Validation of nuclear models in Geant4 using the dose distribution of a 177 MeV proton pencil beam

David C Hall$^{1,4}$, Anastasia Makarova$^{2,4}$, Harald Paganetti$^1$ and Bernard Gottschalk$^3$

1 Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA
2 NCTeam, Department of Radiation Oncology, Medical Faculty, University of Duisburg-Essen, North Rhine-Westphalia 45122, Germany
3 Harvard University Laboratory for Particle Physics and Cosmology, 18 Hammond St., Cambridge, MA 02138, USA

E-mail: dchall@mgh.harvard.edu and anastasia.makarova@uni-due.de

Received 30 July 2015, revised 28 October 2015
Accepted for publication 29 October 2015
Published 27 November 2015

Abstract

A proton pencil beam is associated with a surrounding low-dose envelope, originating from nuclear interactions. It is important for treatment planning systems to accurately model this envelope when performing dose calculations for pencil beam scanning treatments, and Monte Carlo (MC) codes are commonly used for this purpose. This work aims to validate the nuclear models employed by the Geant4 MC code, by comparing the simulated absolute dose distribution to a recent experiment of a 177 MeV proton pencil beam stopping in water.

Striking agreement is observed over five orders of magnitude, with both the shape and normalisation well modelled. The normalisations of two depth dose curves are lower than experiment, though this could be explained by an experimental positioning error. The Geant4 neutron production model is also verified in the distal region. The entrance dose is poorly modelled, suggesting an unaccounted upstream source of low-energy protons. Recommendations are given for a follow-up experiment which could resolve these issues.

Keywords: proton therapy, pencil beam scanning, Monte Carlo, nuclear halo

(Some figures may appear in colour only in the online journal)

$^4$ Both authors have contributed equally to this work.
1. Introduction

Pencil beam scanning (PBS) is a mode of proton therapy whereby each radiation field is constructed from a large number of narrow pencil beams. It is becoming increasingly popular, and in future most proton therapy treatments are likely to be delivered by this treatment technique.

In a PBS radiation field, the absolute dose at a point of interest (POI) in a patient or phantom is computed by adding the contributions of those pencil beams sufficiently near the POI to influence it. In defining this distance of influence, it is important to consider that the primary core of each pencil beam is laterally surrounded by a low-dose envelope. Following the terminology of Gottschalk et al. (2015), the core describes the bulk energy deposited by primary protons, the surrounding halo is caused by charged secondaries originating from hard nuclear interactions, and the aura is the very low dose bath caused by neutral secondaries. In practical beam lines the low-dose envelope may be further complicated by dose due to scattering from upstream apparatus, known as spray. Unlike the halo and aura, which are due to basic physics, spray is reducible in principle, though this can prove difficult to achieve in practice. The halo, aura and spray remove energy (integrated dose) from the core. Neglecting this redistribution of dose can lead to absolute dose errors of order 10% (Pedroni et al. 2005), though the exact result depends strongly upon field size, spot weighting, energy and depth.

To correctly calculate absolute dose in PBS, as required when commissioning a treatment planning system (TPS), it is necessary to model the halo, and spray if present, with sufficient accuracy. Although the TPS model may be entirely experiment-based (Pedroni et al. 2005), Monte Carlo (MC) simulations commonly play a large role (Soukup et al. 2005, Parodi et al. 2013). MC particle transport codes, such as Geant4 (Agostinelli et al. 2003), FLUKA (Ferrari et al. 2005) and MCNPX (Waters 2002) are regularly used as a gold standard against which TPSs are validated (Clasie et al. 2012, Grevillot et al. 2012, Grassberger et al. 2015).

It is therefore important to validate the nuclear interaction models employed by these MC codes in the clinical incident energy regime 70–250 MeV. Recently, a comprehensive absolute measurement of the low-dose region (Gottschalk et al. 2015, hereinafter Go15) explored its complete radial and depth dependence at 177 MeV. We propose to test the nuclear models employed by Geant4 against this dataset, 295 values of log10(D/(MeV/g/p)), whose range of radii, depths and dose variation (five orders of magnitude) affords a test of unprecedented simplicity and scope.

