



HAL
open science

Evaluation of GATE-RTion (GATE/Geant4) Monte Carlo simulation settings for proton pencil beam scanning quality assurance

Carla Winterhalter, Michael Taylor, David Boersma, Alessio Elia, Susanna Guatelli, Ranald Mackay, Karen Kirkby, Lydia Maigne, Vladimir Ivanchenko, Andreas F Resch, et al.

► To cite this version:

Carla Winterhalter, Michael Taylor, David Boersma, Alessio Elia, Susanna Guatelli, et al.. Evaluation of GATE-RTion (GATE/Geant4) Monte Carlo simulation settings for proton pencil beam scanning quality assurance. *Medical Physics*, 2020, 47 (11), pp.5817-5828. 10.1002/mp.14481 . hal-03039937

HAL Id: hal-03039937

<https://hal.science/hal-03039937>

Submitted on 4 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Evaluation of GATE-RTion (GATE/Geant4) Monte Carlo simulation settings for proton pencil beam scanning quality assurance

Carla Winterhalter^{†a)} and Michael Taylor

*Division of Cancer Sciences, University of Manchester, Manchester M13 9PL, UK
The Christie NHS Foundation Trust, Manchester M20 4BX, UK*

David Boersma

*ACMIT GmbH, Viktor Kaplan-Straße 2, Wiener Neustadt A-2700, Austria
EBG MedAustron GmbH, Marie Curie-Straße 5, Wiener Neustadt A-2700, Austria*

Alessio Elia

EBG MedAustron GmbH, Marie Curie-Straße 5, Wiener Neustadt A-2700, Austria

Susanna Guatelli

Centre For Medical Radiation Physics, University of Wollongong, Wollongong, Australia

Ranald Mackay and Karen Kirkby

*Division of Cancer Sciences, University of Manchester, Manchester M13 9PL, UK
The Christie NHS Foundation Trust, Manchester M20 4BX, UK*

Lydia Maigne

Laboratoire de Physique de Clermont, UMR 6533 CNRS - University Clermont Auvergne, Aubière, France

Vladimir Ivanchenko

*CERN, Geneva 23 1211, Switzerland
Tomsk State University, Tomsk 634050, Russia*

Andreas F. Resch

Department of Radiation Oncology, Medical University of Vienna, Währinger Gürtel 18-20, Vienna 1090, Austria

David Sarrut

Université de Lyon, CREATIS, CNRS UMR5220, Inserm U1044, INSA-Lyon, Université Lyon 1, Centre Léon Bérard, Lyon, France

Peter Sitch

The Christie NHS Foundation Trust, Manchester M20 4BX, UK

Marie Vidal

Institut Méditerranéen de Protonthérapie - Centre Antoine Lacassagne - Fédération Claude Lalanne, Nice 06200, France

Loïc Grevillot*

EBG MedAustron GmbH, Marie Curie-Straße 5, Wiener Neustadt A-2700, Austria

Adam Aitkenhead*

*Division of Cancer Sciences, University of Manchester, Manchester M13 9PL, UK
The Christie NHS Foundation Trust, Manchester M20 4BX, UK*

(Received 20 April 2020; revised 28 August 2020; accepted for publication 29 August 2020; published 17 October 2020)

Purpose: Geant4 is a multi-purpose Monte Carlo simulation tool for modeling particle transport in matter. It provides a wide range of settings, which the user may optimize for their specific application. This study investigates GATE/Geant4 parameter settings for proton pencil beam scanning therapy.

Methods: GATE8.1/Geant4.10.3.p03 (matching the versions used in GATE-RTion1.0) simulations were performed with a set of prebuilt Geant4 *physics lists* (QGSP_BIC, QGSP_BIC_EMY, QGSP_BIC_EMZ, QGSP_BIC_HP_EMZ), using 0.1mm-10mm as *production cuts* on secondary particles (electrons, photons, positrons) and varying the *maximum step size* of protons (0.1mm, 1mm, none). The results of the simulations were compared to measurement data taken during clinical patient specific quality assurance at The Christie NHS Foundation Trust pencil beam scanning proton therapy facility. Additionally, the influence of simulation settings was quantified in a realistic patient anatomy based on computer tomography (CT) scans.

Results: When comparing the different *physics lists*, only the results (ranges in water) obtained with QGSP_BIC (G4EMStandardPhysics_Option0) depend on the maximum step size. There is clinically negligible difference in the target region when using High Precision neutron models (HP) for dose calculations. The EMZ electromagnetic constructor provides a closer agreement (within 0.35 mm) to

measured beam sizes in air, but yields up to 20% longer execution times compared to the EMY electromagnetic constructor (maximum beam size difference 0.79 mm). The impact of this on patient-specific quality assurance simulations is clinically negligible, with a 97% average 2%/2 mm gamma pass rate for both *physics lists*. However, when considering the CT-based patient model, dose deviations up to 2.4% are observed. *Production cuts* do not substantially influence dosimetric results in solid water, but lead to dose differences of up to 4.1% in the patient CT. Small (compared to voxel size) *production cuts* increase execution times by factors of 5 (solid water) and 2 (patient CT).

Conclusions: Taking both efficiency and dose accuracy into account and considering voxel sizes with 2 mm linear size, the authors recommend the following Geant4 settings to simulate patient specific quality assurance measurements: No *step limiter* on proton tracks; *production cuts* of 1 mm for electrons, photons and positrons (in the phantom and range-shifter) and 10 mm (world); best agreement to measurement data was found for QGSP_BIC_EMZ *reference physics list* at the cost of 20% increased execution times compared to QGSP_BIC_EMY.

For simulations considering the patient CT model, the following settings are recommended: No *step limiter* on proton tracks; *production cuts* of 1 mm for electrons, photons and positrons (phantom/range-shifter) and 10 mm (world) if the goal is to achieve sufficient dosimetric accuracy to ensure that a plan is clinically safe; or 0.1 mm (phantom/range-shifter) and 1 mm (world) if higher dosimetric accuracy is needed (increasing execution times by a factor of 2); most accurate results expected for QGSP_BIC_EMZ *reference physics list*, at the cost of 10–20% increased execution times compared to QGSP_BIC_EMY. © 2020 The Authors. *Medical Physics* published by Wiley Periodicals LLC on behalf of American Association of Physicists in Medicine. [<https://doi.org/10.1002/mp.14481>]

Key words: dose calculation, Geant4, GATE-RTion, Monte Carlo, proton pencil beam scanning therapy

1. INTRODUCTION

Geant4 is a C++ based, object-oriented, multi-purpose Monte Carlo (MC) tool to simulate particle transport for energies ranging from eV to TeV scale,^{1,2} with applications ranging from the modeling of high energy particle colliders to space and shielding simulations.³ Geant4 is extensively used for medical physics simulations, in particular to calculate the dose (energy deposited per unit mass) during particle therapy (e.g., see Refs. [4–6] for proton radiotherapy simulations using Geant4 or Geant4 wrappers). This study is focused on GATE/Geant4 simulations for patient dosimetry in proton pencil beam scanning.

Geant4 based Architecture for Medicine Oriented Simulation (GAMOS),⁷ Tool for Particle Simulation (TOPAS)⁸ and Geant4 Application for Tomographic Emission (GATE)^{9–11} provide user-friendly interfaces to Geant4. This research takes place within a recent initiative called GATE-RTion,¹² which is aimed at providing a validated long-term application, based on a GATE and Geant4 version, along with associated software tools for clinical dosimetric implementation to facilitate collaborations between hadron therapy institutes. Current collaborators include the Centre Antoine Lacassagne (Nice, France), the Christie NHS Foundation Trust (Manchester, UK) and MedAustron (Wiener Neustadt, Austria).

