Comparative study of alternative Geant4 hadronic ion inelastic physics models for prediction of positron-emitting radionuclide production in carbon and oxygen ion therapy

Andrew Chacon[1,2], **Susanna Guatelli**[1,4], **Harley Rutherford**[1,2], **David Bolst**[1] **Akram Mohammadi**[5], **Abdella Ahmed**[1,2], **Munetaka Nitta**[5], **Fumihiko Nishikido**[5], **Yuma Iwao**[5], **Hideaki Tashima**[5], **Eiji Yoshida**[5], **Go Akamatsu**[5], **Sodai Takyu**[5], **Atsushi Kitagawa**[5], **Theresa Hofmann**[6], **Marco Pinto**[6], **Daniel R. Franklin**[7], **Katia Parodi**[6], **Taiga Yamaya**[5], **Anatoly Rosenfeld**[1,4], **Mitra Safavi-Naeini**[1,2,3]

[1] Centre for Medical Radiation Physics, University of Wollongong, Wollongong NSW 2522 Australia

[2] Australian Nuclear Science and Technology Organisation (ANSTO), NSW, Australia

[3] Brain and Mind Centre, University of Sydney, Sydney, NSW, Australia

[4] Illawarra Health and Medical Research Institute, University of Wollongong, Wollongong NSW 2522 Australia

[5] National Institute of Radiological Sciences (NIRS), National Institutes for Quantum and Radiological Science and Technology, 4-9-1 Anagawa, Inage-ku, Chiba 263-8555, Japan

[6] Ludwig-Maximilians-Universität München, Medical Physics, Germany

[7] School of Electrical and Data Engineering, University of Technology Sydney, Australia

E-mail: `ac917@uowmail.edu.au`

E-mail: `mitras@ansto.gov.au`

Abstract.

The distribution of fragmentation products predicted by Monte Carlo simulations of heavy ion therapy depend on the hadronic physics model chosen in the simulation. This work aims to evaluate three alternative hadronic inelastic fragmentation physics options available in the Geant4 Monte Carlo radiation physics simulation framework to determine which model most accurately predicts the production of positron-emitting fragmentation products observable using in-beam PET imaging. Fragment distributions obtained with the BIC, QMD, and INCL++ physics models in Geant4 version 10.2.p03 are compared to experimental data obtained at the HIMAC heavy-ion treatment facility at NIRS in Chiba, Japan. For both simulations and experiments, monoenergetic beams are applied to three different block phantoms composed of gelatin, poly(methyl methacrylate) and polyethylene. The yields of the positron-emitting nuclei ^{11}C, ^{10}C and ^{15}O obtained from simulations conducted with each model are compared to the experimental yields estimated by fitting a multi-exponential radioactive decay model to dynamic PET images using the normalised mean square error metric in the entrance, build up / Bragg peak and tail regions. Significant differences in positron-emitting fragment yield are observed among the three physics models with the best overall fit to experimental ^{12}C and ^{16}O beam measurements obtained with the BIC physics model.

1. Introduction

Heavy ion therapy delivers a highly conformal therapeutic radiation dose to a target region while minimising damage to surrounding healthy tissue [1]. This is particularly useful for treating deeply-situated tumours while minimising damage to proximal healthy tissue [2]. However, an unavoidable consequence of its steep dose profile is that treatment with an ion beam is very sensitive to positioning uncertainties - much more so than photon therapy. Small positioning errors may arise due to anatomical changes (e.g., organ motion, tumour regression), patient positioning errors, range errors from uncertainties in measurement of CT Hounsfield units and in the conversion of Hounsfield units into particle stopping power. Any of these may lead to substantial excess radiation exposure to normal tissue and insufficient dose being delivered to the tumour [1, 3]. Intra-fraction and post-fraction quality assurance and treatment validation is therefore a subject of great interest in the particle therapy community, since it offers the opportunity to identify dosing errors and correct them in subsequent fractions.

