The influence of Monte Carlo radiation transport codes and hadron physics models on stray neutron dose estimations in proton therapy: An extended intercomparison

Peer-reviewed author version

COLSON, Dries; Pinasti, Sita; Matamoros Ortega, Andrea Nicole; De Saint-Hubert, Marijke; Van Hoey, Olivier & RENIERS, Brigitte (2024) The influence of Monte Carlo radiation transport codes and hadron physics models on stray neutron dose estimations in proton therapy: An extended intercomparison. In: RADIATION MEASUREMENTS, 170 (Art N° 107046).

DOI: 10.1016/j.radmeas.2023.107046
Handle: http://hdl.handle.net/1942/42032

The influence of Monte Carlo radiation transport codes and hadron physics models on stray neutron dose estimations in proton therapy: an extended intercomparison

Dries Colson[1], Sita Pinasti[2,3], Andrea Matamoros Ortega[2,4], Marijke De Saint-Hubert[2], Olivier Van Hoey[2], Brigitte Reniers[1]

[1] Hasselt University, Faculty of Engineering Technology-Nuclear Technology (NuTeC), Hasselt, Belgium

[2] Belgian Nuclear Research Centre (SCK CEN), Research in Dosimetric Applications, Mol, Belgium

[3] Université libre de Bruxelles, Service de Métrologie Nucléaire, Brussels School of Engineering, Brussels, Belgium

[4] KU Leuven, Department of Oncology, Laboratory of Experimental Radiotherapy, Leuven, Belgium

*Corresponding author: dries.colson@uhasselt.be

Abstract

Purpose: Selecting a Monte Carlo (MC) radiation transport code and nuclear physics model affects the simulation of secondary neutron fluence energy spectra and the estimation of the neutron dose equivalent in proton therapy. This study compares the MC codes FLUKA, MCNPX and GEANT4 from previous work with two additional MC codes, TOPAS (with either the Binary Intra-nuclear Cascade (BIC) or the Bertini intra-nuclear cascade model (Bert)) and PHITS (with either the Intra-Nuclear Cascade of Liege (INCL) or the Bert model), to facilitate a more comprehensive comparison.

Material and methods: Secondary neutron fluence energy spectra were scored at different locations inside a water phantom, exposed to a 10x10 cm^2 parallel mono-energetic proton beam with energies of 110 MeV, 150 MeV, 180 MeV and 210 MeV. The neutron dose equivalent was estimated by applying fluence-to-dose equivalent conversion coefficients.

Results: Discrepancies between MC codes and physics models reach 90 % for the total neutron fluence. Differences are most pronounced in the forward direction of the beam due to a notable presence of high-energy neutrons and the higher uncertainty on the neutron cross-section data for these neutrons. Nuclear model selection within PHITS (INCL or Bert) and TOPAS (BIC or Bert) shows a worse agreement for the Bert model than for the other MC codes and nuclear models. The discrepancies in the neutron spectrum simulations propagate into differences in the estimated neutron dose equivalents. They are more prominent for distal positions (up to a factor of 3) than for lateral positions (up to a factor of 2) and decrease at higher proton energies.

Conclusions: This study shows considerable differences in the out-of-field neutron fluence energy spectra calculated by different MC codes and nuclear physics models. Differences in the neutron dose equivalent reach a factor of 3. The discrepancies are most apparent in distal positions and at lower proton energies.

Keywords

Proton therapy
Monte Carlo
neutron dose equivalent
TOPAS
PHITS

Introduction

The use of proton therapy in comparison with conventional radiotherapy offers benefits in terms of higher tumour coverage and better sparing of normal tissue and organs at risk. This is caused by the specific ballistic properties of protons, characterized by the Bragg peak (Mohan, 2022). The production of secondary neutrons by nuclear interactions of the high-energy protons with the beamline components on the one hand and the patient's body on the other is challenging, as it contributes to the out-of-field (OOF) dose (Gottschalk, 2006), (Schneider & Hälg, The impact of neutrons in clinical proton therapy, 2015), (De Saint-Hubert, et al., 2022). This OOF dose is particularly important as it impacts the risk of late effects on normal tissue, such as secondary cancers or cataract. This is especially important for treating young children or the in utero exposure in the case of the radiotherapy treatment of a pregnant woman (Knežević, et al., 2022), (De Smet, et al., 2017), (Lee, et al., 2017), (Englbrecht, et al., 2021), (De Saint-Hubert, et al., 2022).

Monte Carlo (MC) radiation transport codes are often used to simulate the OOF neutron dose, as the neutron dose is not considered in current proton treatment planning systems. MC codes are also applied for designing the shielding of proton therapy facilities. MC simulations are sometimes the only option, as measuring the OOF neutron dose directly is challenging. De Saint-Hubert et al. (De Saint-Hubert, et al., 2022) highlighted in 2022 that concerning OOF neutron dose in proton therapy, not only the number of studies comparing MC simulations with measurements is very limited, but also that significant discrepancies between MC simulations and experimental results (up to factors 2-3) are observed. These discrepancies are mainly due to large measurement uncertainties on the one hand and limitations of nuclear reaction models and cross sections in MC codes on the other hand. The issue does not apply for neutrons with energies below 20 MeV; here MC codes can perform an accurate calculation as reliable neutron cross section data are available. The problem, however, is that for neutrons with energies higher than 20 MeV, cross section data are scarce or non-existent for certain materials and MC codes. In that case, MC codes use nuclear models that describe the interaction of protons and neutrons with target nuclei (De Saint-Hubert, et al., 2022), (Englbrecht, et al., 2021), (Schneider, Agosteo, Pdroni, & Bessere, 2002), (Farah, et al., 2014).

Earlier work performed by De Saint-Hubert et al. in 2022 (De Saint-Hubert, et al., 2022) investigated the influence of three different multi-purpose MC radiation transport codes on the simulation of secondary neutron fluence energy spectra and its impact on calculating the neutron dose in proton therapy. In this study, FLUKA (with HADROTHErapy default settings), GEANT4 (with the binary intra-nuclear cascade model, BIC) and MCNPX were included. Within the MCNPX simulations, the influence of three different nuclear models was additionally investigated, i.e. the Bertini intra-nuclear cascade model in combination with the Dresner evaporation-fission model (Bert-Dres), Isabel for intra-nuclear cascade modelling together with Dresner and Abla models for the evaporation phase, and the Cascade-Exciton Model, providing a combination of essential features of the exciton and intra-nuclear cascade model) (De Saint-Hubert, et al., 2022).

This study aims to extend this earlier work by adding the MC codes TOPAS Tool for Particle Simulation (TOPAS), with either the Binary Intra-nuclear Cascade (BIC) or the Bertini intra-nuclear cascade (Bert) model, and Particle and Heavy Ion Transport code System (PHITS), with either the Intra-Nuclear Cascade of Liege (INCL) or the Bert nuclear model to the comparison. To allow comparison with previous data, the simulation setup of (De Saint-Hubert, et al., 2022) was reproduced. First, the ranges of monoenergetic proton beams (field size of 10 x 10 cm^2 and energies of 110 MeV, 150 MeV, 180 MeV and 210 MeV) in a water phantom with polymethyl methacrylate (PMMA) walls were calculated. Then, the secondary neutron fluence energy spectra were calculated at various locations inside the water phantom, both distal and lateral, with respect to the proton beam's Bragg peak. Fluence-to-dose-equivalent conversion factors were then used to estimate the neutron dose equivalent. An extended comparison was made between FLUKA (HADROTHE), MCNPX (Bert-Dres), GEANT4 (BIC), TOPAS (BIC or Bert) and PHITS (INCL or Bert).

Materials and methods

1. Simulation setup

As this study is an extension to the earlier work (De Saint-Hubert, et al., 2022) the same simulation setup is used. This means a 10x10 cm^2 parallel beam of monoenergetic protons with energies of 110 MeV, 150 MeV, 180 MeV or 210 MeV, travelling through 50 cm of air before reaching the water phantom's surface. The water phantom is developed by Bordy et al. (Bordy, et al., 2013). It measures 30x60x30 cm^3 and has a 15 mm thick PMMA wall. The proton beam hits the centre of the phantom's 12x12 cm^2 entrance window where the wall consists of only 4 mm PMMA.

The simulation setup is schematically represented in Figure 1. In this figure, numbers 1 to 12 indicate the different scoring locations. Material composition and density for air, PMMA and water (summarised in Table 1) were adjusted in the different MC codes following (De Saint-Hubert, et al., 2022). It may be noted that for the elemental composition, the main isotope is assumed to have an abundance of 100 %. This is expected to have only a minimal effect because the other isotopes have very low abundances for the materials in this study.

Table 1: densities and compositions used for the MC simulations.

Material	**Density (g/cm^3)**	**Isotope**	**Mass fraction**
Air (dry)	0.001205	N-14	0.75527
		O-16	0.23178
		Ar-40	0.01282
		C-12	0.00012
Water	0.998	0-16	0.888106
		H-1	0.111894
PMMA	1.18	C-12	0.59985
		0-16	0.31962
		H-1	0.08054

Firstly, the depth dose curve was examined. The absorbed dose to water was scored in scoring volumes with an area of 2x2 cm^2 and 1 mm thickness along the beam's central axis. The distance is calculated from the outside of the phantom, i.e., with the 4 mm PMMA wall of the entrance window included. The range is the linearly interpolated distance corresponding to the 90 % dose in the distal falloff (R90).

Secondly, the neutron fluence energy spectrum was scored in a volume of 1 cm diameter at 8 of the 12 locations indicated by a black-coloured marker in Figure 1 (a). The distances of these scoring locations relative to the Bragg peak can be found in Figure 1 (b), (c), (d) and (e). These figures depict the different monoenergetic proton beams with energies of 110 MeV, 150 MeV, 180 MeV and 210 MeV, respectively, entering the water phantom and the corresponding R90, indicated with a red star. The neutron fluence was scored in 114 logarithmically spaced energy bins from 0.001 eV to 200 MeV. The energy bins and associated energy range were selected in accordance with the fluence-to-dose conversion factors mentioned below to calculate the neutron dose equivalent. Two additional bins were added for neutrons with an energy lower than 0.001 eV and higher than 200 MeV, respectively, to check that all scored neutrons are included in the defined energy range. This was the case for the simulations performed. The neutron fluence data were analysed for four neutron energy regions, i.e., thermal ($E \leq 0.4$ eV), epithermal (0.4 eV $< E \leq$ 100 keV), fast (100 keV $< E \leq$ 19.6 MeV) and high energy ($E >$ 19.6 MeV) neutrons. The segmentation was made by choosing the nearest energy bin from the simulation.