GATE/Geant4 offers a set of prebuilt *physics lists* which the user may adopt in his/her specific simulation scenario. *Physics lists* may use different models to describe a specific electromagnetic or hadronic physics interaction and different parameters such as the *production cut* and *maximum step size*. In order to support the Geant4 medical physics

community, the Geant4 Medical Physics Benchmarking Group^{13,14} has been recently established to benchmark Geant4 for medical physics and to provide recommendations in terms of *physics lists*, *production cuts* and *maximum step sizes*. For proton therapy (using proton energies up to 250 MeV), these settings have been previously investigated for different Geant4 versions: Jarlskog and Paganetti compared simulations (Geant4 v.8.1) in water and a Faraday cup geometry to measurements, and recommended a combination of electromagnetic and nuclear models.¹⁵ That recommendation was used as a basis for Grevillot et al., who compared GATE/Geant4 (Geant4 v.9.2) simulations in water and PMMA to measurements and showed the importance of the *maximum step size* and *production cut*.¹⁶ Kurosu et al. simulated proton treatment nozzles for uniform¹⁷ and for spot scanning¹⁸ with different MC codes, including GATE/Geant4 v.9.5, investigating the influence of both *step size* and *production cut* on depth dose curves. Fuchs et al. investigated multi-Coulomb scattering after a range of materials (GATE, Geant4 versions 9.5.02, 9.6.03, 10.0.02, 10.1, 10.2.02), concluding that agreement to measurements substantially depended on the multiple scattering model set in the specific Geant4 version used.¹⁹ Recently, Resch et al. compared simulated lateral dose profiles and field size factors to measurements to investigate the accuracy of electromagnetic and nuclear scattering models (GATE, Geant4 v.10.03.01), showing a high sensitivity of the elastic scattering cross sections.²⁰

This study extends and complements the work above by comparing a wide range of combinations of *physics options*, *step limiter* and *production cuts* to a comprehensive set of measurements, aimed at clinical applications of Geant4/

GATE for proton therapy dosimetry. It takes the whole (clinical) process of beam characterization into account and all experimental data were acquired as part of clinical commissioning or quality assurance. Additionally, the influence of different settings is quantified in simulations performed in a realistic CT-based patient model. This is therefore a comprehensive validation of Geant4 v.10.3.3 in combination with GATE v.8.1, and therefore also of GATE-RTion v.1.0, which is a long-term stable branch of GATE v.8.1 intended for clinical purposes. A set of parameter recommendations is presented, taking both dose distributions and calculation efficiency into account.

The first part of the paper gives an overview of the simulation setup and the Geant4 settings under investigation. Next, the influence of the different settings on the beam-modeling process and on spot sizes in air is investigated by comparison of simulations of individual pencil beams and measurement data acquired during the commissioning of a proton therapy facility. Then the effect of the settings on simulations in solid water is evaluated in comparison to patient specific quality assurance (PSQA) measurements. Finally, the effect of the MC settings on calculations in the CT-based patient model is assessed for a set of patients, which represent a range of potential dose calculation issues.

2. MATERIALS AND METHODS

2.A. AUTOMC simulations

In the following, MC settings are summarized according to Ref. [21]. A set of in-house developed scripts (AUTOMC²²) written in GNU Octave²³ (v.4.4.1), which automatically create GATE macro files, launch simulations and analyse simulation results, have been developed for clinical proton therapy dose calculations at The Christie NHS Foundation Trust. This system performs all its MC calculations using GATE v.8.1/ Geant4 v.10.3.3.² These are the same versions included in GATE-RTion v.1.0,¹² and in keeping with the goals of GATE-RTion, this work aims to provide recommendations in terms of simulation parameters to adopt for proton pencil beam scanning.

The MC settings for simulations performed in this work are summarized in Table I. Simulations and data analysis were performed on a Linux (based on Ubuntu 18.04 LTS) cluster consisting of 1 master CPU (for data analysis) and 10 slave CPUs (for simulations). Each CPU was a quad-core Intel Xeon E3-1240 v6 @ 3.70 GHz, with 16 GB RAM, leading to a total of 40 cores. No variance reduction techniques were used.

2.A.1. Beam-model tuning

For proton pencil beam scanning MC calculations, protons need not be tracked through the whole beam line. Instead a beam description upstream of the patient can be implemented,^{25–27} which parameterizes each spot by its number of protons (for absolute dose), mean energy and energy spread

(shape of the Bragg peak) and beam optics phase-space (beam sizes in air) at the “starting point” of the simulation (in AUTOMC: 89.5 cm upstream of iso-center). Mean energy and energy spread are automatically and iteratively adjusted such that depth dose curves measured with a Bragg peak chamber (diameter 84 mm) are reproduced (tuning process, see Ref. [22] for more details). For this study, error bars on the mean energy were determined by repeating this tuning process with different initial energies. That is, for one setting, the tuning process was performed once with initial energies above the expected mean energy and once with initial energies below the expected mean energy. The error bars show the maximum difference between the obtained final mean energies, and as such the error due to both statistics and the tolerances in the beam-modeling process.

2.A.2. Lateral spot profiles in air

Within a MC simulation, the beam optics phase-space consisting of spot size, emittance and divergence of the protons at the source plane defines the initial optics of the beam. Additionally, the Lexan “range-shifter,” which can be introduced into the beam to lower the proton energy, also influences spot size and divergence. Table I summarizes the range-shifter options modeled within this study. Beam sizes in air (sigma, modeled with the beam spot following a Gaussian distribution) were simulated with and without range-shifter and compared to beam sizes measured with the Lynx detector (IBA-dosimetry, Schwarzenbruck) at different distances from iso-center.

2.A.3. PSQA measurements

PSQA measurements are routinely performed for each radiation field before the first fraction is delivered to the patient (e.g., Ref. [28]). At The Christie NHS Foundation Trust, doses are measured at multiple depths in a Solid Water phantom using a the Octavius 1500 XDR array (PTW, Freiburg) for relative dose distributions and a PTW 31021 Semiflex 3D ion chamber for absolute dose values (see Ref. [22] for more details). Two example fields (one inhomogeneous field with range-shifter and one homogeneous field without range-shifter, patient 1 and patient 2 of Table II) were simulated to 0.25% statistical uncertainty at the 90–100% dose level (see Ref. [24]) to demonstrate differences between the different settings. 34 fields (six patients, all with range-shifter thicknesses 3 and 5 cm) were recalculated to 0.6% statistical uncertainty at the 90–100% dose level, and compared to 200 (relative) array planes and 74 (absolute) chamber measurement points, which were all taken as part of the clinical PSQA. Place of measurement was at the shallow/central/deep part of the high dose region (first three months of operation of The Christie NHS Foundation Trust proton beam therapy centre), shallow/deep part (next six months) and at the central part only (afterwards). Additionally, measurements were partly repeated in different treatment rooms (see Ref. [22]).

TABLE I. Overview of the AUTOMC simulation settings. First dimension is along the beam direction.

	Beam-model generation	Lateral spot size in air	PSQA phantom	Patient CT
Volume	400 mm × 100 mm × 100 mm, upstream surface at iso-center	0.1 mm × 101 mm × 101 mm, various positions between 40 cm up- and downstream of iso-center	According to phantom CT	According to patient CT
Material	Water (density = 1.0 g/cm ³ , I = 75.3 eV)	Air	solid water (density = 1.057 g/cm ³ , I = 67.1 eV)	Hounsfield units converted to elemental composition & density
Scoring voxel dimensions	0.5 mm × 1 mm × 1 mm, laterally integrated over Bragg peak chamber area.	0.1 mm × 1 mm × 1 mm	2 mm × 2 mm × 2 mm	2 mm × 2 mm × 2 mm
Initial beam	19 energies between 70 and 245 MeV, 4 · 10 ⁷ protons per spot	90 MeV, 150 MeV, 230 MeV, 2 · 10 ⁷ initial protons per spot	Energies according to patient plan. Number of protons scaled to reach 0.25%/0.6% statistical uncertainty at the 90–100% isodose ²⁴	Energies according to patient plan. Number of protons scaled to reach 0.25% statistical uncertainty at the 90–100% isodose ²¹
Lexan range-shifter	None	None/2 cm/3 cm/5 cm. Downstream surface located 46 cm upstream of iso-center	According to patient plan	According to patient plan

2.A.4. Patient CT model

Finally, dose distributions are simulated in the patient CT model for six different patients (1 field per patient, see Table II). The CT calibration was based on the stoichiometric calibration by Schneider et al (Phys Med Biol 29), and Ref. [22] in detail describes the procedure to ensure matching CT calibrations (based on same scanner parametrization, reference tissues, and ionization values) between treatment planning system and AUTOMC. Patients were chosen to represent a range of several potential dose calculation issues, that is, different range-shifter options, heterogeneous anatomical sites, and treatments with metal implants.