For quality assurance and treatment validation, much research in particle therapy is aimed at developing new methods to measure particle range in patients and accurately estimate the spatial distribution and magnitude of the delivered dose. One approach to verifying the delivered dose distribution is to image the short-lived positron-emitter fragmentation radionuclides produced by the beam as it travels through the patient [4, 5, 6]. During heavy ion therapy, a fraction of the ions in the incident beam will undergo inelastic collisions with nuclei in the target volume, resulting in the production of a range of fragments [1]. Some of these fragments will be positron-emitting radionuclides, which continue to travel a short distance in the target before coming to a stop, where they will eventually decay. Measurement and visualisation of the distribution of these secondary positron-emitting fragments offers a valuable opportunity for non-invasive quality assurance in heavy ion therapy [7, 8, 9, 10, 11, 12, 13]. As these radionuclides decay by positron emission, and the resulting positrons annihilate with electrons in the target, the spatio-temporal distribution of annihilations can be imaged using a PET scanner. For commonly-used ion species (e.g. ^{12}C, ^{16}O), PET imaging is normally performed as a post-treatment quality assurance (QA) procedure. This could also be extnded to real-time QA for online correction of range errors if either a very high-sensitivity PET scanner is employed and/or if the signal is enhanced by using a positron-emitting radioactive ion beam. Although the PET image is subject to blurring due to non-zero positron range, this degradation can be corrected by separating the positron-emitting radioisotopes through temporal analysis and performing image deconvolution on each image [14, 15]. The resulting image may then be compared to predictions from the treatment planning system and/or Monte Carlo simulations to confirm proper treatment delivery.

Monte Carlo modelling of heavy ion therapy systems is a critical aspect of the development of reliable range verification and dose distribution estimation techniques. As such, it is necessary to establish the accuracy and precision of the physics models used by these simulations. Modelling nuclear interactions and the resulting secondary particle production is highly complex, because it involves high-energy nuclear physics interactions of a diverse range of nuclei, for which no fully validated models currently exist. Several Monte Carlo toolkits are suitable for this application, including Geant4, MCNP6 and FLUKA [16, 17, 18, 19]. Non-invasive in vivo range monitoring methods frequently make use of Monte Carlo predictions of the distribution of secondary particles to infer primary range and estimate dose from the observed image [20, 21].

In this work, the spatial distributions of positron-emitting fragmentation products produced by irradiating a variety of homogeneous phantoms with ^{12}C or ^{16}O beams at different energies are experimentally measured (indirectly) and compared to results obtained by Monte Carlo simulations using Geant4 with three different hadronic ion inelastic physics models. Here, the absolute yields of the dominant positron-emitting fragmentation products (^{10}C, ^{11}C and ^{15}O) are estimated by fitting a multi-exponential radioactive decay model to experimental data obtained using the a high-resolution in-beam whole-body DOI-PET imaging system at NIRS, Japan, during irradiation of gelatin, PMMA and polyethylene block phantoms with beams of ^{12}C ions with energies of 148.5, 290.5 and 350 MeV/u and ^{16}O ions with energies of 148 and 290 MeV/u [22]. The resulting yields are compared against those obtained from Monte Carlo simulations performed with each of the three evaluated Geant4 hadronic ion inelastic physics models: binary ion cascade (BIC), quantum molecular dynamics (QMD) and the Liege intranuclear cascade (INCL++). Experimental and simulation yields were evaluated across the full width at half maximum (FWHM) and full width at tenth maximum (FWTM) of the beam in the irradiated volume. The normalised mean square error (NMSE) of the experimentally estimated yields per primary particle of each positron-emitting fragment to the values obtained via simulation was calculated at the entrance, build-up and Bragg peak, and tail regions.

Section 2 presents a summary of the key related work in this field. The specific details of the experiment and Monte Carlo simulations are presented in Section 3. Experimental and simulation results are presented and discussed in Section 4 including an overall summary of the relative performance of each model, with final conclusions presented in Section 5.

2. Related Work

To date, no comprehensive analysis has been published comparing the accuracy of the various hadronic inelastic physics models available in Geant4 in terms of positron-emitting fragment production. However, numerous studies have partially addressed different aspects of this problem. This section summarises the most significant of these, in particular those studies where some experimental validation has been performed.

Geant4's models for electromagnetic interactions were validated for carbon ion therapy for energies between 90 and 400 MeV/u by Lechner et al., who compared simulated and experimentally-obtained depth dose curves produced by ^{12}C beams incident upon water and polyethylene phantoms [23]. The location of the Bragg peak predicted by Geant4 was found to be in good agreement with experimental results; however, only ^{12}C is evaluated, and the validation is strictly limited to validation of Geant4's models for electromagnetic interactions, since the location of the Bragg Peak depends only on the electromagnetic physics model.