To justify that claim we discuss some competing studies, limiting ourselves to ones whose signal is largely or entirely from nuclear interactions. That excludes studies such as Kurosu et al. (2014) which concern percent depth-dose (PDD) distributions in broad beams, where the contribution of nuclear reactions is of order 10% (Pedroni et al. 2005).

Experiments with multi-layer Faraday cups (Gottschalk et al. 1999, Paganetti and Gottschalk 2003) satisfy our criterion. They are absolute, and the nuclear and EM signals are completely separated. However, they only measure the range distribution of charged secondaries, a 1D test with transverse information integrated out. And they measure the number of charged secondaries rather than the dose.

More directly comparable to Go15 are measurements of the transverse dose profile which, beyond the core, is attributable either to nuclear secondaries or spray. Sawakuchi et al. (2010a) measured transverse profiles in air at isocenter and at selected depths in water at 72.5, 148.8 and 221.8 MeV. They used a small cylindrical ion chamber as well as optically stimulated luminescent detectors and radiochromic film. Their data were later compared with MCNPX using the Bertini cascade model followed by the preequilibrium model (Sawakuchi et al. 2010b). Good agreement of the shape of the MC simulation with experiment was found, over three decades of dose, for the transverse profiles in water at midrange.
More recently Lin et al (2014) measured transverse profiles at six energies between 100 and 225 MeV using radiochromic film at selected depths in a Solidwater phantom. Their IBA dedicated PBS nozzle was not modelled. Rather, the beam at the phantom entrance was approximated by adding to the core two subsidiary Gaussians whose energy dependent widths and amplitudes were fitted to in-air profile measurements. Simulations were performed with TOPAS using the default physics list and two options for the Geant4 EM model. Limiting ourselves to aspects relevant here, one can conclude from their figure 4 and text that TOPAS agrees well with the shape of their measured 2D transverse dose distribution except for the 0.01% isodose at shallow depths where the double Gaussian approximation to spray is inadequate.

In short, nuclear models in MCNPX and Geant4 have already been tested to the extent that any significant failure to predict the shape of selected transverse dose profiles would have been detected. They also confirm the expected cylindrical symmetry of the profiles wherever nuclear interactions dominate over spray. Go15, however, provides a test which is both simpler and more comprehensive. Measurements are in water, avoiding issues of the water equivalence of Solidwater. They employ a thimble chamber whose absolute calibration is $^{60}$Co-based as in clinical practice, avoiding issues that arise when using radiochromic film for absolute dosimetry. Measurements span the entire low-dose volume, instead of selected depths. Most important, the absolute dose per incident proton is reported. These data are therefore uniquely suited to answering the fundamental yet simple question: does the MC under test correctly predict the dose per incident proton in the region where that dose is dominated by nuclear interactions?

2. Methods

2.1. The experimental dataset

The two-dimensional dose distribution of a 177 MeV proton pencil beam stopping in water was measured using the experimental beam line at Massachusetts General Hospital (MGH). The experiment is described in Gottschalk et al (2015) and further detailed in Gottschalk et al (2014).

The experimental setup comprised a proton beam directed at a water phantom. The beam exited a 6.35 cm diameter beam pipe through a 0.008 cm Kapton® vacuum window. It then traversed a 150 cm helium bag and a 30 cm air gap before entering the water through a 0.86 cm PMMA tank wall (1.02 cm water-equivalent thickness). An air-filled plane-parallel ionisation chamber containing a total thickness of 0.0064 cm of aluminium was placed on the upstream surface of the tank wall, acting as a beam monitor. An Exradin T1 thimble ionisation chamber measured the dose at various positions within the phantom. The design minimised spray since the beam pipe was significantly larger than the beam width and only homogeneous materials were traversed.