2.B. Geant4 physics and simulation settings

The following prebuilt Geant4 *physics lists*, which are considered suitable for hadron therapy, were investigated:

- QGSP_BIC: QGSP_BIC with G4EMStandardPhysics (often referred as G4EMStandardPhysics_Option0, Wentzel multiple scattering model for protons).
- QGSP_BIC_EMY: QGSP_BIC with G4EMStandardPhysics_Option3, Urban multiple scattering model for protons.

- QGSP_BIC_EMZ: QGSP_BIC with G4EMStandardPhysics_Option4, Wentzel multiple scattering model for protons.
- QGSP_BIC_HP_EMZ: Same electromagnetic interactions as QGSP_BIC_EMZ, with high precision (HP) neutron libraries for neutrons below 20 MeV.

The *physics lists* differ in terms of electromagnetic and hadronic physics models and processes. Details and parameters can be found at.^{2,14,30,31} Multiple studies^{15,20} showed that the Binary Cascade model is adequate to model the intranuclear cascade in proton therapy,¹⁴ and therefore the QGSP_BIC was chosen to model hadronic interactions. The EM physics component was modeled using one among G4EMStandardPhysics_Option0, G4EMStandardPhysics_Option3 (EMY), and G4EMStandardPhysics_Option4 (EMZ). G4EMStandardPhysics_Option4 is deemed to be the most accurate EM physics constructor for medical physics applications, at the cost of longer computational times.¹⁴ QGSP_BIC_EMZ in combination with HP (QGSP_BIC_HP_EMZ) neutron data libraries was also studied. For readability, these lists are in the following abbreviated as Option0, EMY, EMZ, HP_EMZ.

The dosimetric effects of varying the *production cuts* (threshold on the production of secondary photons, electrons and positrons) and *step limiter* (set by the user, restricting the

TABLE II. Patient selection. 1 field has been recalculated per patient. Number of phases & fields, CTV volume and prescription are given for the whole treatment, range-shifter for the recalculated field

	Treatment site	Prescription [RBE.Gy] (#fractions)	Number of phases & Fields	CTV vol [cm ³]	Range-shifter	Comment
Patient 1	Neck	50.4 (28); 9 (5)	2 phases, 3 + 2 fields	308; 56	5 cm	—
Patient 2	Brainstem	52.2 (29)	1 phase, 2 fields	44	None	—
Patient 3	Base of skull	54 (30); 16.2 (9)	2 phases, 3 fields each	29; 7	5 cm	—
Patient 4	Paranasal	63; 70 (35)	Simultaneous integrated boost, 4 fields	47; 28	None	—
Patient 5	Spine	50.4 (28); 19.8 (11)	2 phases, 3 fields each	935; 37	None	Titanium implant
Patient 6	Spine	18 (25); 21.6 (12); 9 (5)	3 phases, 4 + 3 + 3 fields	654; 202; 9	2 cm	Titanium implant

maximum step length of protons) parameters were investigated as well. A maximum step length (*step limiter*) has been set for protons. The *step limiter* was set to 0.1, 1 mm or was not set. In the case of no user-defined step limiter, the step size is calculated in Geant4 with consideration of either the distance to the next geometrical boundary or the next interaction as dictated by the implemented physics processes. *Production cuts* set thresholds on the production of secondary particles. Only secondaries with a range exceeding the *production cut* are explicitly produced and tracked, otherwise their energy is considered local energy deposition. *Production cuts* were set for electrons, photons and positrons within the world (1 mm/10 mm) and within the phantom/range-shifter (0.1 mm/1 mm). These cuts were chosen to use one value comparable to calculation voxel size (see Table I), and another value much smaller than the calculation voxel size.

2.C. Overview of study

Simulation results were compared to experimental data for the combinations of *production cuts* and *step limiters* summarized in Table III. From here on, these will be referred to by the scenario numbers shown in the table. Each of these six scenarios was combined with the four prebuilt *physics lists* (Option0, EMY, EMZ, HP_EMZ) leading to a total of 24 combinations.

First, the effect of these settings was evaluated for the energy tuning process (Section 3.A) and for beam size simulations in air (Section 3.B). The effect of each *physics list* was analyzed while setting both *production cuts* and *step limiter* to the smallest values considered, which should provide the most accurate dosimetric results (reference scenario #1). Next, *production cuts* in phantom and world were varied while leaving the *step limiter* constant (scenarios 1, 2 and 3). Then, the effect of the *step limiter* was evaluated while keeping the cuts fixed at the smallest value considered (scenarios 1, 4 and 5). Finally, the scenario with smallest *cuts* and *step*

TABLE III. Overview of the combinations of production cuts and step limiter investigated in this study. In air, range cuts of 1 mm/10 mm correspond to production thresholds of 1 keV/1 keV (photons) 1 keV/16 keV (electrons) and 1 keV/16 keV (positrons). In water, range cuts of 0.1 mm/1 mm correspond to production thresholds of 1.1 keV/2.9 keV (photons), 85 keV/350 keV (electrons) and 83 keV/343 keV (positrons). The tracking cut for electrons is 0.1 keV.

Scenario	Production cuts for electrons, photons and positrons (phantom and range-shifter) [mm]	Production cuts for electrons, photons and positrons (world) [mm]	Step limiter [mm]
1 "Reference"	0.1	1	0.1
2	1	1	0.1
3	0.1	10	0.1
4	0.1	1	1
5	0.1	1	None
6	1	10	None

limiter was compared to the combinations of no *step limiter* with 0.1 mm/1 mm and 1 mm/10 mm *production cuts* (scenarios 1, 5 and 6).

Based on these results, the three *physics lists* EMY, EMZ and HP_EMZ in combination with scenario 5 ("small cuts" compared to 2 mm voxel linear size) and scenario 6 ("large cuts") were further investigated and compared to measurement data acquired with the PSQA phantom (Section 3.C), and compared when considering realistic, CT scan based patient models (Section 3.D).

3. RESULTS

3.A. Beam-model definition for the Monte Carlo simulations

For the full proton energy range, the energy fine tuning (i.e., the mean energy and energy spread defined in the MC simulation for each nominal TPS energy) does not substantially depend on the chosen *physics lists* [mean energy in Fig. 1(a), agreement within 0.08 MeV and energy spread (not shown), agreement within 0.03 MeV]. Subsequently, the tuning process was repeated for a 245 MeV (nominal energy) spot with different *production cuts* [Fig. 1(b), scenario 1–3], *step limiter* [Fig. 1(c), scenarios 1,4 and 5] and combination of *production cuts* and *step limiter* [Fig. 1(d), scenarios 1,5 and 6]. Simulations performed with the EMY and EMZ did not depend on the step limiter, whereas for simulations performed with Option0 energies varied by up to 0.5 MeV, indicating that for this physics option the range in water depends on the step limiter.

After the beam-modeling process, local dose differences between MC based calculations and depth dose measurement were analyzed for the 245 MeV spot [Fig. 2(a)]. There was no substantial difference between different *physics lists* [Fig. 2(b), scenario 1]. It is worth noting that for Option0 the differences are dependent on the *step limiter* [Fig. 2(c), scenario 5], and that for all *physics lists* the pattern of differences changes with larger *production cuts* [Fig. 2(d), scenario 6]. All observed differences between MC simulations and depth dose measurement lie within $\pm 3\%$.