Napoli et al. and Haettner et al. performed a series of experimental studies in which a ΔE-E telescope is used to identify the fragment species, such as carbon or oxygen, produced during particle irradiation with the resulting fragment momentum and angular distribution characterised for ^{12}C beams incident on a range of thin and thick water and PMMA targets [24, 25]. This work was then extended by Bohlen et al., Dudouet et al. and Bolst et al., in separate studies comparing the predictions of Geant4 fragmentation models: Binary Ion Cascade, Quantum Molecular Dynamics and the Liege Intranuclear Cascade model with experimental results [26, 27, 28]. The ΔE-E telescope is able to distinguish between fragments with differing atomic number; however, it is unable to differentiate between different isotopes for ions heavier than helium (such as ^{10}C and ^{11}C), which is of critical importance for PET quality assurance.

A pioneering series of studies comparing Monte Carlo simulation results with experimentally-measured yields of positron-emitting nuclear fragments produced during proton and carbon therapy was conducted at GSI by Parodi et al. and Pönisch et al. [5, 21, 29, 30]. In these studies, experimental positron yields were obtained by imaging a PMMA target during irradiation by pencil proton and ^{12}C beams using a PET system with a spatial resolution of approximately 7 mm. The FLUKA Monte Carlo simulation framework was used to simulate the proton beam, while a specialised in-house simulation code was developed to model the fragmentation process during carbon ion therapy. This work demonstrated the feasibility of imaging a phantom during and after irradiation with proton and ^{12}C beams and obtaining a positron activity profile along the beam axis; it also introduced the idea of fitting the observed activity profile to a multi-exponential radioactive decay model to estimate the proportions of different positron-emitting fragmentation products. This work provided valuable experimental data which was used in many subsequent studies [20, 31, 32, 33, 34, 35]. Further investigations by Sommerer et al. extended the work using FLUKA by conducting a more comprehensive analysis and comparison of the yield of positron-emitting fragments with the experimentally obtained results [18, 19, 36].

Experimental work by Priegnitz et al. demonstrated an approach for predicting positron-emitting fragment distributions during carbon and proton therapy using a PET scanner with 7 mm spatial resolution [31, 32]. The yields of positron emitting nuclei ^{10}C, ^{11}C and ^{15}O were estimated by transversally integrating the observed activity over the whole phantom. A further study by Pshenichnov et al. attempted to compare the predictions of an equivalent Geant4 simulation with experimental estimates of positron-emitting fragments [34, 35]. This work was able to demonstrate that using the Binary Ion Cascade model, coupled with the Geant4 (version 8.0) Radioactive Decay model, the positron activity profile generated using ^{12}C beams inside several different homogeneous phantoms is able to be estimated.

Lau et al. explored the yields of positron-emitting fragments produced during carbon and proton therapy using Geant4 [37]. Different yields were obtained when alternative Geant4 fragmentation models were used. The Quantum Molecular Dynamics (QMD) physics model gave the closest agreement to the experimental results when compared to the BIC model; however, the total yields were averaged over the entire phantom and did not account for the spatial distribution of the fragmentation products.

A study comparing the distributions of secondary particles predicted by different Monte Carlo codes undertaken by Robert et al. did find some notable differences between the results obtained with Geant4 (version 9.4) and FLUKA, especially in the gamma spectrum yields and distribution when using incident proton or carbon beams [33].

Li et al., used Monte Carlo simulations to provide a method for range verification [20]. Their approach was validated using experimental data provided by Parodi et al., which was compared to results from their Geant4 simulations using the Bertini Cascade physics model [21]. When the positron activity profile was normalised to the maximum, good agreement was achieved between the simulation and experimental results.

In summary, there remains a significant knowledge gap concerning the best Geant4 hadronic inelastic ion fragmentation models for simulation of heavy ion therapy. We intend to address this gap by comparing the spatial distributions of positron-emitting fragmentation products resulting from the irradiation of a variety of homogeneous phantoms with ^{12}C or ^{16}O beams at different energies, since these are most relevant for quality assurance methods based on in-beam in vivo PET.

3. Materials and Methods

The evaluation of the three alternative hadronic ion fragmentation models in Geant4 was performed by comparing the predicted depth-dose curves and fragmentation product distributions resulting from simulations conducted with each of the three models (BIC, QMD, INCL++) to that measured experimentally using the normalised mean squared error performance metrics. These comparisons have been performed for carbon ion beams at three incident energies and oxygen ion beams at two incident energies in three different homogeneous phantoms.