In order to make absolute measurements of dose, the beam monitor was calibrated using a Faraday cup and the thimble chamber was calibrated by an Accredited Dosimetry Calibration Laboratory. Together, these could precisely measure a distribution of dose per proton. By recording depth dose curves at different radial distances from the beam axis (rather than lateral dose profiles at different depths) the signal range was limited. Then the beam fluence could be adjusted at each radial distance, in order to optimise the sensitivity of the experiment. This yielded a dose uncertainty of 3% and a positional uncertainty of 1 mm.

The model dependent fit of Go15 to their data yielded (among other things) properties of the incident beam at the tank entrance: beam range $d_{90} = 21.32 \pm 0.13$ cm and beam width...
\[ \sigma_z = 5.9 \pm 2.8 \text{ mm}^4. \] The angular spread (beam divergence), emittance and width of the Bragg peak were so poorly constrained as to be essentially unmeasured. Using radiochromic film, the beam was found to be cylindrically symmetric (\( \sigma_x = \sigma_y \)) to within a few percent.

2.2. Monte Carlo simulation

We simulated the two-dimensional dose distribution using Geant4 version 10.1p02 (Agostinelli et al. 2003), which was the latest release at the time of writing. A total of \( 1 \times 10^8 \) proton histories were simulated.

The simulated geometry was simplified with respect to the experimental setup. The beam source was positioned on the edge of a large volume of water, with the tank wall replaced by its water-equivalent thickness. Upstream apparatus was neglected. The dose distribution was scored in a set of concentric tubes centred on the beam axis, in order to exploit the symmetry of the simulation. The depth and radial bin widths were 1 mm and 4 mm respectively, to match the dimensions of the ionisation chamber. Since the beam source was placed at the surface of the phantom, the absolute dose was simply computed by dividing the scored dose by the number of proton histories simulated.

In the experimental paper, a beam energy of 177 MeV was determined via Janni’s range-energy tables (Janni 1982). However, the mean ionisation energy of water employed by default in Geant4 is taken from revised ICRU-73 tables (Sigmund et al. 2009), which provide shorter ranges than Janni’s tables. Therefore the beam energy was tuned to 177.9 MeV, where the simulated range agreed with experiment. It has been shown in a Monte Carlo study (Peeler and Titt 2012) that the dose distribution at large radial distances is unrelated to the size of the core. The width of the Gaussian beam energy distribution is calculated using an experiment-based parametrisation derived specifically for the MGH beam line (Clasie et al. 2012),

\[
\frac{\Delta E}{E} = (-8.650 \times 10^5 + 51.869 \times 10^3 \cdot d_{80} + 9337.7444) \times 10^{-6}
\]

where \( d_{80} \) is the beam range. This yields a relative energy spread of 0.65%.

The lateral and angular distributions of the beam were described by Gaussian distributions, with standard deviations \( \sigma_x \) and \( \sigma_\theta \) respectively. Although these parameters were estimated by the fitting procedure in the experimental paper, they were poorly constrained. In the simulation, they were tuned to \( \sigma_x = 5.4 \text{ mm} \) and \( \sigma_\theta = 18.4 \text{ mrad} \) in order to yield the best description of the experimental core data. Positional and angular distributions were assumed uncorrelated.

Appropriate Geant4 physics settings for proton therapy applications have been identified in version 8.1p01 by Jarlskog and Paganetti (2008) and in version 9.2 by Grevillot et al. (2010). In this study, similar physics modules were selected in Geant4 version 10.1p02. Electromagnetic interactions were described by the standard model (G4EmStandardPhysics_option4), elastic nuclear interactions were described by the standard hadronic elastic physics model (G4HadronElasticPhysics), and inelastic nuclear interactions were described by a binary cascade model followed by a pre-compound model of nuclear de-excitation (G4HadronPhysicsQGSP_BIC_HP). Additionally, the G4EmExtraPhysics, G4StoppingPhysics, G4IonBinaryCascadePhysics

Errors taken directly from the fitting program. Even the detailed paper (Gottschalk et al. 2014) did not give estimated parameter errors as the authors considered them unreliable in such a complicated fit. Of course, a targeted experiment could more precisely determine the beam parameters.