The execution times for the 245 MeV spot decreased by factors of 4.5–5.9 when deactivating the *step limiter* (scenario 5 vs scenario 1) and by factors of 5.3–7.2 when increasing the *production cuts* (scenario 6 vs scenario 5). Such results are to be expected because smaller steps and a higher number of secondaries to track inevitably increases the execution time. Compared to HP_EMZ (scenarios 5/6), shorter execution times are achieved with Option0 (factors of 1.4/2.0), followed by EMY (factors of 1.2/1.5) and EMZ (factors of 1.0/1.1).

3.B. Lateral pencil beam profiles in air

For the 150 MeV beam without range-shifter, beam sizes in air (Fig. 3) agreed within 0.35 mm to measurements,

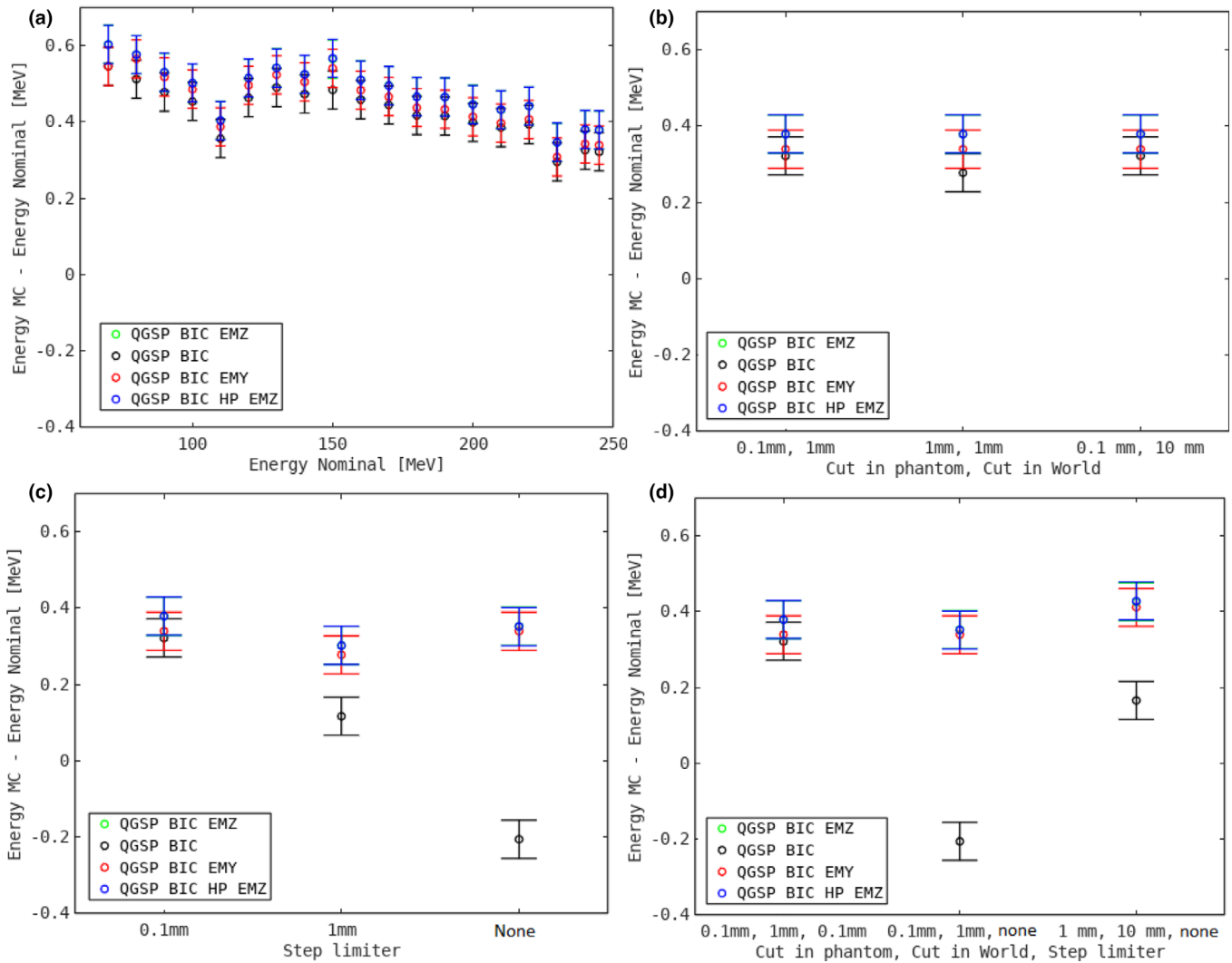


FIG. 1. Results for all energies and physics lists (a, scenario 1), and as a function of production cuts in phantom and world (b, scenario 1–3), as a function of step limiter (c, scenario 1,4,5) and as a function of combinations of cuts and step limiter (d, scenario 1,5,6) for a 245 MeV (nominal energy) spot. EnergyNominal is the energy reported by the TPS, EnergyMC is the mean energy used to define the source within Gate (i.e., EnergyMC is derived by the tuning process.) Error bars are determined by using different starting conditions (initial energy) in the energy tuning process. [Color figure can be viewed at wileyonlinelibrary.com]

showing that the beam optics have been matched well. With the 5 cm range-shifter, simulations performed with EMZ and HP_EMZ [green/blue lines in Figs. 3(a)–3(c)] agreed within 0.25 mm to measurements (magenta triangles), with negligible differences when using the HP module. EMY lead to smaller beam sizes compare to the other physics lists and measurements [up to 0.79 mm difference to measurements, red lines in Figs. 3(a)–3(c)]. As the beam optics have been defined without range-shifter, and the range-shifter been simulated as a physical component, these differences between different settings can be attributed to the differences in scattering in the range-shifter. Option0 simulated beam sizes without step limiter and with large cuts differed by up to 0.51 mm from measurements [black line in Fig. 3(c)]. Simulations performed with the 5 cm range-shifter and a very low energy proton beam (90 MeV) also showed less scattering when using EMY compared to EMZ/HP_EMZ [Figs. 3(d)–

3(e)]. For a high energy proton beam (230 MeV), less scattering occurs in the range-shifter and differences between the different physics lists was therefore negligible.

3.C. PSQA measurements

The previous two sections showed that dose results depend on the step limiter for Option0. As the step limiter substantially increases simulation times, for the remainder of the study, we concentrate on EMY, EMZ and HP_EMZ without a step limiter and in combination with “small” (0.1 mm/1 mm) and “large” (1 mm/10 mm) cuts (phantom/world, in comparison to the 2 mm calculation grid, scenarios 5 and 6).

Two example fields [patient 1, Figs. 4(a)–4(d) and patient 2, Figs. 4(e)–4(h)] fields were simulated in solid water to 0.25% statistical uncertainty at the 90–100% dose level to demonstrate the differences caused by the MC simulation

