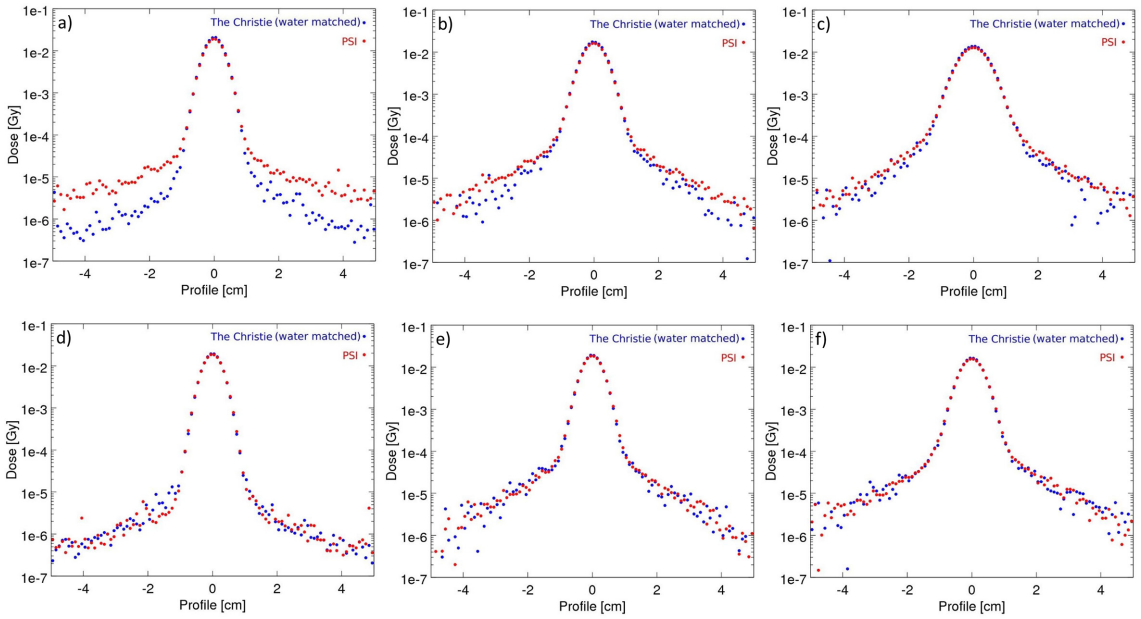


Figure 6. For a clinical proton field with pre-absorber, doses calculated with the PSI MC algorithm (physical pre-absorber, (a)), the water matched Christie MC model (descriptive pre-absorber, (b)) and the dose difference PSI – Christie on one example slice (c) and as a histogram for the whole distribution (d). The Christie doses are scaled by 7% to show the relative agreement. MC, Monte Carlo.