If this assumption were invalid, the model-dependent fit in Go15 would have yielded a statistically significant emittance.
and G4NeutronTrackingCut physics modules were used. The range threshold for secondary particle production was set to 0.1 mm.

3. Results

Figure 1 presents ten absolute depth dose curves produced by a 177 MeV proton pencil beam stopping in water. These were recorded at radial distances of 0, 1, 2, 3, 4, 5, 6, 7, 8 and 10 cm from the central beam axis, with greater radial distance corresponding to lower dose. Emphasis is given to the distal data in figure 2. The data are also displayed on a linear dose scale in figure 3. Lateral dose profiles at depths of 12 cm (mid-range) and 21 cm (near end of range) are shown in figure 4. In analytical TPSs, these lateral distributions are usually described by a double Gaussian (Li et al 2012).

Disagreement between simulated and experimental depth dose curves is quantified using a figure-of-merit. This is the root-mean-square dose deviation normalised to the maximum measured dose for the depth dose curve under consideration,

$$f = \frac{\sqrt{\langle (D_{\text{data}} - D_{\text{MC}})^2 \rangle}}{\text{max}(D_{\text{data}})}$$

(2)

where the calculation includes only depths at which an experimental measurement $D_{\text{data}}$ is available. At each depth, the simulated dose $D_{\text{MC}}$ of the nearest voxel was used. This is displayed in figure 3 for each depth dose curve.

4. Discussion

Before addressing issues evident in the results, we note that overall agreement between Geant4 and experiment is striking, especially considering that neither dataset is normalised.
Figure 2. Zoomed view of the distal data of figure 1. Data along the beam axis ($r = 0$ cm) are not shown as no experimental measurements were made in this region.

Figure 3. The data of figure 1 in linear presentation. Each panel is separately normalised to the maximum experimental dose, but the relation between simulation (line) and experiment (circles) is unaltered and absolute. The figure-of-merit $f$ (see text) quantifies the disagreement, expressed in percent.
By considering the agreement in different regions of the dose distribution, it is possible to verify the modelling of each component: core, halo, aura and spray.

4.1. Core

Agreement in the core (radii 0 and 1 cm), which tests the Geant4 EM model, is almost exact. This is unsurprising since the incident beam parameters were optimised to fit the experimental data in this region (enabling the Geant4 nuclear model to be tested at larger radii).

4.2. Halo

The halo involves two distinct physics mechanisms. Coherent scattering (the proton interacts with the nucleus as a whole) gives rise to a Bragg peak, a hint of which is seen as far out as \( r = 8 \) cm. Incoherent scattering (the proton interacts with constituents of the nucleus) produces the mid-range bump which appears at \( r = 4 \) cm and persists beyond \( r = 10 \) cm. Figure 3 demonstrates that Geant4 accurately describes both the shape and normalisation of these features, with notable exceptions at \( r = 3 \) and 6 cm.

Gottschalk et al (2015) already commented on the \( r = 6 \) cm scan being an outlier, evidence of positioning errors of order \( \pm 1 \) mm in the experiment. The much larger disagreement at \( r = 3 \) cm was unexpected and would require a radial positioning error of 3 mm to achieve agreement. On the other hand, it seems unlikely that Geant4 would mismodel these two features, of very different physical origin, by nearly the same factor. This question would be resolved by improved positioning precision and intervening scans out to \( r = 5 \) cm.

The halo produced by other Geant4 nuclear models was also investigated, though the binary cascade model clearly offered the best description. The Bertini cascade model was found to overestimate incoherent scattering at intermediate radii (increasing the figure-of-merit \( f \) by factors up to 3). The Liege intranuclear cascade model was found to underestimate incoherent scattering at large radii (increasing \( f \) by factors up to 5). The choice of de-excitation model was not found to have a significant effect.
4.3. Aura

Both experiment and simulation exhibit a clear plateau just downstream of the distal edge of the Bragg peak (see figures 1 and 2). This flat region is attributed to neutral secondaries (neutrons and γ-rays). Good agreement is observed, though Geant4 is generally higher than experiment by ∼10%. This mismodelling is reduced at \( r = 3 \) cm, lending further support to the positioning error hypothesis described in section 4.2. This is a qualitatively new test of MC neutron production.