The neutron dose equivalent was then calculated based on fluence-to-dose conversion factors. This is done by multiplying the scored neutron fluence for each energy bin by the corresponding kerma factor $k(E)$ and quality factor $Q(E)$, both as a function of the neutron energy (Romero-Expósito, Domingo, Sánchez-Doblado, Ortega-Gelabert, & Gallego, 2016), (Chadwick, et al., 1999), (Siebert & Schuhmacher, 1995). Some of the MC codes allow this neutron dose equivalent to be calculated instantly. To ensure that the same method of calculation and associated fluence-to-dose conversion factors were used across the different MC codes, this calculation was performed manually, outside of the MC code, starting from the scored neutron fluence energy spectra.

The number of histories was adapted to obtain a statistical uncertainty (1) for the simulations of the depth-dose curves of lower than 1 % on the absorbed dose to water in each bin in depth from the surface to the corresponding R10 (the depth corresponding to the 10% dose in the distal falloff) for the different proton energies and (2) for the simulations of the neutron fluence energy spectra of lower than 3 % for the energy bins that significantly contribute to the neutron spectrum, i.e. the energy bins with a neutron fluence within 95% of the maximum neutron fluence from the simulation (coverage factor $k = 1$). Error bars were not added to the graphs presented in this manuscript because they overlap with the markers and thereby make the visualisation and further comparison difficult.

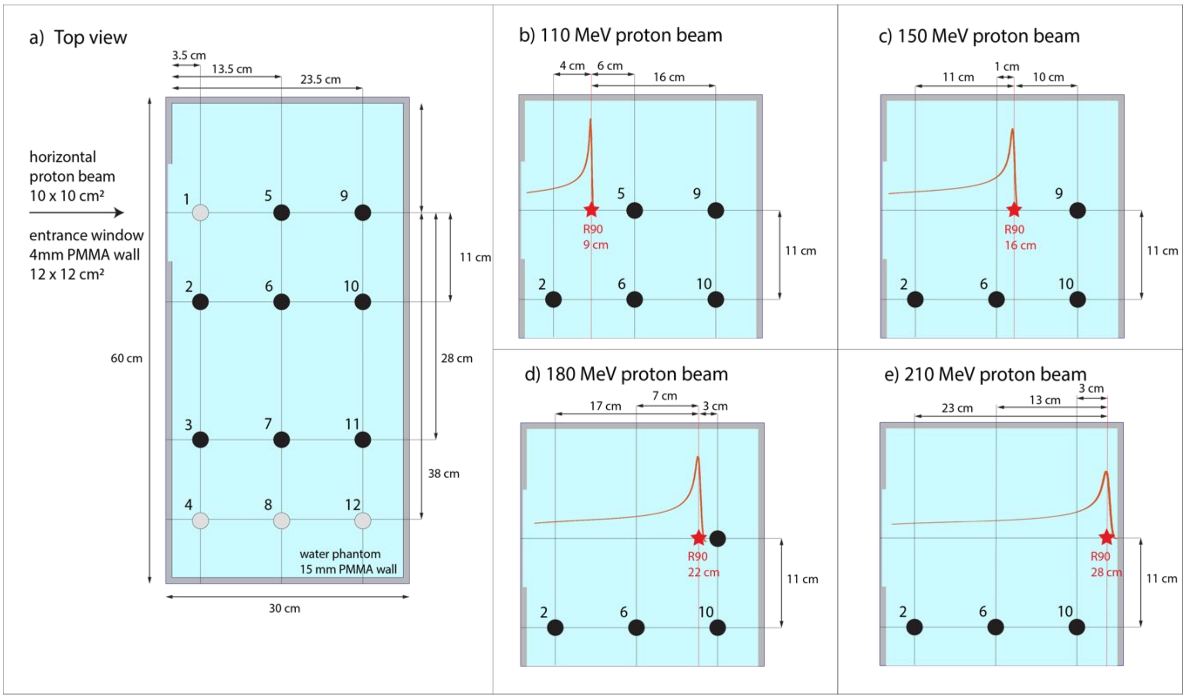

Figure 1: (a) Representation of the 30x60x30 cm^3 water phantom with 15 mm thick PMMA walls and 12x12 cm^2 beam entrance window with 4 mm PMMA wall. Numbers 1 to 12 indicate the locations of the spherical scoring volumes of 1 cm diameter. The neutron fluence energy spectra are scored at locations marked in solid black.

Figure 1 (b), (c), (d) and (e) illustrate the distance of the scoring locations with respect to the Bragg peak for proton energies 110 MeV, 150 MeV, 180 MeV and 210 MeV, respectively. R90 is indicated with a red star.

2. Monte Carlo codes

The simulations in this study were performed with two different MC radiation transport codes, i.e., TOPAS and PHITS. Each simulation was performed with one billion source particles. No variance reduction techniques were applied.

a. TOPAS

TOPAS, which stands for 'Tool for Particle Simulation', is based on the Geant4 Simulation Toolkit, a widely used package for simulating the passage of particles through matter (Perl, Shin, Schumann, Faddegon, & Paganetti, 2012) - (Faddegon, et al., 2020). It offers the power of GEANT4, but with a user-friendly interface, by applying scripting language and providing a graphical user interface (GUI). This research used version 3.9 (release date December 16, 2022), which relies on GEANT4 version 10.7 patch 03 (November 19, 2021). It should be noted that De Saint-Hubert et al (De Saint-Hubert, et al., 2022) used GEANT4 version 10.1.2 (release date June 10, 2016). The TOPAS results from the current study were thus generated by using more recent GEANT4 data.

In GEANT4, the code 'Physics' specifies the particles and physics processes, including the corresponding cut-offs used for the simulation. The default physics library used in TOPAS has been proven to adequately perform for proton therapy research (Zacharatou & Paganetti, 2008), (TOPAS MC Inc., 2022), (Lansonneur, et al., 2020). This modular list considers both protons and secondary particles of interest (e.g. electrons, neutrons, photons, alpha particles, deuterons, tritons), consisting of the following seven modules with Geant4 Class Names:

- G4EMStandarPhysics_option4,
- HadronPhysicsQGSP_BIC_HP (default) or HadronPhysicsQGSP_BERT_HP (see below)
- G4DecayPhysics,
- G4IonBinaryCascadePhysics,
- G4HadronElasticPhysicsHP,
- G4StoppingPhysics, (Zacharatou & Paganetti, 2008), (TOPAS MC Inc., 2022)
- G4RadioactiveDecay (Lansonneur, et al., 2020).

No changes were made to the default selection of the different modules for the simulations performed in this work due to its proven accuracy for research in proton therapy (Zacharatou & Paganetti, 2008) within an acceptable computing time, except for the hadron physics model. Each simulation was performed twice with either the BIC or Bert physics model enabled. This was done to assess the influence of the physics model on nuclear interactions that determine the OOF dose. These two hadron physics models, BIC or Bert, were suggested by multiple references to be used for medical applications and more specifically for proton therapy (Zacharatou & Paganetti, 2008), (TOPAS MC Inc., 2022), (Lansonneur, et al., 2020), (Wright & Incerti, s.d.), (Novak, 2019), (Arce, et al., 2020), (Arce, et al., 2021).

Two users used this code to study the inter-user variability.

b. PHITS

PHITS (Particle and Heavy Ion Transport code System) is a general MC package developed by the Japan Atomic Energy Agency that allows for the simulation of various types of particles (i.e. neutrons, protons, photons, electrons and heavy ions) through matter and can be used for a wide range of applications including radiation shielding and medical physics (Sato, et al., 2018). PHITS version 3.31 (release date April 19, 2023) was employed for this study. For all simulations, the neutron cross-sections for energies up to 20 MeV were taken from JENDL-4.0 (Japanese Evaluated Nuclear Data Library version 4.0) nuclear dataset. Meanwhile, the INCL and Bert physics models were used for neutron cross-sections with energy above 20 MeV and all proton cross-sections. As the parameter for the physics model, we selected the GEM version 1 for the de-excitation model, turned on the energy and angle straggling options by setting *nedisp* to 1 and *spread* to 2, and used the event generator mode version 2. The PHITS developer recommended these parameter settings to correctly simulate and record the quantities retrieved in the case of proton therapy (Japan Atomic Energy Agency, 2007). The INCL model was employed by setting *inclg* to 2, whereas the Bert model was turned on by setting *ejamnu* and *eqmdnu* to 1000.

Results

1. Range verification

First, the ranges of the mono-energetic proton beams were compared. Table 2 reports on the linearly interpolated R90 values obtained with the different MC codes and physics models. The results from this study (highlighted in grey) were compared to the results of (De Saint-Hubert, et al., 2022). The results of both TOPAS users agree for both physics models within 0.1 mm (detailed results were added in the Supplementary material); hence an average value is represented for this MC code. It can be noted that the values for R90 from this study agree with the previously published results within 1.5 mm. The difference in R90 values between the different MC codes and physics models is smallest at the lowest proton energy and increases with increasing energies.

Table 2: Linearly interpolated R90 values (mm) calculated with different MC codes and physics models for the monoenergetic proton beams
(results of this study are highlighted in grey).

		Proton energy			
MC code	**Physics model**	110 MeV	150 MeV	180 MeV	210 MeV
GEANT4	**BIC**	89.8	156.2	215.8	280.6
MCNPX	**Bert-Dres**	90.1	156.2	215.1	280.7
FLUKA	**HADROTHErapy**	89.8	156.2	214.7	280.2
TOPAS	**BIC**	90.1	156.7	215.5	281.2
	Bert	90.1	156.7	215.5	281.2
PHITS	**INCL**	89.7	155.8	214.4	279.8
	Bert	89.7	155.8	214.4	279.8

Figure 2 shows the absolute absorbed dose to water per source particle as a function of the depth in the water phantom for 110 MeV (left) and 210 MeV (right) proton beams. Although only two proton energies are shown here, the findings described below are similar for the two intermediate proton energies. The depth dose curves of these two intermediate energies can be found in the Supplementary material. A fair agreement in the plateau region between the different simulations is observed for the 110 MeV and 210 MeV proton beams. Although a close match of R90 values was previously noted, differences between the depth dose curves appear in the dose build-up region, the Bragg peak and the distal dose falloff region. The largest discrepancies arise in the dose build-up region. PHITS (INCL and Bert) data exhibit a slightly faster build-up and show a systematically higher absorbed dose to water in the Bragg peak compared to TOPAS (BIC and Bert) data. The simulations with MCNPX (Bert-Dres), on the other hand, show a lower absorbed dose to water per source particle in the Bragg peak compared to TOPAS (BIC and Bert).