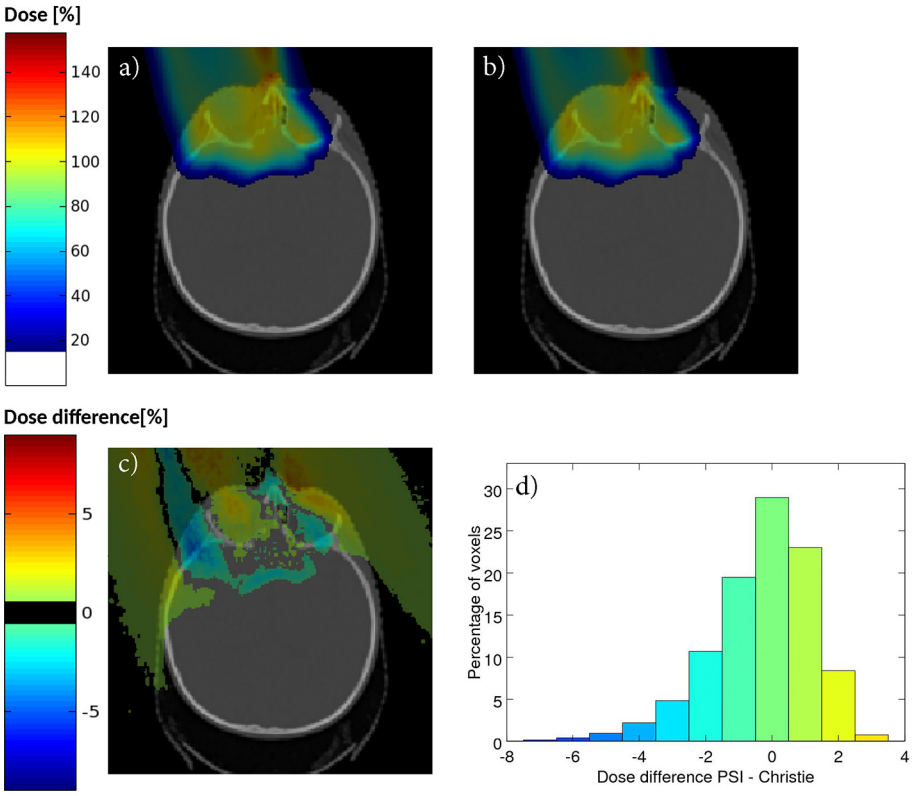
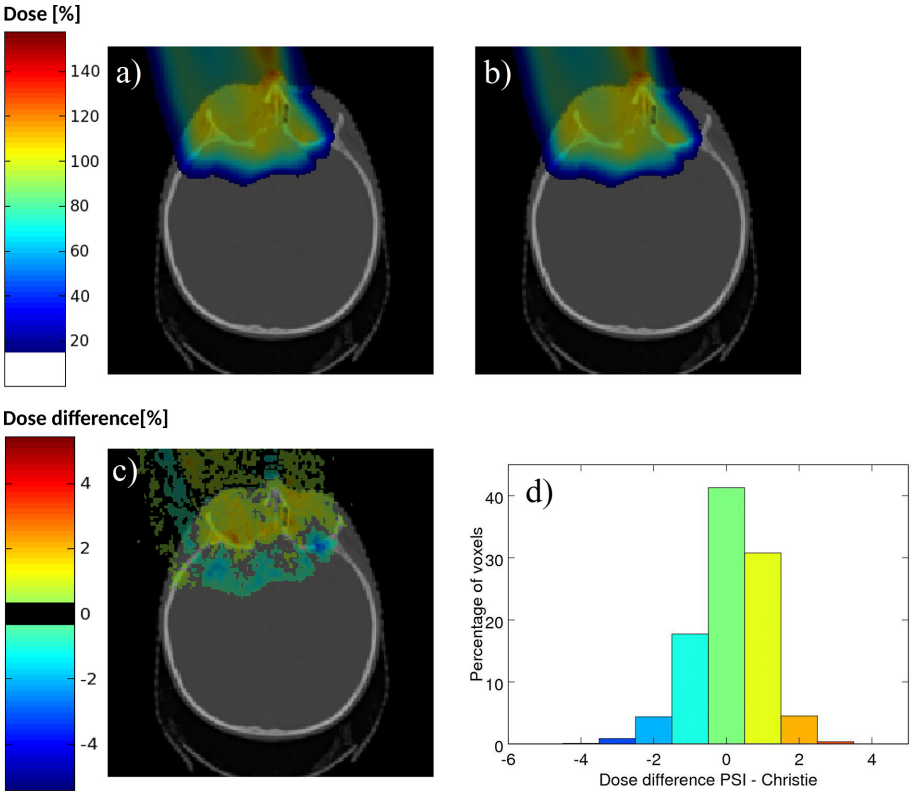


Figure 7. For a clinical proton field with pre-absorber, doses calculated with the PSI MC algorithm (a), the water matched Christie MC model (b) when both use a physical pre-absorber and the dose difference PSI – Christie on one example slice (c) and as a histogram for the whole distribution (d). In contrast to the descriptive approach, the doses are not scaled and absolute dose agreement is shown. MC, Monte Carlo.



heterogeneous phantoms. Even then, such measurements are subject to many experimental uncertainties, and often evaluated using for example 3%/3 mm γ analysis—which would not be sufficient to distinguish these kind of differences. This clearly shows that before MC calculation systems are used routinely in clinical practice, there is a need for a standard validation procedure. This might include an inter comparison as has been done in this study, and/or measurements in non-water materials.