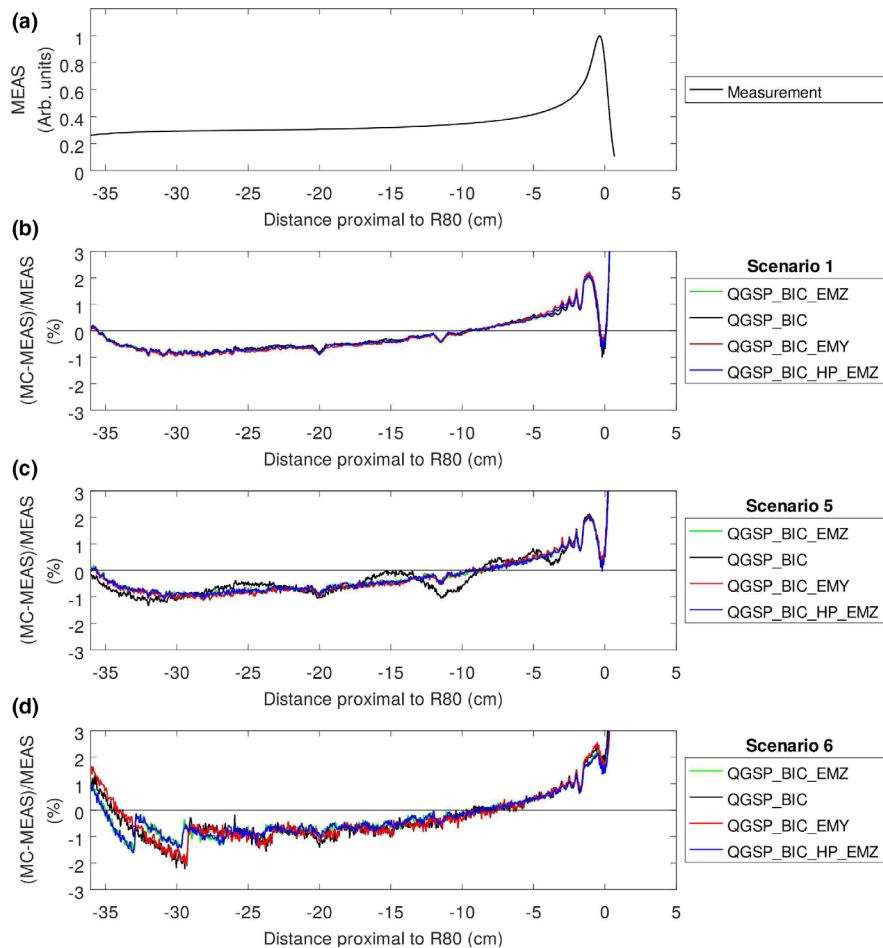


FIG. 2. Measured integral depth dose curve for the 245 MeV spot (a), and differences between MC and measurements after the tuning process for the 245 MeV spot for all physics lists with range cuts and step limiter defined according to scenario 1(b), scenario 5(c) and scenario 6(d). [Color figure can be viewed at wileyonlinelibrary.com]

settings. For patient 1, an inhomogeneous field with a 5 cm range-shifter, there was no substantial difference between EMZ and HP_EMZ [Fig. 4(b)]. EMZ and EMY simulated dose results differed by up to 1% of the prescription dose [Fig. 4(c)], whereas different *production cuts* did not substantially influence dose results [Fig. 4(d)]. For patient 2, a homogenous field without range-shifter, magnitudes of differences were comparable to statistical fluctuations [Figs. 4(e)–4(h)]. Execution times (Table IV) decreased by factors of up to 4.7 when changing from small to large *cuts*, and faster execution times were achieved with EMY (up to factors of 1.6/1.3 faster than HP_EMZ/EMZ).

Finally, 34 fields were recalculated in solid water to a 0.6% statistical uncertainty at the 90–100% dose level and compared to PSQA experimental data (see Section 2.A). Relative measurements [Fig. 5(a)] show a mean 2%/2 mm local gamma agreement (10% low dose threshold) over all 34 fields (200 planes) of 97% for all three scenarios (EMZ, small cuts; EMZ, large cuts; EMY, large cuts). Absolute dose measurements [74 points, Fig. 5(b)] show a dose offset of -1.2% (EMZ, small cuts), -1.0% (EMZ, large cuts) and -0.9% (EMY, large cuts).

3.D. Patient CT

In patient CT (Fig. 6), there was no substantial difference in terms of dose calculation between EMZ and HP_EMZ in the target region (Fig. 6, column 2). Differences between EMZ and EMY (Fig. 6, column 3) were substantial (within 2.4%), especially when using a range-shifter [Figs. 6(c), 6(k), 6(w)] and after tissue heterogeneities [e.g., implant in Fig. 6(s)]. *Production cuts* (column 4) cause dose differences of up to 4.1% at air-tissue interface [Fig. 6(l)]. Execution times were decreased by up to a factor of 2.2 when changing from small to large cuts, and faster execution times were achieved with EMY (up to factors of 1.4/1.2 faster than HP_EMZ/EMZ, Table IV).

4. DISCUSSION

Comparison of simulated dose distributions in water shows that for QGSP_BIC with G4EMStandardPhysics_Option0 the results depend on the step limiter. In the Option0 configuration, steps are larger than in EMY and EMZ. If the step limiter is reduced in Option0, step sizes

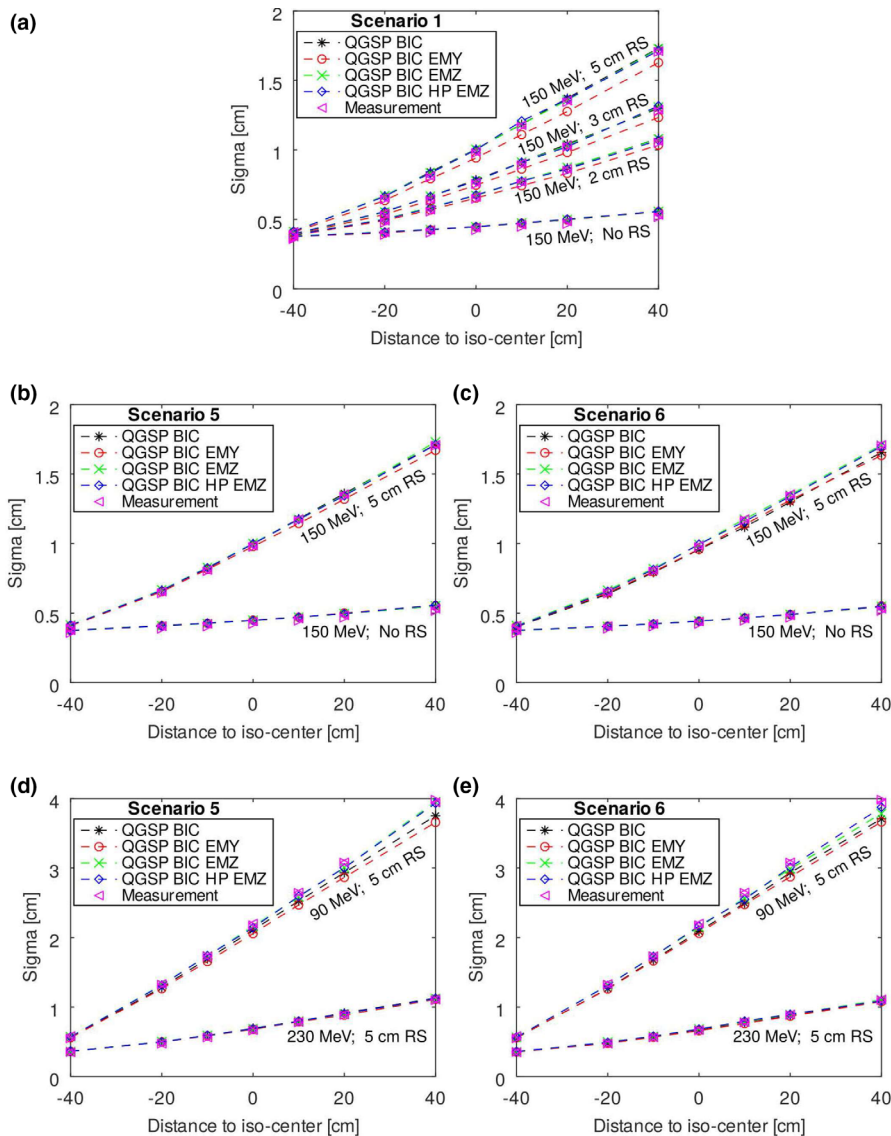


FIG. 3. Beam size (sigma of the Gaussian distribution of the beam spot) of a 150 MeV (a–c), a 90 MeV and a 230 MeV (d,e) spot as a function of distance from iso-center in air for all physics lists, (a) without range-shifter and with 2/3/5 cm Lexan range-shifter for small cuts and step limiter (scenario 1), (b,d) without and with a 5 cm range-shifter for small cuts and step limiter default (scenario 5) and (c,e) for large cuts and step limiter default (scenario 6). Measurements are represented with magenta triangles. [Color figure can be viewed at wileyonlinelibrary.com]

are therefore much more reduced compared to EMY and EMZ. As the step limiter substantially increases calculation times (factors of 4.5–5.9), EMY, EMZ or HP_EMZ without *step limiter* are therefore better suited for dose simulations in proton therapy.