The models evaluated were the Binary Ion Cascade (BIC), Quantum Molecular Dynamics and Liège Intranuclear Cascade (INCL++) models [38, 39]. BIC tracks interactions between primary/secondary particles and target nucleons sequentially (hence "binary"), using experimental cross-section data to determine the probability of each type of interaction. Secondary particles are then tracked in turn until both the maximum and average energy of the particles falls below a threshold; in this manner, a single primary results in a tree-like probability graph until all particles are below the minimum energy threshold [38]. By contrast, the QMD model considers multi-body interactions between all nucleons in both projectile and target nuclei. This is intended to offer greater fidelity in the simulation at the cost of computational complexity [38]. Finally, INCL++ is a newer spallation-based model suitable for "light ion" nucleus-nucleus interactions (note: rather confusingly, in this context, the term "light ion" includes "heavy ions" such as carbon and oxygen, due to the different nomenclature used in the high energy physics and medical physics communities) [39, 38].

Monte Carlo simulations were performed using Geant4 toolkit version 10.2.p03 [16] ‡. Electromagnetic interactions were modelled using the standard Geant4 physics option 3 list (`G4EmStandardPhysics_option3`), while the hadronic physics models used are listed in Table 1.

Experimental measurements were performed at the physics beamline of the Heavy Ion Medical Accelerator in Chiba (HIMAC), at Japan's National Institute for Radiological Science (NIRS) in January 2018 with beam parameters for each ion species and energy listed in Table 2.

3.1. Depth-Dose Relationship in Water

Experimental dosimetric measurements were performed using a water phantom and a cruciform ionisation chamber array (Figure 1) [40]. The ionisation chamber consists of two intersecting arms at right angles, both at right angles to the beam, each featuring

‡ In this version of Geant4, the use of the G4IonBinaryCascadePhysics model results in the use of G4BinaryLightIonReaction model (Binary Light Ion Cascade); throughout the rest of this paper, this physics model will be referred to as Binary Ion Cascade (BIC).

Table 1: Hadronic physics processes and models used in all simulations.

Interaction	**Energy Range**	**Geant4 Model**
Radioactive Decay	All energies	G4RadioactiveDecayPhysics
Particle Decay	All energies	G4Decay
Hadron Elastic	0–100 TeV	G4HadronElasticPhysicsHP
Ion Inelastic	<100 MeV 100 MeV–10 GeV	Binary Light Ion Cascade BIC or QMD or INCL++
Neutron Capture	0–20 MeV >19.9 MeV	NeutronHPCapture nRadCapture
Neutron Inelastic	0–20 MeV >19.9 MeV	NeutronHPInelastic Binary Cascade
Proton Inelastic	990 eV–10 TeV	Binary Cascade

Table 2: Beam parameters for each ion species and energy. All beams had an energy spread of 0.2 % of the nominal energy; 95% confidence intervals are listed for beam flux.

Ion	**Energy (MeV/u)**	σ_x **(mm)**	σ_y **(mm)**	**Beam flux (pps)**
^{12}C	148.5	2.77	2.67	$1.8\times10^9 \pm 3.8\times10^7$
^{12}C	290.5	3.08	4.70	$1.8\times10^9 \pm 6.4\times10^7$
^{12}C	350	2.50	2.98	$1.8\times10^9 \pm 4.6\times10^7$
^{16}O	148	2.79	2.89	$1.1\times10^9 \pm 2.8\times10^7$
^{16}O	290	2.60	4.90	$1.1\times10^9 \pm 7.0\times10^7$

65 miniature ionisation chambers with a uniform spacing of 2 mm in both horizontal and vertical dimensions. Each individual ionisation chamber has a depth of 100 μm and the array is positioned with a geared stepper motor with a precision of 100 μm. Energy deposition is measured on the central ionisation chamber only and normalised to the entrance value to produce a normalised dose. The horizontal and vertical transverse beam profiles were obtained by fitting a 2D Gaussian function to the values obtained from the ionisation chamber array; these measurements were used to determine the beam dimensions for the simulation study.