Indeed, choosing the best ionisation potential for water during the tuning process is not necessarily straightforward. On the one hand, the current ICRU report recommends $I = 78$ eV,[33] updated from the previously recommended $I = 75$ eV.[34] On the other hand, in both GATE and TOPAS, the ionisation potentials used for materials in the patient CT are calculated from elemental ionisation potentials using the Bragg additivity rule. Therefore, it is arguably more consistent to modify/adjust the ionisation potentials for each element rather than to adjust the ionisation potential of water alone. For instance, although the ionisation potential of water was changed during the beam modelling process of the model, the CT conversion from HU to stopping power was not altered. As such, although a consistency in water was achieved between the two systems, this may lead to inconsistences for range calculations in the patient.

In practice, and based on the derivations of Schneider,[27] there are two ways to derive the scanner specific HU conversion scheme necessary for the MC simulation:

(1) Repeating the calibration process described by Schneider et al[27] for the institute-specific scanner and scanner settings.
(2) Using the original curve by Schneider et al[27] while adjusting the densities of the individual materials to match the MC calculated stopping powers relative to water to those of the institute specific scanner.[2]

The second approach aims to generate stopping powers for MC which are consistent with the treatment planning system. This has for example been chosen by Schuemann et al[35,36] to investigate differences between the analytical and the MC simulations. It is however important to note that an ionisation potential of water has to be assumed to derive the stopping powers relative to water of the institute specific scanner.[28] To be self-consistent, *i.e.* achieve the same ranges in the patient CT as the treatment planning system, the same I value should be used for the tuning process of the MC simulation.

We also observed major differences in results between the two MC systems as a result of modelling of the pre-absorber. In the literature, pre-absorbers in MC simulations have been modelled either as physical components[17,19,23] or their effects included in the initial beam description.[18] When performed

carefully, both are valid approaches, with the descriptive approach having the advantage of efficiency for particle tracking. However, in our study, when including the pre-absorber in the beam optics and using the same number of protons, integrated doses (Figure 4) and point to point doses in the core (Table 1) were higher when compared to the physical modelling.

There are two potential reasons for the observed absolute dose differences, namely the loss of fluence in the pre-absorber and wide angle scattered protons/secondaries originating in the pre-absorber. When analysing beam profiles on a logarithmic scale (Figure 5), the difference in the contribution of these secondaries on the beam halo (see terminology introduced in Gottschalk et al[37]) is clearly visible at low depths (Figure 5a), an effect not observed without pre-absorber (Figure 5d–f). At deeper depths, this difference in the low dose halo, however, vanishes (Figure 5b and c), as here the secondaries produced in water are more prevalent. The results of this study, however, show a systematic absolute dose offset over all depths (Figure 4), and the authors therefore conclude that the difference in absolute dose is mainly due to the loss of fluence in the pre-absorber, and is only partly due to differences in the modelling of wide angle scattered protons originating in the pre-absorber, both of which are not taken into account correctly in such a descriptive approach. Nevertheless, it is interesting to point out the differences in the low dose halo between the descriptive and the physical approaches also visible in Figure 6, which are not observed when using the same pre-absorber approach (Figure 7). This is consistent with the work by Grassberger et al,[19] who show that not tracking the particles through the treatment head results in an underestimation of the dose halo. Investigating the effect of the different pre-absorber descriptions on secondaries would be an interesting continuation of this work.