For the endpoints of this study (clinical measurement data consisting of depth dose curves, beam size in air, dose distributions in a solid water phantom and patient CT), there was no substantial difference between HP_EMZ and EMZ in-field. This is expected since the in-field region is dominated by the incident proton beam, while a bigger effect of the neutron distribution is expected out-of-field and beyond the distal edge of the Bragg peak. Using the high precision neutron libraries, however, increased execution times by up to 30%.

It has been demonstrated that EMY (G4EMStandardPhysics_Option 3, Urban scattering model)

underestimates the beam sizes in air after a scattering material when compared to EMZ (G4EMStandardPhysics_Option 4, Wentzel scattering model) and to measurements. This is mainly due to differences in multiple scattering modeling, and confirms in Geant4.10.3.3 the results of Fuchs et al.,¹⁹ who showed that multi-Coulomb scattering angles were better reproduced when using the Wentzel instead of the Urban scattering model. Consequently, dose distributions simulated in a solid water phantom differ (within 1% of prescription dose) when comparing EMY and EMZ, with the EMY option resulting in less lateral spread of the fields, just as it resulted in less lateral spread of single spots. However, as this lies within the measurement uncertainty of clinical PSQA measurements, and the calculation discrepancies are significant only at very located areas of the dose map, no substantial difference between *physics lists* is observed when comparing

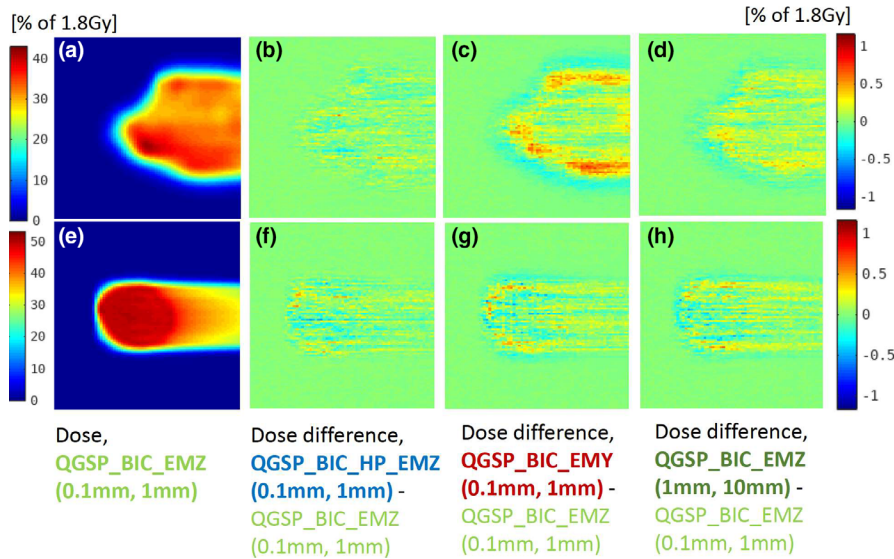


FIG. 4. Dose to water³² simulated with EMZ with 0.1 mm production cuts in the range-shifter/phantom and 1 mm production cuts in the world (scenario 5, a,e), dose differences between HP_EMZ and EMZ (b,f), dose differences between EMY and EMZ (c,g), and dose differences due to the settings of small (scenario 5, 0.1 mm/1 mm) or large (scenario 6, 1 mm/10 mm) production cuts (d,h), for two different patient fields (a–d and e–h). Results have been scaled to the prescribed fraction dose of 1.8 Gy. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE IV. Execution times for the fields displayed in Figs. 4(a)–4(d)/Figs. 6(a)–6(d) for each scenario (0.25% statistical uncertainty at the 90–100% dose level).

Physics list	Production cuts in phantom/world	Patient 1, solid water [CPU.h]	Patient 1, CT [CPU.h]
EMY	0.1 mm/1 mm	475.0	895.2
EMY	1 mm/10 mm	104.1	448.2
EMZ	0.1 mm/1 mm	605.8	1068.1
EMZ	1 mm/10 mm	127.9	491.0
HP_EMZ	0.1 mm/1 mm	644.9	1233.2
HP_EMZ	1 mm/10 mm	163.7	646.4

simulations to a wide range of relative array and absolute chamber measurements, with an average gamma index agreement (2%/2 mm) of 97% and absolute dose offset to

measurements of 1% (Fig. 5). Gamma analysis was chosen to evaluate the agreement to measurements in solid water as this is the standard clinical procedure to ensure treatment safety. When considering patient models based on CT acquisition, the same effect is observed, with dose differences of up to 2.4% between EMZ and EMY, which are especially pronounced when using a range-shifter. Execution times are up to 10–30% higher for EMZ when compared to EMY.

Production cuts influence the difference between simulated depth dose curves and measurements, and therefore have to be chosen carefully to ensure correct absolute dose scaling. For PSQA in solid water, production cuts from 0.1 to 1 mm do not substantially affect dose distributions and agreement to measurements, but increase execution times by up to a factor of 5. In contrast, in patient CT, production cuts cause dose differences of up to 4%, especially at air-tissue

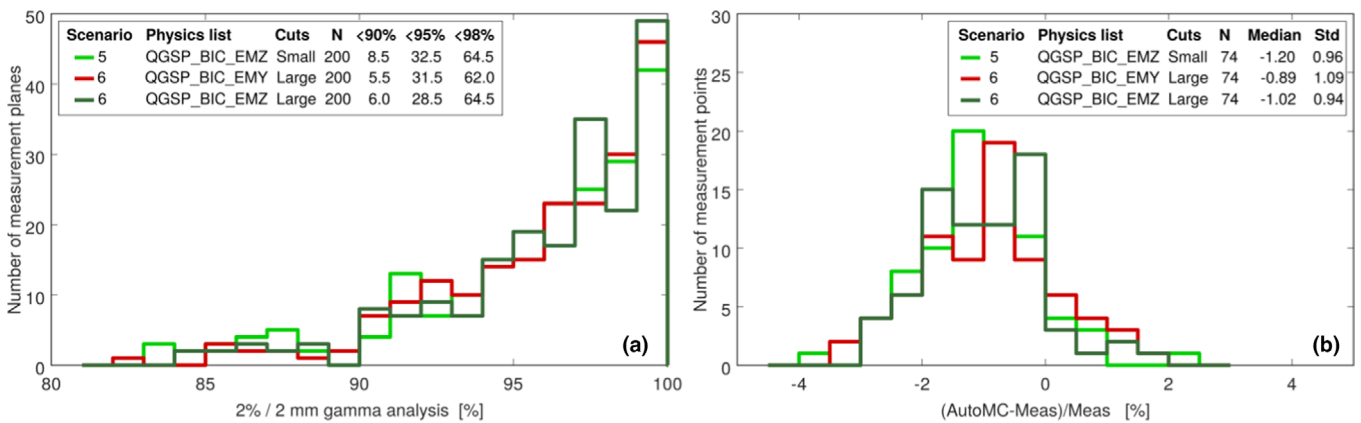


FIG. 5. Histograms of the gamma index values obtained when comparing simulations to relative measurements for 200 array planes (a) and histograms of absolute dose differences for 74 chamber measurements (b) for EMZ with small cuts (light green), EMZ with large cuts (dark green), EMY with large cuts (red). In (a), the percentage of planes which have <90%, <95% and <98% of pixels having $\gamma \leq 1$ are indicated in the figure legend. [Color figure can be viewed at wileyonlinelibrary.com]