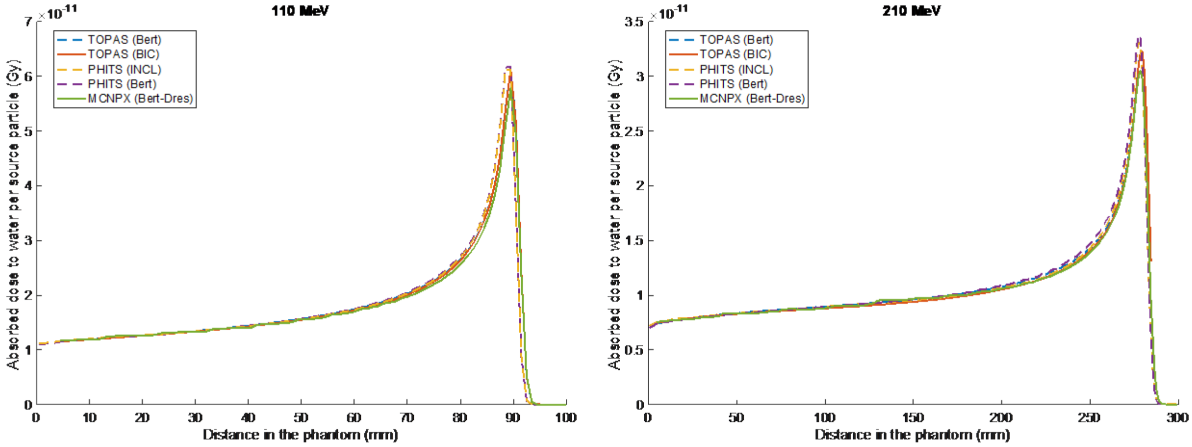

Figure 2: Depth dose curves for 110 MeV (left) and 210 MeV (right) proton beams simulated with TOPAS (Bert or BIC) and PHITS (INCL or Bert) and MCNPX (Bert-Dres).

2. Neutron fluence energy spectra

a. Comparison of TOPAS (BIC) simulations for 110 MeV and 210 MeV proton beam

The neutron fluence energy spectra for the 110 MeV and 210 MeV proton beams are discussed in detail for different locations in the water phantom. Only the two extreme proton energies are included in this discussion as the findings of the intermediate energies are similar. The neutron fluence energy spectra for 150 MeV and 180 MeV can, however, be found in the Supplementary materials. Data obtained with TOPAS (BIC) was chosen for this discussion because it was performed and verified independently by two different users. The observations are however similar for the other MC codes and physics models.

The results of both TOPAS users show very close agreement for both physics models and all locations investigated. The comparison between the two users of the simulated neutron fluences for 110 MeV and 210 MeV was added in the Supplementary material. The differences in the total neutron fluence do not exceed 1 %. These differences are within the statistical uncertainties. Because of this very close agreement, only one user's results are shown for the remainder of this study.

Figure 3 represents the neutron fluence energy spectra at different positions in the water phantom simulated using TOPAS (BIC) for the 110 MeV proton beam (left side) and 210 MeV proton beam (right side). Data obtained with TOPAS (BIC) was arbitrarily chosen for this visualisation.

For the 110 MeV proton beam, a thermal neutron peak can be noticed ($E \leq 0.4$eV), while peaks fast (100 keV $<E \leq 19.6$MeV) and high energy neutrons ($E > 19.6$MeV) are clearly distinguished for both the distal and lateral locations. When both distal locations, 5 and 9, respectively 6 and 16 cm from the Bragg peak, are compared, similar neutron energy distribution can be noticed. Still, as expected, the neutron fluence at position 9 is lower due to the increased distance out-of-field. The lateral positions 2, 6 and 10 yield an overall decreased neutron fluence compared to the distal positions. The neutron energy distribution at lateral positions is also shifted towards the thermal-energy peak compared to the distal positions. As can be seen in Figure 3, the lateral positions 3, 7 and 11 show an analogous energy distribution to lateral positions 2, 6 and 10, yet with a lower fluence due to the increased lateral distance out-of-field. These positions are therefore omitted from further simulations, as conclusions will not provide additional information.

For the simulations with the 210 MeV proton beam, only the lateral positions are considered, as distal positions 5 and 9 are in-field for this proton energy. As with the 110 MeV proton beam simulations, the thermal neutron peak and the fast and high-energy neutron peaks can also be noticed here. For positions 2, 6 and 10, the fluence of fast and high-energy neutrons increases the shorter the scoring location is relative to the Bragg peak. The thermal neutron fluence is similar for positions 2 and 10, while position 6 shows enhanced fluence of thermal neutrons. Positions 3, 7 and 11 show a decreased neutron fluence with corresponding energy distribution due to the increased lateral distance out-of-field and are therefore no longer discussed in further simulations.

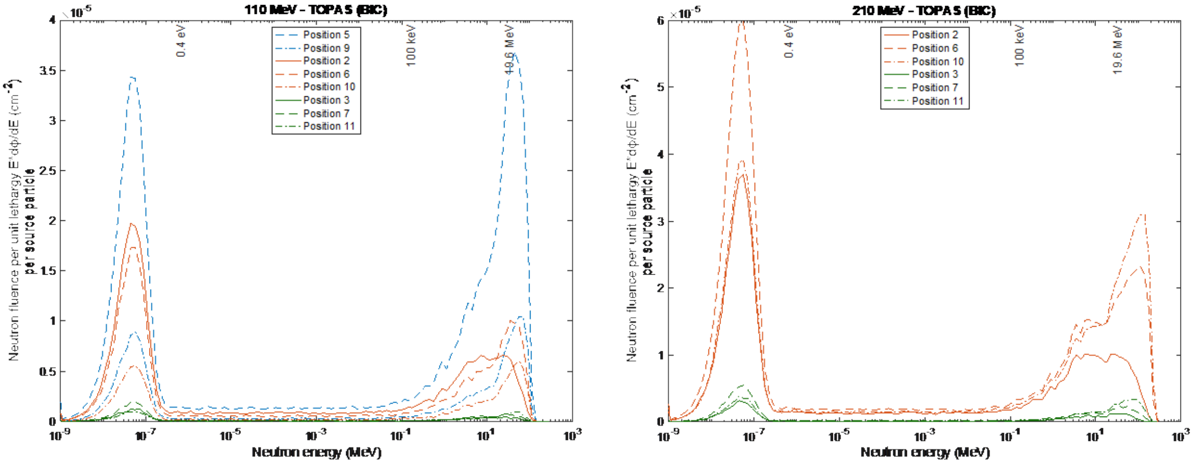

Figure 3: Neutron fluence energy spectra simulated with TOPAS (BIC) for 110 MeV (left) and 210 MeV (right) proton beams at different positions in the water phantom.

b. MC code and hadron physics models intercomparison for distal positions (110 MeV proton beam)

Figure 4 compares the simulated neutron fluence energy spectra of the 110 MeV proton beam at distal positions 5 and 9, calculated with different MC codes and physics models. This graph includes the data simulated with GEANT4 (BIC), MCNPX (Bert-Dres) and FLUKA from (De Saint-Hubert, et al., 2022). The results from TOPAS (BIC or Bert) and PHITS (INCL or Bert) are new data in this study.

Table 3 summarises the neutron fluence data simulated with the different MC codes and physics models as mentioned above, in four neutron energy regions, i.e. thermal (E ≤ 0.4eV), epithermal (0.4 eV <E ≤ 100 keV), fast (100 keV <E ≤ 19.6MeV) and high-energy (E > 19.6MeV). The results from this study are highlighted in grey. The differences between the applied physics models are compared within a specific MC code. As TOPAS uses the data from GEANT4, the TOPAS results with BIC are compared to those of GEANT4, which were also performed with the same BIC model, from (De Saint-Hubert, et al., 2022). The final column of Table 3 represents the relative difference between the maximum and minimum calculated neutron fluence across all different MC codes and physics models. Kindly note that a comparison for the distal positions for a 150 MeV and 180 MeV proton beam between the MC codes and hadron physics models included in this study was provided as Supplementary material.

A good agreement can be noted for the distal positions for the fast and high-energy neutrons between GEANT4 (BIC) and TOPAS (BIC), where PHITS (INCL) has a lower fluence. For MCNPX (Bert-Dres), PHITS (Bert), and TOPAS (Bert) an even more pronounced reduction is observed for the fast and high-energy neutron fluence. In contrast, FLUKA notes a higher fast and high-energy neutron fluence than the other codes. The differences between MC codes are lower for the thermal neutron part of the spectrum. For position 5, there is a fair agreement between MCNPX (Bert-Dres), FLUKA and TOPAS (BIC) and PHITS (INCL and Bert), where GEANT4 (BIC) has a slightly higher and TOPAS (Bert) has a somewhat lower thermal neutron fluence. For position 9, there is a fair agreement between all the MC codes and physics models, except MCNPX (Bert-Dres) and TOPAS (Bert) for which a lower thermal neutron fluence was calculated.

As can be seen in Figure 4, the TOPAS (BIC) neutron spectrum simulations agree within uncertainties with the simulation results of GEANT4 (BIC). A percentage difference for the total neutron fluence of a maximum of 3 % was observed between TOPAS (BIC) and GEANT4 using the same physics model across all distal positions. The differences are most prominent in the high-energy neutron range, as presented in Table 3.

Within TOPAS, the Bert model yields a systematically lower neutron fluence compared to the BIC model, which was on average 39 %. This underestimation is more pronounced for the fast and high-energy neutron fluence (up to 56 %) than for the thermal neutron fluence (on average 24 %).

For PHITS, the influence of using two different physics models, INCL and Bert, showed a lower neutron fluence for Bert compared to INCL, which was on average 13 %. The differences are most pronounced in the fast and high-energy ranges, compared to the thermal and epithermal energy ranges. For the fast and high-energy ranges, the differences between INCL and Bert were up to 29 % for position 5 and 39 % for position 9.