A depth-dose water phantom simulation study was performed using ^{12}C and ^{16}O ion beams using each combination of parameters specified in Table 2 with each of the three hadronic ion inelastic fragmentation physics models under evaluation (BIC, QMD and INCL++). All simulation parameters (phantom geometry and composition, beam energies and dimensions) were configured to match the parameters of the experimental

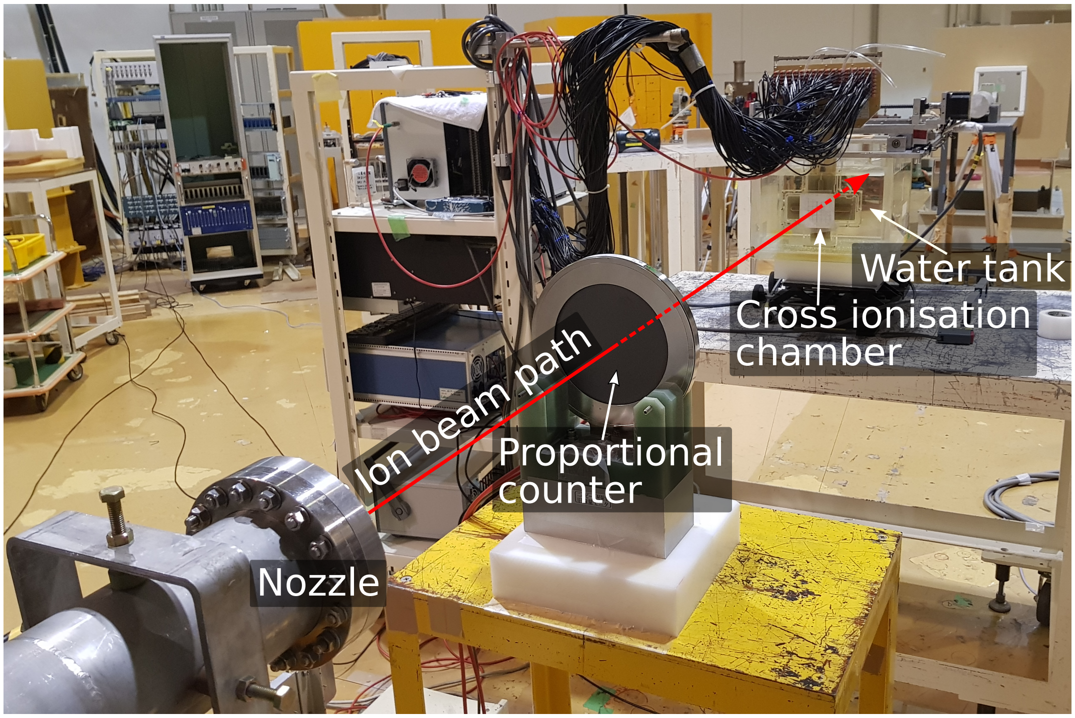


Figure 1: The experimental configuration used for depth-dose measurements.

depth-dose measurements.

3.2. Positron-Emitting Fragment Yield

PMMA, polyethylene and gelatin (encased in a PMMA container) phantoms with dimensions of 100 mm×100 mm×300 mm were used for the positron yield experiments. Transaxial phantom dimensions were ten times the beam diameter, while the axial dimension was sufficient to encompass the maximum particle range for all ion species and energy ranges evaluated. The gelatin phantom comprised a 4 mm thick open rectangular prism PMMA container with internal dimensions of $92\times92\times292$ mm^3, which was then filled with gelatin. As a phantom material for heavy ion therapy, gelatin is essentially equivalent to water (the gel is 98% water by mass), while preventing migration of fragmentation products due to convection. An air gap of 1.75 m was present from the end of the nozzle to the surface of the phantoms.

Positron annihilations were imaged using a whole-body DOI-PET scanner prototype developed at NIRS [22]. Each phantom was positioned so that the expected location of the Bragg peak was approximately located at the centre of the whole-body DOI-PET scanner's field of view (CFOV), as shown in Figure 2. Three repeated irradiations and image acquisitions were performed for each phantom type. Two instances of each phantom type were used in these experiments, such that one phantom of each type could be irradiated while the positron-emitting radionuclides in the other phantoms were allowed to fully decay.