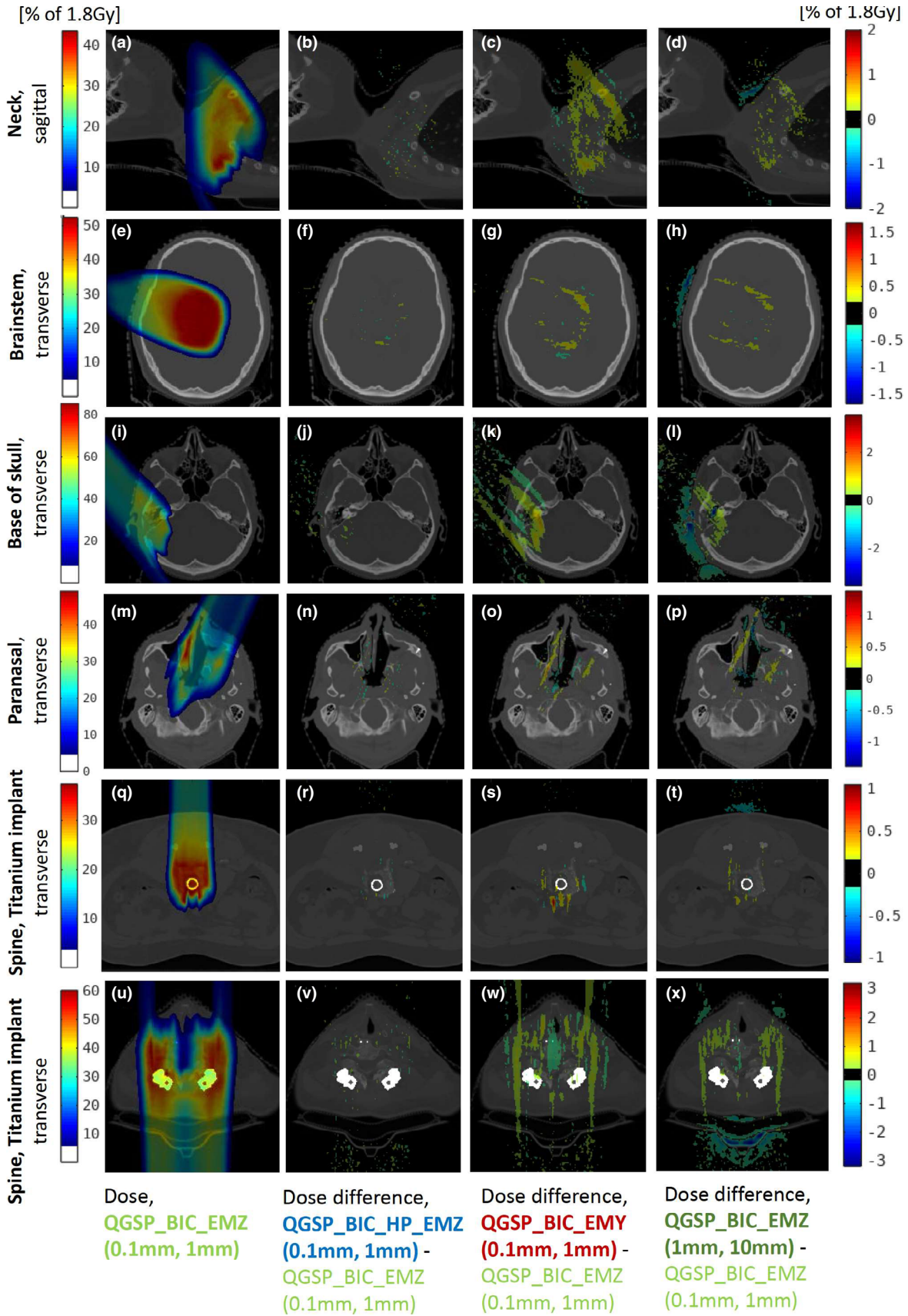


FIG. 6. Dose to medium simulated with EMZ with 0.1 mm production cuts in the range-shifter/phantom and 1 mm production cuts in the world (scenario 5, first column), dose differences between HP_EMZ and EMZ (second column), dose differences between EMY and EMZ (third column), and dose differences due to the settings of small or large production cuts (1 mm production cuts in the range-shifter/phantom and 10 mm production cuts in the world, fourth column), for six different patient fields. Results have been scaled to the prescribed fraction dose of 1.8 Gy. [Color figure can be viewed at wileyonlinelibrary.com]

interfaces. This can be explained by the production of secondaries with ranges between 0.1mm–1mm in the denser tissue, which have larger ranges in the surrounding air. Execution times in CT-based patient models are increased by a factor of 2 when using small *production cuts*.

Physics settings selection are strongly dependent on the end-user application. Establishing recommendations for simulation parameters for the GATE-RTion framework is of paramount importance however, the final choice of simulation and physics settings remains the responsibility of the user.

On the one hand, if using a full Geant4 based MC system, the user might decide for the settings providing the most accurate dosimetric calculations, independently from time constraints. On the other hand, quicker MC simulations might be essential for a clinical environment, and depending on clinical tolerances and workload a compromise between speed and dosimetric accuracy may be required to the user.

As such, the authors would recommend the following Geant4 settings for PSQA dosimetry, when considering voxels of 2 mm size:

- No *step limiter* for proton tracks.
- *Production cuts* on electrons, photons and positrons of 1 mm in the phantom and range-shifter, while adopting a 10 mm value in the surrounding geometry (world).
- Best agreement to measurement data was found for QGSP_BIC_EMZ *reference physics lists* at the cost of 20% increased execution times compared to QGSP_BIC_EMY.

For simulations in CT-based patient models, the following settings are recommended:

- No *step limiter* on proton tracks.
- *Production cuts* on electrons, photons and positrons of 1 mm (phantom/range-shifter) and 10 mm (world) if the goal is to achieve sufficient dosimetric accuracy to ensure that a plan is clinically safe; or 0.1 mm (phantom/range-shifter) and 1 mm (world) if higher dosimetric accuracy is needed. However, these more accurate simulations are subject to a factor 2 increase in the execution time.
- Most accurate results are expected for QGSP_BIC_EMZ *reference physics list*, at the cost of 10–20% increased execution times compared to QGSP_BIC_EMY.

These recommendations are consistent with those of the Geant4 Medical Simulation Benchmarking Group,¹³ which recommends G4EMStandardPhysics_Option4 and QGSP_BIC_HP for hadron therapy.¹⁴ In addition, using the recommendations established in the present study increase efficiency by omitting the HP neutron libraries for this application with no substantial impact on the simulated dose distributions.

It is important to note that these results are specific to the geometries and quantities of interest investigated in this study. As such, the HP module might for example be relevant for out-of-field dose and neutron contributions. However, no

measurements are available from our clinical facility to provide a benchmark for this. Furthermore, there are additional combinations of parameters (for example setting cuts differently in range-shifter and phantom) and Geant4 settings which have not been investigated within the scope of the study, as for example *production cuts* on protons and *step limiter* for other particles (alpha, electrons).

This study has been performed using Geant4 v.10.3.3 as this is the Geant4 version underlying the first GATE-RTion release. The results are applicable to any users of Geant4 v. 10.3.3, including wrappers other than GATE.

5. CONCLUSIONS

GATE-RTion v.1.0 (GATE v.8.1/Geant4 v.10.3.p03) and multiple settings of *step limiter*, *production cuts* and reference *physics lists* have been evaluated against measurement data and optimized for independent dose calculations for proton therapy. For this application, increasing *production cuts* can substantially decrease calculation times. When investigating *physics lists*, High Precision neutron models did not substantially influence the in-field dose. The Geant4 EMZ electromagnetic *physics list* leads to most accurate dose results. Depending on the institute's clinical tolerances and simulation workload however, EMY, which further reduces computation time, might be an acceptable alternative for PSQA purposes. This study has provided recommendation in terms of physics-setting for clinical use of GATE-RTion 1.0 for proton pencil beam scanning PSQA.

ACKNOWLEDGMENTS

This study was performed within the GATE-RTion framework, and we acknowledge the support of both the GATE and Geant4 community. We are thankful for the computation time provided on the proton therapy development cluster, and the computation support by Ian Porter. Furthermore, we thank Edward Smith for providing GATE support.