When subsequently including the physical pre-absorber in The Christie system, absolute doses agreed well, which validates that the observed differences are indeed due to the pre-absorber modelling approaches and not due to potential underlying code differences. This emphasised the importance of absolute dose measurements with a pre-absorber when validating the MC setup, which have shown that absolute doses with a physical pre-absorber reproduce measurements with only a 1% offset.[17] When the pre-absorber cannot be modelled as a physical component however, *e.g.* due to overlap problems with the CT or to reduce simulation times and a descriptive model is preferred, it is therefore important to take this proton loss into account. The descriptive model has, *e.g.* been successfully used by Fracchiolla et al[18] to predict absolute doses. However, this approach succeeded due to additional measurements that were performed with the pre-absorber to introduce an energy dependent scaling factor to modify the fluence of each proton pencil beam.

In summary, when following the beam model process as described in "Setting up a PBS MC system", ionisation potentials and the modelling of objects have been shown to be critical parameters. This study therefore proposes an extended beam modelling procedure (*new/altered steps marked in italic*):

(1) *Choose the water definition, especially the ionisation potential. This should be the same as used for the creation of the stopping power curve in the TPS.*
(2) Without pre-absorber:
    - Characterise the beam optics (initial phase space: beam size, angular spread and correlation) of pencil beams in air (for each nozzle extension if applicable).
    - Tune energy and energy spread to reproduce measured depth dose curves in water.
(3) With pre-absorber, physical description
    - Tune material properties (*e.g.* density) to match depth–dose cuves in water

    OR With pre-absorber, optical description.
    - Characterise initial phase space of pencil beams in air (for each nozzle extension).
    - Tune energy and energy spread *and proton number* to reproduce measured depth–dose curves in water.
(4) *Perform validation procedures, e.g. the intercomparison between MC algorithms and/or measurements in non-water materials.*
(5) *For the CT conversion, use the original curve by Schneider et al[27] while adjusting the densities of the individual materials to match the MC calculated stopping powers relative to water to those of the institute-specific scanner.*

Conclusions

This study highlights some potential uncertainties that can occur purely from the process of beam modelling for MC systems for proton PBS therapy. As such, new steps in the beam modelling process are proposed, to potentially minimise the influence of ionisation potentials and the modelling of objects, which have been shown to be critical parameters. Standard validation procedures, including an intercomparison between different MC algorithms and/or MC-calculated and measured dose distributions in non-water materials, should be considered when introducing MC calculations in clinical practice.

ACKNOWLEDGEMENTS

We are thankful for the computer support of Gilles Martin and the Medical IT group and the calculation time provided by the Merlin4 cluster. This work forms part of Carla Winterhalter's PhD thesis (Protons Do Play Dice: Validating, Implementing and Applying Monte Carlo Techniques for Proton Therapy, Diss. ETH No. 25698, https://doi.org/10.3929/ethz-b-000314036).

FUNDING

This work was supported by a research grant from Varian Medical Systems - Particle Therapy, Germany. First work has been done within the scope of the ESTRO mobility grant (Technology Transfer Grant TTG).

NOTES

The TOPAS website states: "Our mission is to make TOPAS into a tool that fits comfortably into the hands of every medical physicist." (url: http://www.topasmc.org, 10th July 2019).