MC neutron production is, of course, an ingredient in every shielding study, but there it is complicated by neutron transport through complex geometries of various intervening materials. Factor-of-two accuracy is the norm, and is in fact acceptable from a practical viewpoint (Clasie et al 2010). In contrast, figure 2 is the most direct test possible of the neutron source model in Geant4, with minimal absorbing material, itself water. It shows that Geant4 performs well in both producing and transporting neutrons within water.

4.4. Spray

In contrast, Geant4 is significantly lower than experiment in the proximal region at intermediate and large radii (see figure 1). One immediately suspects spray, and this was investigated. We modelled the upstream apparatus described in section 2.1, and moved the MC beam source from the tank wall to the vacuum window. The beam parameters were reoptimised to fit the experimental core dose, but no significant improvement was observed in the proximal region. Either Geant4 is indeed wrong here, or there is an unmodelled source of spray, perhaps scattering of protons in the beampipe originating far upstream. This could be investigated using in-air transverse scans in the entrance region.

5. Conclusions

We have compared the full dose distribution of a 177 MeV proton beam stopping in water with Geant4 simulations and found absolute agreement over five orders of magnitude, typically within a few percent. The depth dose curves at \( r = 3 \) and 6 cm are outliers, with Geant4 lower than experiment by as much as 40%, though even here the shape is well modelled. The discrepancy at \( r = 3 \) cm could be due to experimental positioning error \( \approx 3 \) mm, unlikely but not impossible, or it could mean that Geant4 underpredicts secondaries at intermediate radii.

Agreement in the aura is also good, marking a new test of MC neutron production. In contrast, measurements at the very lowest entrance doses are substantially higher than Geant4. That suggests a broad unaccounted-for source of low-energy incident protons (spray), although it is conceivable instead that Geant4 underpredicts back- and transverse scattering.

Our findings call for an improved experiment with better positional accuracy, scans at more intermediate radii, and in-air scans in the entrance region to investigate spray. Eventually, measurements should also be made at other beam energies. That being said, with the exceptions just noted, this study validates the Geant4 nuclear model to a few percent absolute error at a typical proton therapy energy, justifying its use in commissioning PBS TPSs. It also shows that Geant4 is accurate enough to provide kernels modelling the halo for TPSs. It would be interesting to see whether other widely used MC codes, such as MCNPX and FLUKA, are as accurate.
Acknowledgments

The authors thank Dr J Daartz and Dr B Clasie for their description of the MGH experimental beam line and Dr J Schumann for comments that greatly improved the manuscript. DH and HP wish to acknowledge support from National Cancer Institute grant U19CA21239. BG thanks the Physics Department of Harvard University for continuing support. AM thanks her PhD supervisor Dr W Sauerwein and the Department of Radiation Oncology at University of Duisburg-Essen.