This work was funded by the Science and Technology Facilities Council (STFC) Advanced Radiotherapy Network, grant number ST/N002423/1 and the Engineering and Physical Sciences Research Council, grant number EP/R023220/1 supported by the NIHR Manchester Biomedical Research Council.

AUTHORS' CONTRIBUTIONS

CW performed the simulations and data analysis and wrote the manuscript draft. AA provided the AUTOMC simulation and analysis framework and supported the data analysis. CW, MT, LM, PS, DS, MV, LG, and AA obtained funding for this study. MT, DB, AE, RM, KK, LM, AR, PS, DS, and MV regularly reviewed the simulation results and the data interpretation. SG and VI reviewed the study and provided knowledge on the underlying GEANT4 code. LG and AA supervised the project. All authors reviewed and approved the final manuscript.

CONFLICT OF INTEREST

The authors have no relevant conflict of interest to disclose.

*Equal contributions.

†Present address: Paul Scherrer Institut, Forschungsstrasse 111, WMSA/C27, Villigen PSI 5232, Schweiz.

^{a)}Author to whom correspondence should be addressed. Electronic mail: carla.winterhalter@psi.ch.

REFERENCES

- Agostinelli S, Allison J, Amako K, et al. Geant4—a simulation toolkit. *Nucl Instrum Methods Phys Res Sect A*. 2003;506:250–303.
- Allison J, Amako K, Apostolakis J, et al. Recent developments in Geant4. *Nucl Instrum Methods Phys Res Sect A*. 2016;835:186–225.
- Allison J, Amako K, Apostolakis J, et al. Geant4 developments and applications. *IEEE Trans Nucl Sci*. 2006;53:270–278.
- Winterhalter C, Zepter S, Shim S, et al. Evaluation of the ray-casting analytical algorithm for pencil beam scanning proton therapy. *Phys Med Biol*. 2019;64:065021.
- Saini J, Maes D, Egan A, et al. Dosimetric evaluation of a commercial proton spot scanning Monte-Carlo dose algorithm: comparisons against measurements and simulations. *Phys Med Biol*. 2017;62:7659.
- Paganetti H, Jiang H, Lee S-Y, Kooy HM. Accurate Monte Carlo simulations for nozzle design, commissioning and quality assurance for a proton radiation therapy facility. *Med Phys*. 2004;31:2107–2118.
- Arce P, Lagaresa IJ, Harkness L, et al. Gamos: a framework to do Geant4 simulations in different physics fields with an user-friendly interface. *Nucl Instrum Methods Phys Res, Sect A*. 2014;735:304–313.
- Perl J, Shin J, Schümann J, Faddegon B, Paganetti H. TOPAS: an innovative proton Monte Carlo platform for research and clinical applications. *Med Phys*. 2012;39:6818–6837.
- Strulab D, Santin G, Lazaro D, Breton V, Morel C. GATE (Geant4 Application for Tomographic Emission): a PET/SPECT general-purpose simulation platform. *Nucl Phys B-Proc Suppl*. 2003;125:75–79.
- Jan S, Benoit D, Becheva E, et al. GATE V6: a major enhancement of the GATE simulation platform enabling modelling of CT and radiotherapy. *Phys Med Biol*. 2011;56:881.
- Sarrut D, Bardès M, Bousson N, et al. A review of the use and potential of the GATE Monte Carlo simulation code for radiation therapy and dosimetry applications. *Med Phys*. 2014;41:064301.
- Grevillot L, Boersma DJ, Fuchs H, et al. Technical note: GATE-RTion: a GATE/Geant4 release for clinical applications in scanned ion beam therapy. *Med Phys*. 2020;47:3675–3681.
- Twiki web page of the G4MSBG initiative: G4-Med, <https://twiki.cern.ch/twiki/bin/view/Geant4/G4MSBG>. Accessed on the 10/01/2020.
- Arce P, Bolst D, Cutajar D, et al. Report on G4-Med, a Geant4 benchmarking system for medical physics applications developed by the Geant4 Medical Simulation Benchmarking Group. *Med Phys*. 2020. <https://doi.org/10.1002/mp.14226>
- Jarlskog CZ, Paganetti H. Physics settings for using the Geant4 toolkit in proton therapy. *IEEE Trans Nucl Sci*. 2008;55:1018–1025.
- Grevillot L, Frisson T, Zahra N, et al. Optimization of Geant4 settings for proton pencil beam scanning simulations using GATE. *Nucl Instrum Methods Phys Res Sect B*. 2010;268:3295–3305.
- Kurosu K, Takashina M, Koizumia M, Dasc IJ, Moskvina VP. 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 Phys Res Sect B*. 2014;336:45–54.
- Kurosu K, Das IJ, Moskvina VP. Optimization of GATE and PHITS Monte Carlo code parameters for spot scanning proton beam based on simulation with FLUKA general-purpose code. *Nucl Instrum Methods Phys Res Sect B*. 2016;367:14–25.
- Fuchs H, Vatnitsky S, Stock M, Georg D, Grevillot L. Evaluation of GATE/Geant4 multiple Coulomb scattering algorithms for a 160 MeV proton beam. *Nucl Instrum Methods Phys Res Sect B*. 2017;410:122–126.
- Resch AF, Elia A, Fuchs H, et al. Evaluation of electromagnetic and nuclear scattering models in GATE/Geant4 for proton therapy. *Med Phys*. 2019;46:2444–2456.
- Sechopoulos I, Rogers DWO, Bazalova-Carter M, et al. RECORDS: improved reporting of Monte Carlo radiation transport studies: report of the AAPM Research Committee Task Group 268. *Med Phys*. 2018;45:e1–e5.
- Aitkenhead A, Sitch P, Richardson JC, Winterhalter C, Patel I, Mackay RI. Automated Monte-Carlo re-calculation of proton therapy plans using Geant4/Gate: implementation and comparison to plan-specific quality assurance measurements. *Br J Radiol*. 2020;93:20200228.
- Eaton JW, Bateman D, Hauberg S, Wehbring R. GNU Octave version 4.4.1 manual: a high-level interactive language for numerical computations. <https://www.gnu.org/software/octave/doc/v4.4.1/>; 2018.
- Chetty IJ, Rosu M, Kessler ML, et al. Reporting and analyzing statistical uncertainties in Monte Carlo-based treatment planning. *Int J Radiat Oncol Biol Phys*. 2006;65:1249–1259.
- Grevillot L, Bertrand D, Dessy F, Freud N, Sarrut D. A Monte Carlo pencil beam scanning model for proton treatment plan simulation using GATE/Geant4. *Phys Med Biol*. 2011;56:5203–5219.
- Grassberger C, Lomax A, Paganetti H. Characterizing a proton beam scanning system for Monte Carlo dose calculation in patients. *Phys Med Biol*. 2015;60:633–645.
- Winterhalter C, Fura E, Tian Y, et al. Validating a Monte Carlo approach to absolute dose quality assurance for proton pencil beam scanning. *Phys Med Biol*. 2018;63:175001.
- Trnkova P, Bolsi A, Albertini F, Weber DC, Lomax AJ. Factors influencing the performance of patient specific quality assurance for pencil beam scanning IMPT fields. *Med Phys*. 2016;43:5998.
- Schneider U, Pedroni E, Lomax A. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Physics in Medicine & Biology*. 1996;41:111.
- Physics Reference Manual, Version: geant4 10.3 (9 December 2016), <http://geant4-userdoc.web.cern.ch/geant4-userdoc/UsersGuides/PhysicsReferenceManual/BackupVersions/V10.3/fo/PhysicsReferenceManual.pdf>. Accessed on the 09/01/2020.
- https://geant4.web.cern.ch/support/user_documentation.09/01/2020.
- Paganetti H. Dose to water versus dose to medium in proton beam therapy. *Phys Med Biol*. 2009;54:4399–4421.