REFERENCES

1. Lomax A. Intensity modulation methods for proton radiotherapy. *Phys Med Biol* 1999; **44**: 185–205. doi: https://doi.org/10.1088/0031-9155/44/1/014
2. Paganetti H. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Phys Med Biol* 2012; **57**: R99–117. doi: https://doi.org/10.1088/0031-9155/57/11/R99
3. Yepes P, Adair A, Grosshans D, Mirkovic D, Poenisch F, Titt U, et al. Comparison of Monte Carlo and analytical dose computations for intensity modulated proton therapy. *Phys. Med. Biol.* 2018; **63**: 045003. doi: https://doi.org/10.1088/1361-6560/aaa845
4. Tommasino F, Fellin F, Lorentini S, Farace P. Impact of dose engine algorithm in pencil beam scanning proton therapy for breast cancer. *Physica Medica* 2018; **50**: 7–12. doi: https://doi.org/10.1016/j.ejmp.2018.05.018
5. Widesott L, Lorentini S, Fracchiolla F, Farace P, Schwarz M. Improvements in pencil beam scanning proton therapy dose calculation accuracy in brain tumor cases with a commercial Monte Carlo algorithm. *Phys. Med. Biol.* 2018; **63**: 145016. doi: https://doi.org/10.1088/1361-6560/aac279
6. Agostinelli S, Allison J, Amako K, Apostolakis J, Araujo H, Arce P, et al. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 2003; **506**: 250–303. doi: https://doi.org/10.1016/S0168-9002(03)01368-8
7. Jan S, Santin G, Strul D, Staelens S, Assié K, Autret D, et al. Gate: a simulation toolkit for PET and SPECT. *Phys Med Biol* 2004; **49**: 4543–61. doi: https://doi.org/10.1088/0031-9155/49/19/007
8. Perl J, Shin J, Schümann J, Faddegon B, Paganetti H. Topas: an innovative proton Monte Carlo platform for research and clinical applications. *Med Phys* 2012; **39**: 6818–37. doi: https://doi.org/10.1118/1.4758060
9. Arce P, Ignacio Lagares J, Harkness L, Pérez-Astudillo D, Cañadas M, Rato P, et al. Gamos: a framework to do Geant4 simulations in different physics fields with an user-friendly interface. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 2014; **735**: 304–13. doi: https://doi.org/10.1016/j.nima.2013.09.036
10. Iwase H, Niita K, Nakamura T. Development of general-purpose particle and heavy ion transport Monte Carlo code. *J Nucl Sci Technol* 2002; **39**: 1142–51. doi: https://doi.org/10.1080/18811248.2002.9715305
11. Briesmeister JF. MCNPTM-A general Monte Carlo N-particle transport code. Version 4C, LA-13709-M, Los Alamos National Laboratory 2; 2000
12. Battistoni G. "The FLUKA code: Description and benchmarking". *AIP Conference proceedings.* **Vol. 896. No. 1**; 2007.
13. Titt U, Bednarz B, Paganetti H. Comparison of MCNPX and Geant4 proton energy deposition predictions for clinical use. *Phys Med Biol* 2012; **57**: 6381–93. doi: https://doi.org/10.1088/0031-9155/57/20/6381
14. Grevillot L, Frisson T, Zahra N, Bertrand D, Stichelbaut F, Freud N, et al. Optimization of GEANT4 settings for proton pencil beam scanning simulations using gate. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 2010; **268**: 3295–305. doi: https://doi.org/10.1016/j.nimb.2010.07.011
15. Zacharatou Jarlskog C, Paganetti H. Physics settings for using the Geant4 toolkit in proton therapy. *IEEE Trans Nucl Sci* 2008; **55**: 1018–25. doi: https://doi.org/10.1109/TNS.2008.922816
16. Andreo P. On the clinical spatial resolution achievable with protons and heavier charged particle radiotherapy beams. *Phys Med Biol* 2009; **54**: N205–15. doi: https://doi.org/10.1088/0031-9155/54/11/N01
17. Winterhalter C, Fura E, Tian Y, Aitkenhead A, Bolsi A, Dieterle M, et al. Validating a Monte Carlo approach to absolute dose quality assurance for proton pencil beam scanning. *Phys. Med. Biol.* 2018; **63**: 175001. doi: https://doi.org/10.1088/1361-6560/aad3ae
18. Fracchiolla, Lorentini S, Widesott L, Schwarz M. Characterization and validation of a Monte Carlo code for independent dose calculation in proton therapy treatments with pencil beam scanning. *Phys Med Biol* 2015; **60**: 8601–19. doi: https://doi.org/10.1088/0031-9155/60/21/8601
19. Grassberger C, Lomax A, Paganetti H. Characterizing a proton beam scanning system for Monte Carlo dose calculation in patients. *Phys Med Biol* 2015; **60**: 633–45. doi: https://doi.org/10.1088/0031-9155/60/2/633
20. Grevillot L, Bertrand D, Dessy F, Freud N, Sarrut D. A Monte Carlo pencil beam scanning model for proton treatment plan simulation using GATE/GEANT4. *Phys Med Biol* 2011; **56**: 5203–19. doi: https://doi.org/10.1088/0031-9155/56/16/008
21. Verburg JM, Grassberger C, Dowdell S, Schuemann J, Seco J, Paganetti H. Automated Monte Carlo simulation of proton therapy treatment plans. *Technol Cancer Res Treat* 2016; **15**: NP35–46. doi: https://doi.org/10.1177/1533034615614139
22. Sawakuchi GO, Titt U, Mirkovic D, Ciangaru G, Zhu XR, Sahoo N, et al. Monte Carlo investigation of the low-dose envelope from scanned proton pencil beams. *Phys Med Biol* 2010; **55**: 711–21. doi: https://doi.org/10.1088/0031-9155/55/3/011
23. Saini J, Maes D, Egan A, Bowen SR, St James S, Janson M, et al. Dosimetric evaluation of a commercial proton spot scanning Monte-Carlo dose algorithm: comparisons against measurements and simulations. *Phys. Med. Biol.* 2017; **62**: 7659–81. doi: https://doi.org/10.1088/1361-6560/aa82a5
24. Pedroni E, Meer D, Bula C, Safai S, Zenklusen S. Pencil beam characteristics of the next-generation proton scanning gantry of psi: design issues and initial commissioning results. *The European Physical Journal Plus* 2011; **126**: 66. doi: https://doi.org/10.1140/epjp/i2011-11066-0
25. Wright D, Incerti S. A short guide to choosing physics lists. 2018. Available from: http://geant4.in2p3.fr/IMG/pdf_PhysicsLists.pdf [12.11.208].
26. Reference Physics Lists. GEANT4: A Simulation Toolkit. 2018. Available from: https://geant4.web.cern.ch/node/155 [12.11.2018].