References

Agostinelli S et al 2003 Geant4—a simulation toolkit Nucl. Instrum. Methods A 506 250–303
Clasie B, Depauw N, Fransen M, Gomà C, Panahandeh H R, Seco J, Flanz J B and Kooy H M 2012 Golden beam data for proton pencil-beam scanning Phys. Med. Biol. 57 1147–58
Clasie B, Wroe A, Kooy H, Depauw N, Flanz J, Paganetti H and Rosenfeld A 2010 Assessment of out-of-field absorbed dose and equivalent dose in proton fields Med. Phys. 37 311–21
Ferrari A, Sala P R, Fasso A and Ranft J 2005 FLUKA: a multi-particle transport code Technical Report CERN-2005-010, SLAC-R-773, INFN-TC-05-11
Gottschalk B, Cascii E W, Daartz J and Wagner M S 2014 Nuclear halo of a 177 MeV proton beam in water: theory, measurement and parameterization (arXiv: 1409.1938)
Gottschalk B, Cascii E W, Daartz J and Wagner M S 2015 On the nuclear halo of a proton pencil beam stopping in water Phys. Med. Biol. 60 5627–54
Gottschalk B, Platais R and Paganetti H 1999 Nuclear interactions of 160 MeV protons stopping in copper: a test of Monte Carlo nuclear models Med. Phys. 26 2597–601
Grassberger C, Lomax A and Paganetti H 2015 Characterizing a proton beam scanning system for Monte Carlo dose calculation in patients Phys. Med. Biol. 60 633–45
Grevillot L, Bertrand D, Despy F, Freund N and Sarrut D 2012 GATE as a GEANT4-based Monte Carlo platform for the evaluation of proton pencil beam scanning treatment plans Phys. Med. Biol. 57 4223–44
Grevillot L, Frisson T, Zahra N, Bertrand D, Stichelbaut F, Freud N and Sarrut D 2010 Optimization of GEANT4 settings for proton pencil beam scanning simulations using GATE Nucl. Instrum. Methods B 268 3295–305
Janni J F 1982 Proton range-energy tables, 1–10 GeV At. Data Nucl. Data Tables 27 147–339
Jarlskog C Z and Paganetti H 2008 Physics settings for using the Geant4 toolkit in proton therapy IEEE Trans. Nucl. Sci. 55 1018–25
Kuros K, Takashima M, Koizumi M, Das I J and Moskvin V P 2014 Optimization of GATE and PHITS Monte Carlo code parameters for uniform scanning proton beam based on simulation with FLUKA general-purpose code Nucl. Instrum. Methods B 336 45–54
Li Y, Zhu R X, Sahoo N, Anand A and Zhang X 2012 Beyond gaussians: a study of single-spot modeling for scanning proton dose calculation Phys. Med. Biol. 57 983–97
Lin L, Kang M, Solberg T D, Ainsley C G and McDonough J F 2014 Experimentally validated pencil beam scanning source model in TOPAS Phys. Med. Biol. 59 6859–73
Paganetti H and Gottschalk B 2003 Test of GEANT3 and GEANT4 nuclear models for 160 MeV protons stopping in CH2 Med. Phys. 30 1926–31
Parodi K, Mairani A and Sommerer F 2013 Monte Carlo-based parametrization of the lateral dose spread for clinical treatment planning of scanned proton and carbon ion beams J. Radiat. Res. 54 91–6
Pedroni E, Scheib S, Börhringer T, Coray A, Grossmann M, Lin S and Lomax A 2005 Experimental characterization and physical modelling of the dose distribution of scanned proton pencil beams Phys. Med. Biol. 50 541–61
Peeler C R and Titt U 2012 Monte Carlo study of radial energy deposition from primary and secondary particles for narrow and large proton beamlet source models Phys. Med. Biol. 57 3785–92
Sawakuchi G O, Zhu X R, Poenisch F, Suzuki K, Ciangaru G, Titt U, Anand A, Mohan R, Gillin M T and Sahoo N 2010a Experimental characterization of the low-dose envelope of spot scanning proton beams Phys. Med. Biol. 55 3467–78
Sawakuchi G O, Titt U, Mirkovic D, Ciangaru G, Zhu X R, Sahoo N, Gillin M T and Mohan R 2010b Monte Carlo investigation of the low-dose envelope from scanned proton pencil beams Phys. Med. Biol. 55 711–21
Sigmund P, Schinner A and Paul H 2009 Errata and addenda for ICRU Report 73, stopping of ions heavier than helium J. ICRU 5 1
Soukup M, Fippel M and Alber M 2005 A pencil beam algorithm for intensity modulated proton therapy derived from Monte Carlo simulations Phys. Med. Biol. 50 5089–104
Waters L S 2002 MCNPX User’s Manual Technical Report LA-UR-02-2607