Considering in general the discrepancies between the different MC codes and physics models, represented as the relative difference between the maximum and minimum calculated neutron fluence, one can observe differences in the total neutron fluence of up to approximately 50 % at position 5 and almost 90 % at position 9 for proton energy 110 MeV. Hence, the discrepancy increases with a larger distal distance from the Bragg peak. As also described above, these differences are particularly pronounced in the fast- and high-energy neutron energy range. For example, for position 9, a difference of up to 131 % on the high-energy neutron fluence can be obtained, depending on the choice of MC code or physics model.

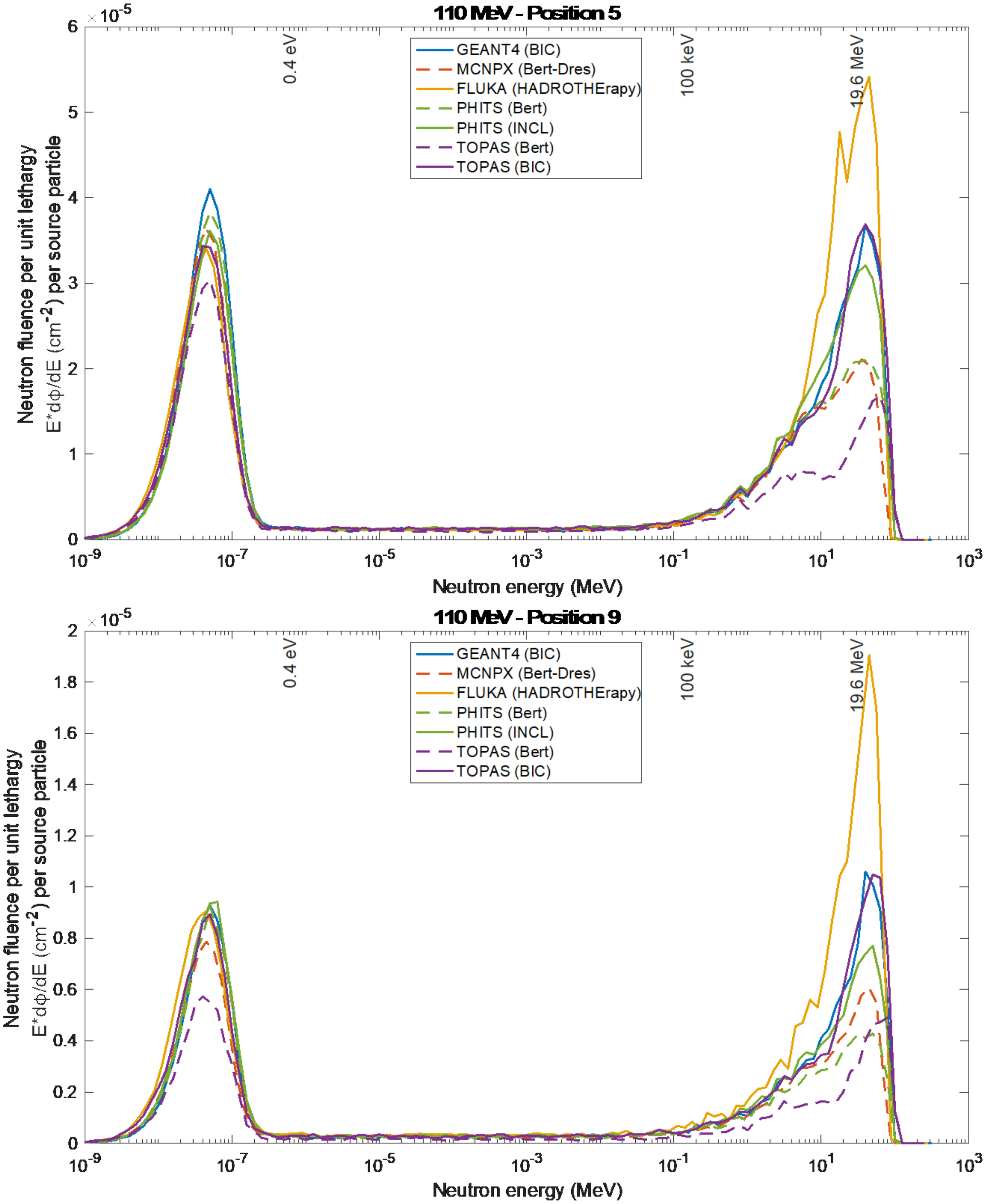

Figure 4: Neutron fluence energy spectra simulated with different MC codes and physics models for 110 MeV proton beam in a water phantom at positions 5 (top) and 9 (bottom) located distal relative to the Bragg peak.

Table 3: Secondary neutron fluence per source particle (cm^{-2}) at distal positions relative to the Bragg peak simulated with different MC codes and physics models for a 110 MeV proton beam in a water phantom, including percentage differences of TOPAS (BIC) to GEANT4 (BIC) ,TOPAS (Bert) to TOPAS (BIC) and PHITTS (Bert) to PHITS (INCL). The final column represents the percentage difference between the maximum and minimum calculated neutron fluence across all different MC codes and physics models.

MC code:		**MCNPX**	**GEANT4**	**FLUKA**	**TOPAS**				**PHITS**			
Physics model:		**Bert-Dres**	**BIC**	**HADRO THErapy**	**Bert**	**BIC**			**INCL**	**Bert**		
							Difference of BIC to GEANT4 (%)	Difference of Bert to BIC (%)			Difference of Bert to INCL (%)	$\frac{(\max - \min)}{\text{mean}}$ (%)
Scale factor:		10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}			10^{-6}	10^{-6}		
Position 5	Thermal	69.6	76.4	68.4	60.2	69.0	-10	-13	68.5	73.0	6	23
	Epith	14.6	15.6	15.1	13.4	17.1	10	-22	16.9	17.0	1	24
	Fast	44.2	51.5	64.8	25.6	43.9	-15	-42	50.3	44.3	-12	84
	High	23.9	41.1	62.7	22.4	51.0	24	-56	43.4	30.7	-29	102
	TOTAL	152	185	211	122	181	-2	-33	179	165	-8	52
Position 9	Thermal	15.0	17.1	18.8	11.5	17.5	2	-34	17.8	16.7	-6	45
	Epith	3.07	3.43	4.28	2.05	3.67	7	-44	3.69	3.26	-12	67
	Fast	9.16	11.4	16.5	4.86	9.52	-16	-49	10.3	7.92	-23	117
	High	6.74	11.5	20.2	6.33	14.1	23	-55	10.2	6.15	-39	131
	TOTAL	34.0	43.4	59.8	24.7	44.8	3	-45	41.9	34.0	-19	87

c. MC code and hadron physics models intercomparison for lateral positions (110 MeV and 210 MeV proton beam)

Figure 5 represents the simulated neutron fluence energy spectra at lateral positions 2, 6 and 10 for the 110 MeV and 210 MeV proton beams, while Table 4 presents the fluence data in four neutron energy regions. The differences between the applied physics models are compared within a specific MC code. As TOPAS uses the data from GEANT4, the TOPAS results with BIC are compared to those of GEANT4, which were also performed with the same BIC model, from (De Saint-Hubert, et al., 2022). The final column of Table 4 represents the relative difference between the maximum and minimum calculated neutron fluence across all different MC codes and physics models. A comparison for the lateral positions for a 150 MeV and 180 MeV proton beam between the MC codes and hadron physics models included in this study is provided as Supplementary material.

For the 110 MeV proton beam, the differences between the simulated neutron fluence energy spectra are generally smaller for the lateral positions compared to two distal positions.

For the 110 MeV proton beam, a reasonable agreement for the thermal peaks can be observed between all MC codes and physics models, except for TOPAS (Bert), which calculated a lower thermal neutron fluence. This difference seems to increase with larger distal distances from the Bragg peak. Considering positions 2 and 6 for 110 MeV, also FLUKA calculated a lower thermal neutron fluence. For the 210 MeV proton beam, GEANT4 (BIC), PHITS (Bert and INCL) and TOPAS (BIC) are consistent regarding thermal neutron fluence. It should, however, be noted that at this proton energy, FLUKA shows a considerably lower, while MCNPX (Bert-Dres) shows a higher thermal neutron fluence, which has been further described and discussed in the previous paper (De Saint-Hubert, et al., 2022). For positions 2 and 6, TOPAS (Bert) also yields a higher thermal neutron fluence than the other codes and physics models.

Concerning the fast and high-energy neutron fluence, a fair agreement was observed between GEANT4 (BIC), FLUKA, PHITS (Bert and INCL) and TOPAS (BIC). However, the differences for the 110 MeV proton beam extend when the distal distance from the Bragg peak increases. This becomes apparent for positions 6 and 10 where TOPAS (Bert), MCNPX (Bert-Dres) and PHITS (Bert) calculate a reduced fast and high-energy neutron fluence. For the 210 MeV proton beam, FLUKA yields a slightly reduced fast neutron fluence, and TOPAS (Bert) yields a reduced high energy neutron fluence. With MCNPX (Bert-Dres), the fast neutron fluence is somewhat higher than the other codes and physics models.

As already noted for the distal positions, TOPAS (BIC) agrees within uncertainties with GEANT4 (BIC), as can be observed in Figure 5. Table 4 notes percentage differences for the total neutron fluence of on average 2 % for 110 MeV and 5 % for 210 MeV between TOPAS and GEANT4 using the same model across all lateral positions. The differences are most pronounced in the high-energy neutron energy range, as also concluded with the distal locations.

When the differences between both physics models within TOPAS are studied, a distinction can be made between the 110 MeV and 210 MeV proton beam simulations. For 110 MeV, large differences between BIC and Bert can be observed, which was on average 29 % across the three lateral positions. The discrepancies are most pronounced in the high-energy neutron energy range and tend to increase at increasing distances. These differences are mainly because the Bert model yields a systematically lower fast and high-energy neutron fluence than the BIC model. This is consistent with the distal positions. For 210 MeV, the differences between BIC and Bert diminish to, on average 6 % across the three lateral positions. Here, the agreement is worse the closer the location is to the Bragg peak.