27. Schneider W, Bortfeld T, Schlegel W. Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions. *Phys Med Biol* 2000; **45**: 459–78. doi: https://doi.org/10.1088/0031-9155/45/2/314
28. Schneider U, Pedroni E, Lomax A. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Phys Med Biol* 1996; **41**: 111–24. doi: https://doi.org/10.1088/0031-9155/41/1/009
29. ICRU_46. photon, electron, proton and neutron interaction data for body tissues.*ICRU report* 1992; **46**.
30. Chetty IJ, Rosu M, Kessler ML, Fraass BA, Ten Haken RK, Kong F-MS, Kong Feng-Ming (Spring), et al. Reporting and analyzing statistical uncertainties in Monte Carlo-based treatment planning. *Int J Radiat Oncol Biol Phys* 2006; **65**: 1249–59. doi: https://doi.org/10.1016/j.ijrobp.2006.03.039
31. Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys* 1998; **25**: 656–61. doi: https://doi.org/10.1118/1.598248
32. Wilson JW, Xu Y, Kamaratos E, Chang C. Mean excitation energies for stopping powers in various materials using local plasma oscillator strengths.. *NASA Technical Paper* 1984; **2271**.
33. ICRU_90. key data for Ionizing-Radiation dosimetry: measurement standards and applications.*ICRU report* 2016; **90**.
34. ICRU_37. stopping powers for electrons and positrons.*ICRU report* 1984; **37**.
35. Schuemann J, Dowdell S, Grassberger C, Min CH, Paganetti H. Site-Specific range uncertainties caused by dose calculation algorithms for proton therapy. *Phys Med Biol* 2014; **59**: 4007–31. doi: https://doi.org/10.1088/0031-9155/59/15/4007
36. Schuemann J, Giantsoudi D, Grassberger C, Moteabbed M, Min CH, Paganetti H. Assessing the clinical impact of approximations in analytical dose calculations for proton therapy. *Int J Radiat Oncol Biol Phys* 2015; **92**: 1157–64. doi: https://doi.org/10.1016/j.ijrobp.2015.04.006
37. Gottschalk B, Cascio EW, Daartz J, Wagner MS. On the nuclear halo of a proton pencil beam stopping in water. *Phys Med Biol* 2015; **60**: 5627–54. doi: https://doi.org/10.1088/0031-9155/60/14/5627