When considering the differences between the two physics models within PHITS, i.e. INCL or Bert, a differentiation must also be made between the 110 MeV and 210 MeV proton beam. For 110 MeV, Bert calculates a higher neutron fluence close to the Bragg peak (position 2), while in general a lower neutron fluence is obtained with Bert further away from the Bragg peak (position 10). Here, differences are most noticeable in the fast and high-energy neutron range. For 210 MeV, it should be noted that Bert calculates a lower high-energy neutron fluence, but an elevated thermal neutron fluence, which results in a higher total neutron fluence calculated with Bert compared to INCL. The agreement between both models at 210 MeV is better at locations closer to the Bragg peak.

The general discrepancy between the different MC codes and physics models is smaller for the lateral positions compared to the distal positions. For the 110 MeV proton beam, the largest disagreement on the total neutron fluence for the lateral positions rises with increasing distances after the Bragg peak from 26 % to more than 60 %. As for the distal positions, the differences in the lateral positions for the 110 MeV proton beam are particularly pronounced in the fast- and high-energy neutron energy range, up to almost 100 % for the fast neutron fluence for position 10. For the 210 MeV proton beam, the disagreement remains approximately constant among the various positions. Still, differences of up to approximately 40 % on the total neutron fluence can be obtained, depending on the choice of MC code or physics model. For the 210 MeV proton beam, unlike the 110 MeV proton beam, there is no pronounced variation in the neutron fluences split into the four energy ranges.

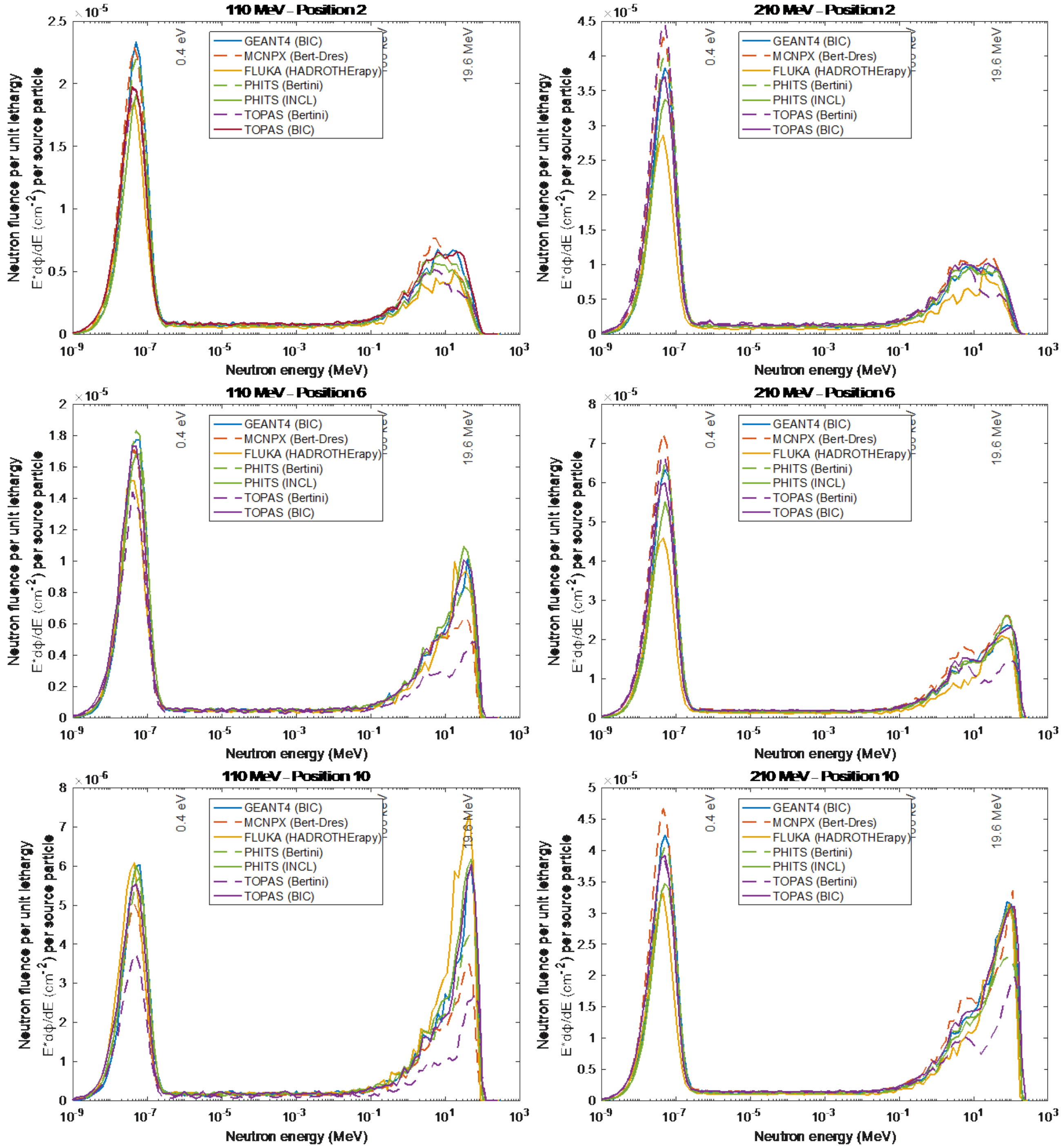

Figure 5: Neutron fluence energy spectra simulated with different MC codes and physics models for 110 MeV (left side) and 210 MeV (right side) proton beam in a water phantom at positions 2 (top), 6 (centre) and 10 (bottom) located lateral relative to the Bragg peak.

1

Table 4: Secondary neutron fluence per source particle (cm^{-2}) at lateral positions relative to the Bragg peak simulated with different MC codes and physics models for a 110 MeV and 210 MeV proton beam in a water phantom, including percentage differences of TOPAS (BIC) to GEANT4 (BIC) ,TOPAS (Bert) to TOPAS (BIC) and PHITTS (Bert) to PHITS (INCL). The final column represents the percentage difference between the maximum and minimum calculated neutron fluence across all different MC codes and physics models.

MC code		MCNPX	GEANT4	FLUKA	TOPAS				PHITS			
Physics model		Bert-Dres	BIC	HADRO THErapy	Bert	BIC	Difference of BIC to GEANT4 (%)	Difference of Bert to BIC (%)	INCL	Bert	Difference of Bert to INCL (%)	(max – min) / mean (%)
Scale factor:		10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}			10^{-6}	10^{-6}		
Position 2	Thermal	43.9	43.0	36.4	39.5	40.6	-6	-3	35.9	42.3	18	20
	Epith	10.3	9.6	7.29	10.3	10.8	13	-5	8.80	10.4	18	36
	Fast	23.0	21.0	14.9	16.8	20.8	-1	-19	17.3	20.3	18	42
	High	4.09	4.82	3.82	4.16	7.32	52	-43	4.45	5.42	22	72
	TOTAL	81.3	78.4	62.4	70.7	79.5	1	-11	66.4	78.5	18	26
Position 6	Thermal	32.6	34.1	29.9	27.7	34.4	1	-20	33.0	34.9	6	22
	Epith	5.98	6.33	5.52	4.84	6.87	8	-30	6.42	6.73	5	33
	Fast	15.8	17.7	16.0	9.42	15.9	-10	-41	16.6	15.7	-6	54
	High	7.02	10.1	10.9	6.59	13.5	34	-51	13.1	11.2	-14	67
	TOTAL	61.4	68.2	62.4	48.5	70.8	4	-31	69.1	68.6	-1	35
Position 10	Thermal	9.59	11.2	11.6	7.39	11.3	1	-35	11.5	10.9	-5	40
	Epith	1.95	2.31	2.54	1.34	2.34	1	-43	2.41	2.12	-12	56
	Fast	5.70	7.20	8.70	3.01	6.00	-17	-50	6.91	5.87	-15	92
	High	3.84	6.21	8.30	3.62	7.67	24	-53	7.62	5.94	-22	76
	TOTAL	21.1	26.9	31.2	15.4	27.3	2	-44	28.5	24.9	-13	63
Position 2	Thermal	81.8	71.9	55.3	88.3	73.9	3	19	65.3	78.5	20	45
	Epith	16.7	14.6	10.2	19.4	17.2	18	13	14.1	16.3	16	60
	Fast	35.4	31.6	21.2	32.2	32.8	4	-2	29.5	31.8	8	46
	High	14.2	12.7	10.8	9.82	16.4	29	-40	14.1	14.4	3	50
	TOTAL	148	131	97.6	150	140	7	7	123	141	15	39
Position 6	Thermal	138	119	91.1	133	120	1	11	104	123	19	40
	Epith	25.6	21.5	16.0	25.3	23.8	11	6	19.7	23.2	18	43
	Fast	57.4	48.1	35.2	42.7	48.1	0	-11	43.2	46.6	8	48
	High	44.6	41.3	37.6	28.3	47.2	14	-40	45.3	40.4	-11	46
	TOTAL	266	230	180	229	239	4	-4	212.2	233.8	10	38
Position 10	Thermal	89.2	79.5	62.2	78.3	79.0	-1	-1	67.0	78.4	17	35
	Epith	19.7	17.1	13.6	17.7	19.3	13	-8	15.6	17.9	15	35
	Fast	50.8	45.6	35.8	32.1	43.9	-4	-27	37.9	41.3	9	46
	High	49.6	53.8	52.6	33.7	60.1	12	-44	54.0	46.0	-15	53
	TOTAL	209	196	164	162	202	3	-20	174	184	5	26

3. Neutron dose equivalent

a. Comparison of TOPAS (BIC) simulations for 110 MeV, 150 MeV, 180 MeV and 210 MeV proton beam

The neutron dose equivalent was estimated based on the simulated neutron fluence energy spectra for distal positions 5 and 9 and lateral positions 2, 6 and 10 for proton energies 110 MeV, 150 MeV, 180 MeV and 210 MeV. In Figure 6, the left-hand side shows the neutron dose equivalent per source particle for positions 5 and 9, located on the beam's central axis. The right-hand side of Figure 6 represents the neutron dose equivalent per source particle for positions 2, 6 and 10, located at 6 cm lateral to the edge of the field. Since the distance out-of-field for the considered locations differs for the different proton energies, the neutron dose equivalent is shown as a function of the distance with respect to the Bragg peak. A positive distance corresponds to a location behind the Bragg peak, whereas a negative distance corresponds to a location before the Bragg peak. All the results in Figure 6 are simulated with TOPAS (BIC) as they were performed and verified independently by two different users. The observations are however similar for the other MC codes and physics models.

When the two distal positions on the central axis are considered (left-hand side, Figure 6), one can observe that the neutron dose equivalent inversely correlates to the distance to the Bragg peak. Due to this limited number of positions investigated, it is difficult to relate the impact of the proton energy to the neutron dose equivalent here.

For the three positions located at 6 cm lateral to the edge of the field (right-hand side, Figure 6), the neutron dose equivalent estimation is higher when closer to the Bragg peak (which corresponds to a 0 cm distance), within each energy. However, for the 180 MeV proton beam, the neutron dose equivalent is 15 % higher (3.5 10^{-11} mSv versus approximately 3.1 10^{-11} mSv) at -13 cm versus 11 cm. Positions closer to the entrance window of the beam in the water phantom (position 2) further yield lower neutron dose equivalents when compared to positions that are positioned towards the back of the water phantom (position 10). Furthermore, the proton energy correlates positively with the neutron dose equivalent. For example, position 2 for the 110 MeV proton beam (blue circular marker) with 1.13 10^{-11} mSv yields an approximately four times lower neutron dose equivalent than position 10 for the 210 MeV proton beam (purple triangular marker) with 4.38 10^{-11} mSv, even while the latter is located at a greater distance from the Bragg peak. Position 6 for the 150 MeV proton beam (red square marker) and 180 MeV proton beam (yellow square marker) have values between the previous two extreme energies. The same can also be seen at position 6, 110 MeV (blue square marker), position 10, 150 MeV (red triangular marker) and position 10, 180 MeV (yellow triangular marker), where the neutron dose equivalent increases from 1.27 10^{-11} mSv, to 2.75 10^{-11} mSv and 3.51 10^{-11} mSv, respectively.

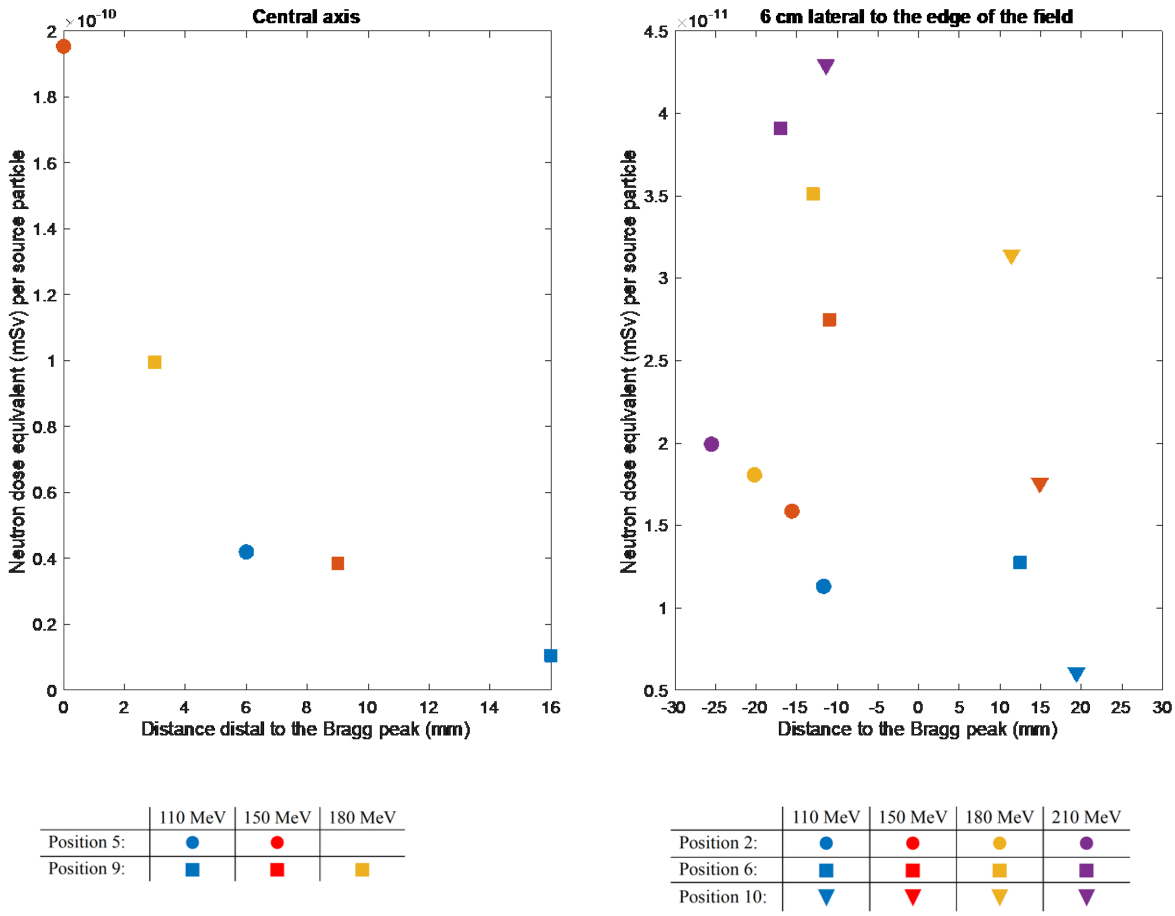

Figure 6: Neutron dose equivalent (mSv) per source particle as function of the distance to the Bragg peak (where a positive distance corresponds to a distance posterior to the Bragg peak, a negative distance anterior to the Bragg peak) on the beam's central axis (left) and 6 cm lateral to the edge of the field (right) for 110 MeV (blue), 150 MeV (orange), 180 MeV (yellow) and 210 MeV (purple) proton beams simulated with TOPAS (BIC).

b. MC code and hadron physics models intercomparison

Table 5 compares the neutron dose equivalents calculated based on the neutron fluence energy spectra simulated with the different MC codes and physics models at the different locations for the 110 MeV and 210 MeV proton beam. The percentage difference of the maximum and minimum neutron dose equivalent estimated by the different MC codes and physics models compared to the mean was added to the results. This table includes the data simulated with GEANT4 (BIC), MCNPX (Bert-Dres) and FLUKA from (De Saint-Hubert, et al., 2022). The results from TOPAS (Bert or BIC) and PHITS (INCL or Bert), highlighted in grey, are the new data in this study. A comparison of the neutron dose equivalents for a 150 MeV and 180 MeV proton beam of the MC codes and hadron physics models included in this study is provided as Supplementary material.

Generally, one can notice substantial discrepancies between the different MC codes and physics models for the neutron dose equivalent estimates. The differences are most pronounced at the distal positions, compared to the lateral positions. For example, for the 110 MeV proton beam at position 5, there is a difference of a factor of 2 between the lowest and highest estimated neutron dose equivalent relative to the average of the different MC codes and physics models; for position 9, this is between a factor of 2 and 3. For the lateral positions at 110 MeV, it is observed that the larger the distance is from the Bragg peak, the larger the discrepancy between the MC codes is. The relative difference between the extrema here ranges from about 40 % at position 2 to 90 % at position 10. For 210 MeV, the discrepancies at the lateral positions remain constant, there is no immediate influence of distance from the Bragg peak to note. Here, the relative difference between the extrema is approximately 40 %. In addition to describing these maximum differences, it should be noted, though, that there is reasonable agreement between the neutron dose equivalent calculated with GEANT4 (BIC), TOPAS (BIC) and PHITS (INCL). These MC codes with corresponding physics models agree within 10 %, except for position 2 for 110 MeV.

Table 5: Neutron dose equivalent per source particle (mSv) simulated with different MC codes and physics models for 110 MeV and 210 MeV proton beams in the water phantom.

	110 MeV Distal position		110 MeV Lateral position			210 MeV Lateral position		
	5	**9**	**2**	**6**	**10**	**2**	**6**	**10**
Scale factor:	10^{-12}	10^{-12}	10^{-12}	10^{-12}	10^{-12}	10^{-12}	10^{-12}	10^{-12}
MCNPX (Bert-Dres)	29.5	6.95	11.0	9.77	4.16	20.8	42.7	42.1
GEANT4 (BIC)	41.3	10.3	10.5	12.1	5.95	18.3	37.7	42.5
FLUKA (HADROTHErapy)	58.5	16.9	7.65	12.0	7.75	13.5	31.5	38.7
TOPAS (Bert)	20.3	4.84	8.08	6.70	2.88	16.2	28.4	26.8
TOPAS (BIC)	42.0	10.6	11.3	12.7	6.07	20.0	39.5	43.8
PHITS (INCL)	40.8	8.93	8.56	12.9	6.43	17.7	37.0	38.9
PHITS (Bert)	31.8	6.01	10.1	11.5	5.13	17.7	37.0	38.9
$\frac{(\text{max}-\text{min})}{\text{mean}}$ (%)	101	131	38	56	89	41	39	44

Discussion

1. Range verification

To verify if the source is modelled correctly within the different codes and discrepancies found out-of-field are not related to major disagreements in the field, we first looked at the proton range and depth-dose profiles in the field for the four different mono-energetic proton beams, 110 MeV, 150 MeV, 180 MeV and 210 MeV. The R90 values for all MC codes and physics models agree closely, i.e., within 1.5 mm. Minor differences can also be noted in the absolute absorbed dose per history in depth. These depth-dose profiles were studied for MCNPX (Bert-Dres), TOPAS with BIC and Bert and PHITS with INCL and Bert. No depth-dose profiles were available for GEANT4 and FLUKA from the previous study (De Saint-Hubert, et al., 2022). Although all codes are consistent in the plateau region of the curve, differences are noticeable in the Bragg peak and most considerable in the dose build-up region. There is a faster dose build-up for PHITS (Bert), whereas for MCNPX (Bert-Dres), a lower absolute dose per history is noted in the Bragg peak. This was also described for MCNPX (Bert-Dres) in the study by Jette et al (2011) (Jette & Chen, 2011). A fair agreement is pointed out for the other simulations. Differences in electromagnetic physics between the different MC codes, i.e., differences in range straggling, may explain these differences in the steep dose gradient regions. This is compatible with the data of Vedelago et al. from 2022 (Vedelago, et al., 2022), who also observed a good agreement with minor differences in simulations of the depth-dose profiles of mono-energetic proton beams with FLUKA, TOPAS and MCNP. This study concluded that the differences at 221.1 MeV were larger than at 100.77 MeV. This is not noticeable in our results. This study focuses on secondary neutrons out-of-field and we accepted these small differences for the further purpose of this study.

2. Neutron fluence energy spectra

The simulated neutron fluence energy spectra show a similar distribution, i.e., a thermal – epithermal and fast - high-energy component. Distal locations show a more prominent fast and high-energy peak compared to lateral locations. This can be explained by the fact that most fast and high-energy neutrons are emitted in the forward direction of the beam, as a result of prompt neutron emission. This contrasts with thermal neutrons and evaporation (fast) neutrons, which have a rather isotropic emission as they are emitted by excited target nuclei. This is consistent with the findings of Vedelago (Vedelago, et al., 2022), who investigated the angular distribution of secondary neutrons by MC simulations.

The study performed by (De Saint-Hubert, et al., 2022) concluded that the most considerable differences between the MC codes in their research, i.e., GEANT4, FLUKA and MCNPX (Bert-Dres), arose in the distal positions relative to the beam. This study added two MC codes with two different physics models to the comparison. New data included simulations with TOPAS with BIC or Bert and PHITS with INCL or Bert physics models. Also, this extended comparison shows the most prominent differences in the high-energy neutron fluence at the distal scoring positions. Whereas PHITS (Bert) agrees well with MCNPX (Bert-Dres), the other codes PHITS (INCL), TOPAS (BIC), GEANT4 and FLUKA calculate higher high-energy neutron fluences. As also stated by (De Saint-Hubert, et al., 2022), these differences can be explained by the more prominent presence of the high-energy component in the neutron fluence energy spectra that originates in the forward direction of the beam. This is also visible in Figure

3, which shows the neutron fluence energy spectra simulated with TOPAS (BIC) at different OOF locations. Differences in this high-energy neutron fluence originate from whether neutron cross-section data are available and activated within an MC code. De Saint-Hubert et al (2022) (De Saint-Hubert, et al., 2022) pointed out that cross-section data are missing for neutrons with an energy higher than 20 MeV in GEANT4 and FLUKA. Since TOPAS uses the data from GEANT4, the simulation results of TOPAS (BIC) in this study are consistent with those of GEANT4 from (De Saint-Hubert, et al., 2022), which also used the BIC as a hadron physics model. MCNPX (Bert-Dres), on the other hand, does have nuclear cross-section data up to 150 MeV (De Saint-Hubert, et al., 2022). PHITS does also include neutron cross-section data for a broad energy spectrum. It employs JENDL below 20 MeV and INCL or a modified version of Bert above 20 MeV (Japan Atomic Energy Agency, 2007) (Boudard, Cugnon, David, Leray, & Mancusi, 2013) (Sawada, et al., 2012) (Takada, Yoshizawa, Kosako, & Ishibashi, 1998). The differences in the thermal peak are lower for the distal positions when the different codes are compared.

For the lateral positions, it was observed by (De Saint-Hubert, et al., 2022) that the variations between the three codes studied, GEANT4, FLUKA and MCNPX (Bert-Dres), are more minor and less pronounced for the high-energy neutron fluence. It was noted that these differences depended mainly on the relative position of the scoring volume to the Bragg peak. A better agreement between the MC codes is observed for the lateral positions. The more the scoring volumes are located in the forward direction, i.e. a smaller angle to the beam direction, the more dominant the high-energy component in the neutron fluence energy spectra is, leading to discrepancies between the different MC codes and physics models, as already described above. This is also valid for the extended comparison. A higher neutron fluence can be observed for scoring volumes located in a more forward position for FLUKA, GEANT4, TOPAS (BIC) and PHITS with either INCL or Bert, compared to MCNPX (Bert-Dres).

In the study of Vedelago et al. (Vedelago, et al., 2022), simulating secondary neutron fluence energy spectra in a water phantom with FLUKA, TOPAS and MCNPX, large discrepancies were also observed between the different MC codes. Although the differences between the MC codes were not discussed in their work, it can be seen that, for example, for a 110.77 MeV proton beam at a location distal to the proton beam, MCNP calculates a higher thermal and fast and high-energy neutron fluence than FLUKA and that TOPAS calculates a higher thermal and fast and high-energy neutron fluence than MCNP. However, making a comparison with the results from this study is impossible because different settings and hadron physics models were used for the respective MC codes.

This work describes the differences in OOF neutron fluence energy spectra between MC codes and nuclear models. Depending on the choice of MC code or physics model, discrepancies of approximately 90 % on the total neutron fluence can be obtained. For the 110 MeV proton beam, these discrepancies are larger for the distal positions than for the lateral positions and rise in both cases with increasing distance from the Bragg peak. Here the disagreement is most noticeable in the fast and high-energy neutron range (up to 131 %). For the 210 MeV proton beam, the discrepancies for the lateral positions seem somewhat lower than for the 110 MeV proton beam, here no immediate dependence on distance can be seen. The discrepancy is also common when considering the entire neutron energy range, rather than pronounced in the fast and high-energy range.

Among both MC codes PHITS and TOPAS, there is not a solid agreement of the neutron fluence energy spectra in terms of the selection of physics model INCL or Bert, and BIC or Bert, respectively. Within PHITS, considerable differences can be observed in the neutron fluence energy spectra. Except for the neutron fluence scored at position 2 for both the 110 MeV and 210 MeV proton beam, Bert calculates a lower high-energy neutron fluence than INCL. For the TOPAS simulations, the differences between the hadron physics models are even more pronounced. Major differences can be noted between BIC versus Bert for the simulated neutron fluence energy spectra with TOPAS. Here, Bert systematically calculates a lower high-energy neutron fluence. Although variations exist between MC codes and models as described above, this still seems anomalous. Previous data have always favoured the BIC model more than Bertini because of a better agreement with experimental data (Englbrecht, et al., 2021), (Arce, et al., 2021).

As cited earlier, TOPAS uses GEANT4 data to perform the simulations. Nevertheless, differences of up to 5 % on the total neutron fluence between TOPAS (BIC) and GEANT4 (BIC) can be noted. The differences are more pronounced at higher proton energies and at locations distal with respect to the Bragg peak and are especially present in the high-energy neutron energy range. The different versions of the nuclear models used can explain these. The TOPAS simulations from this study use GEANT4 data version 10.7, whereas the GEANT4 simulations from (De Saint-Hubert, et al., 2022) use version 10.1.2. By including the results from FLUKA and MCNPX from (De Saint-Hubert, et al., 2022), the limitation arises that for these codes, the most recent version with possibly adapted physics models, is not included in the comparison.

3. Neutron dose equivalent

The neutron dose equivalent at different distal and lateral positions with respect to the Bragg peak, presented in Figure 6, facilitated a detailed study of the variation of the neutron dose equivalent with the position inside the water phantom. The neutron dose equivalent is highest for positions located distal to the Bragg peak, compared to lateral positions. For either lateral or distal positions, the neutron dose equivalent is highest closest to the Bragg peak. A higher proton energy also yields a higher neutron dose equivalent for the lateral positions. This is consistent with the findings of Van Hoey et al. in 2022 (Van Hoey, et al., 2022). Here, it was concluded that the ambient neutron dose inversely correlates with the distance from the isocenter and that it is more pronounced in the forward direction of the beam, compared to the backward direction. An increasing proton energy was also identified as one of the strongest influences.

It is evident that the differences in the neutron spectrum between the MC codes and physics models propagate into differences between the estimated neutron dose equivalents. The differences are more prominent for the distal positions; here, a difference of up to a factor of 3 in the neutron dose equivalent can be obtained, depending on which MC code and physics model was selected. For the lateral positions, discrepancies exist, but they are smaller than for the distal positions. Comparing the lateral positions, the agreement is better for 210 MeV, here the difference between the smallest and largest neutron dose equivalent estimate is about 40 %. For 110 MeV, this is approximately 40 % closest to the Bragg peak, but increases with distance from the Bragg peak. Differences in the simulation of the high-energy neutrons, as described earlier, can largely explain these discrepancies. It can be noted, however, that GEANT4 (BIC), TOPAS (BIC) and PHITS (INCL) are in fairly similar agreement, with in most cases a difference of lower than 10 %.

4. Strengths, limitations and future prospects

It is pointed out that the TOPAS simulations from this study and the MCNPX simulations from (De Saint-Hubert, et al., 2022) were each performed by different users, independent from each other, to study the inter-user variability. This is a very instructive process to identify potential errors and optimise this and future simulations. Simulation results agreed within the statistical uncertainty.

It should however be noted that no experimental validation was carried out. It is therefore not possible to conclude on the performance of any code or nuclear model. Therefore, no comparisons were performed relative to a reference either. Figure 3 and Figure 6 elaborate on the results of TOPAS (BIC), but this MC code and physics model have been chosen arbitrarily. Future work includes validation measurements with, for example, a bubble detector for personal neutron dosimetry (BD-PND), a wide-energy neutron detection instrument (WENDI-2), a Bonner sphere spectrometer or an electrochemically-etched track detector (CR39). However, it should be emphasized that proper validation of MC codes and nuclear models is challenging to perform, mainly because the detector response is energy-dependent and in many cases requires MC simulations for energy response corrections and appropriate calibrations. It should be additionally noted that only limited information exists on the performance of MC codes and nuclear models, which makes an informed selection more difficult. This highlights the importance of benchmarking studies such as this one (De Saint-Hubert, et al., 2022), (Romero-Expósito, Domingo, Sánchez-Doblado, Ortega-Gelabert, & Gallego, 2016), (Hälg, et al., 2012).

Within our study a purely theoretical setup was used, i.e., a mono-energetic parallel proton beam hitting a water phantom. This can be considered a limitation as it limits the translatability to clinical plans. Moreover, the beam divergence was not considered which may result in a larger field size in depth. Therefore, future validation measurements, require the implementation of the beam model in the MC simulations accordingly. It is then also needed to evaluate whether the accelerator, beamline or in-room components should be included in the simulation to obtain an accurate estimate of the secondary neutron dose. In a clinical situation, fast and high-energy neutrons have sufficient energy to be scattered into distant objects and possibly return to the patient (Englbrecht, et al., 2021). Moreover, the effect of adding a range shifter was not included in this study. Adding a range shifter in the beam has a major impact on the neutron dose. Van Hoey et al. (Van Hoey, et al., 2022), recently reported that the ambient neutron dose doubled in measurements and simulations when a range shifter was added to the beam. The impact of a range shifter on the neutron dose equivalent estimated by different MC codes and hadron physics models will be explored in a future study.

An additional limitation is that simulations in this study were performed in water and the impact of different materials and material heterogeneities on calculating the OOF neutron dose was not studied, limiting the clinical translatability Future work will include a study on the impact of different MC codes and hadron physics models on relevant clinical treatment plans in anthropomorphic phantom, e.g. the in-utero dose in the case of the irradiation of a pregnant mother.

Conclusion

Considerable differences are observed in calculating the neutron fluence energy spectra OOF of mono-energetic proton beams in a water phantom between different MC codes and nuclear models. The differences are most pronounced in distal positions, in contrast to lateral locations, due to the more prominent presence of high-energy neutrons there and the higher uncertainty on the neutron cross-section data for these high-energy neutrons in the simulations. Depending on the choice of MC code or physics model, discrepancies of 90 % on the total neutron fluence can be obtained. In most cases, the disagreement is mostly noticeable in the fast and high-energy neutron fluence (up to 131 %). The nuclear model selection within PHITS, either INCL or Bertini, impacts the simulation of the neutron fluence energy spectra, especially for distal positions. The nuclear model selection within TOPAS, BIC or Bertini, produces systematic and substantial differences for distal and lateral locations. For the two MC codes, the Bertini model shows worse agreement compared to the other MC codes and nuclear models for OOF neutron simulations.

Differences in the neutron spectrum simulations propagate into differences in the estimated neutron dose equivalents. The differences between the MC codes and physics models are more prominent for the distal positions (up to a factor of 3) than for the lateral positions (up to a factor of 2) and decrease at higher proton energies.

Acknowledgements

The resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation - Flanders (FWO) and the Flemish Government.

References

Perl, J., Shin, J., Schumann, J., Faddegon, B., & Paganetti, H. (2012). TOPAS: an innovative proton Monte Carlo platform for research and clinical applications. *Medical Physics, 39*(11), 6818-6837.

Bordy, J., Bessieres, I., d'Agostino, E., Domingo, C., d'Errico, F., di Fulvio, A., . . . Vermesse, D. (2013). Radiotherapy out-of-field dosimetry: Experimental and computational results for photons in a water tank. *Radiation Measurements, 57*, 29-34.

Romero-Expósito, M., Domingo, C., Sánchez-Doblado, F., Ortega-Gelabert, O., & Gallego, S. (2016). Experimental evaluation of neutron dose in radiotherapy patients: Which dose? *Medical Physics, 43*(1), 360-367.

Chadwick, M., Barschall, H., Caswell, R., DeLuca, P., Hale, G., Jones, D., . . . Young, P. (1999). "A consistent set of neutron kerma coefficients from thermal to 150 Mev for biologically important materials. *Medical Physics, 26*(1), 974-991.

Siebert, B., & Schuhmacher, H. (1995). Quality factors, ambient and personal dose equivalent for neutrons based on the new ICRU stopping power data for protons and alpha particles. *Radiation Protection Dosimetry, 58*(1), 177-183.

Sato, T., Iwamoto, Y., Hashimoto, S., Ogawa, T., Furuta, T., Abe, S.-i., . . . Niita, K. (2018). Features of Particle and Heavy Ion Transport code System (PHITS) version 3.02. *Journal of Nuclear Science and Technology, 55*(6), 684-690.

Japan Atomic Energy Agency. (2007). *PHITS - Particle and Heavy Ion Transport code System*. Retrieved March 22, 2023, from https://phits.jaea.go.jp

Lansonneur, P., Mammar, H., Nauraye, C., Patriarca, A., Hierso, E., Dendale, R., . . . De Marzi, L. (2020). First proton minibeam radiation therapy treatment plan evaluation. *Scientific reports, 10*(1), 7025.

Boudard, A., Cugnon, J., David, J.-C., Leray, S., & Mancusi, D. (2013). New potentialities of the Liège intranuclear cascade model for reactions induced by nucleons and light charged particles. *Physical Review C, 87*(014606).

Sawada, Y., Uozumi, Y., Nogamine, S., Yamada, T., Iwamoto, Y., Sato, T., & Niita, K. (2012). Intranuclear cascade with emission of light fragment code implemented in the transport code system PHITS. *Nuclear Instruments and Methods in Physics Research Section B, 291*, 38-44.

Takada, H., Yoshizawa, N., Kosako, K., & Ishibashi, K. (1998). An Upgraded Version of the Nucleon Meson Transport Code: NMTC/JAERI97. *Japan Atomic Energy Research Institute, JAERI-Data/Code 98-005*.

Lee, C., Lee, S., Lee, S.-J., Song, H., Kim, D.-H., Cho, S., . . . Kim, J. (2017). Monte Carlo simulation of secondary neutron dose for scanning proton therapy using FLUKA. *PLOS ONE, 12*(10).

Englbrecht, F., Trinkl, S., Mares, V., Rühm, W., Wielunski, M., Wilkens, J., . . . Parodi, K. (2021). A comprehensive Monte Carlo study of out-of-field secondary neutron spectra in a scanned-beam proton therapy gantry room. *Zeitschrift für Medizinische Physik, 31*(2), 215-228.

Schneider, U., Agosteo, S., Pdroni, E., & Bessere, J. (2002). Secondary neutron dose during proton therapy using spot scanning. *International Jornal of Radiation Oncology, 53*(1), 244-251.

Farah, J., Martinetti, F., Sayah, R., Lacoste, V., Donadille, L., Trompier, F., . . . Clairand, I. (2014). Monte Carlo modeling of proton therapy installations: a global experimental method to validate secondary neutron dose calculations. *Physics in Medicine & Biology, 59*(11).

Mohan, R. (2022). A review of proton therapy – Current status and future directions. *Precision Radiation Oncology, 6*, 164–176.

Gottschalk, B. (2006). Neutron dose in scattered and scanned proton beams: In regard to Eric J. Hall (Int J Radiat Oncol Biol Phys 2006;65:1–7). *International Journal of Radiation Oncology*Biology*Physics, 66*(5), 1594.

De Saint-Hubert, M., Farah, J., Klodowska, M., Romero-Expósito, M., Tyminska, K., Mares, V., . . . Trinkl, S. (2022). The influence of nuclear models and Monte Carlo radiation transport codes on stray neutron dose estimations in proton therapy. *Radiation Measurements, 150*.

Faddegon, B., Ramos-Méndez, J., Schümann, J., McNamara, A., Shin, J., Perl, J., & Paganetti, H. (2020). The TOPAS Tool for Particle Simulation, a Monte Carlo Simulation Tool for Physics, Biology and Clinical Research. *Physica Medica, 72*, 114-121.

Zacharatou, C., & Paganetti, H. (2008). Physics Settings for Using the Geant4 Toolkit in Proton Therapy. *IEEE Transactions on Nuclear Science, 55*(3), 1018-1025. Retrieved from https://topas.readthedocs.io/en/latest/parameters/physics/intro.html

TOPAS MC Inc. (2022). *Physics*. Retrieved May 24, 2023, from https://topas.readthedocs.io/en/latest/parameters/physics/index.html

Vedelago, J., Geser, F., Muñoz, I., Stabilini, A., Yukihara, E. G., & Jäkel, O. (2022). Assessment of secondary neutrons in particle therapy by Monte Carlo simulations. *Physics in Medicine and Biology, 67*(015008).

Hälg, R., Besserer, J., Boschung, M., Mayer, S., Clasie, B., Kry, S., & Schneider, U. (2012). Field calibration of PADC track etch detectors for local neutron dosimetry in man using different radiation qualities. *Nuclear Instruments and Methods in Physics Research A, 694*, 205-210.

Schneider, U., & Hälg, R. (2015). The impact of neutrons in clinical proton therapy. *Frontiers in Oncology, 5*(235).

Wright, D., & Incerti, S. (s.d.). *A short guide to choosing physics lists.* Retrieved September 30, 2023, from http://geant4.in2p3.fr/IMG/pdf_PhysicsLists.pdf

Novak, M. (2019, February 4). *Short Guide to Choosing Your Physics LIsts.* Retrieved September 30, 2023, from https://indico.cern.ch/event/776050/contributions/3241826/attachments/1789270/2914266/ChoosingPhysLists.pdf

Arce, P., Bolst, D., Bordage, M.-C., Cirrone, P., Cortes-Giraldo, M. A., Cutajar, D., . . . Wright, D. (2020, July 24). *G4_Med, a Geant4 benchmarking tool for medical physics applications.* Retrieved September 30, 2023, from http://ionimaging.org/assets/talks/ws2020llu-susannaguatelli.pdf

Arce, P., Bolst, D., Cutajar, D., Guatelli, S., Le, A., Rosenfeld, A., . . . Wright, D. (2021). Report on G4–Med, a Geant4 benchmarking system for medical physics applications developed by the Geant4 Medical Simulation Benchmarking Group. *Medical Physics, 48*(1), 19-56.

De Saint-Hubert, M., Verbeek, N., Bäumer, C., Esser, J., Wulff, J., Nabha, R., . . . Brualla, L. (2022). Validation of a Monte Carlo Framework for Out-of-Field Dose Calculations in Proton Therapy. *Frontiers in Oncology, 12*.

De Smet, V., De Saint-Hubert, M., Dinar, N., Manessi, G. P., Aza, E., Cassell, C., . . . Vanhavere, F. (2017). Secondary neutrons inside a proton therapy facility: MCNPX simulations compared to measurements performed with a Bonner Sphere Spectrometer and neutron H*(10) monitors. *Radiation Measurements, 99*, 25-40.

Jette, D., & Chen, W. (2011). Creating a spread-out Bragg peak in proton beams. *Physics in Medicine and Biology, 56*, N131-N138.

Knežević, Ž., Stolarczyk, L., Ambrožová, I., Caballero-Pacheco, M. Á., Davídková, M., De Saint-Hubert, M., . . . Olko, P. (2022). Out-of-Field Doses Produced by a Proton Scanning Beam Inside Pediatric Anthropomorphic Phantoms and Their Comparison With Different Photon Modalities. *Frontiers in Oncology, 12*(904563).

Van Hoey, O., Stolarczyk, L., Lillhök, J., Eliasson, L., Mojzeszek, N., Liszka, M., . . . Olko, P. (2022). Simulation and experimental verification of ambient neutron doses in a pencil beam scanning proton therapy room as a function of treatment plan parameters. *Frontiers in Oncology, 12*(903537).