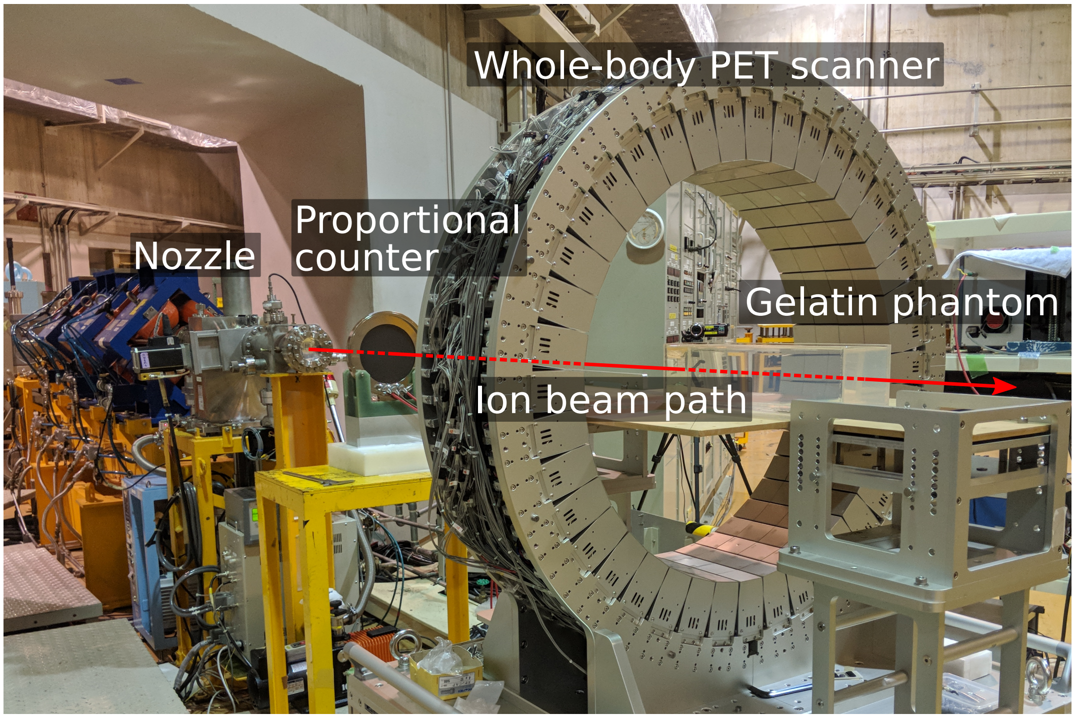
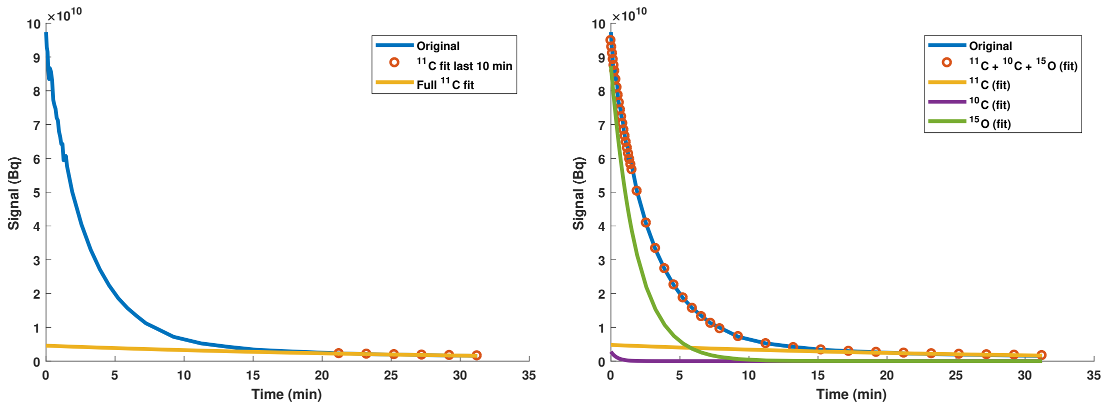


Figure 2: The experimental configuration used for positron-emitting fragment yield estimation. Image acquisition is performed with the whole-body DOI-PET scanner [22].

The beam conditions for the irradiations are detailed in Table 2. Particle therapy irradiation normally consists of a periodic series of beam pulses (called *spills*); in these experiments, a total of 20 spills were used for each beam energy and phantom. Each spill had a beam-on time of 1.9 s followed by a beam-off time of 1.4 s, with a total spill period of 3.3 s.

The whole-body DOI-PET scanner acquired coincidence data in list mode (i.e. a list of coincidence events in which the time of arrival, location and energy deposited by each half of the event is recorded sequentially) during the inter-spill periods and after the final spill post-irradiation, for a total image acquisition time of 30 minutes. Temporal histogramming of the list-mode data was performed in the post-irradiation period with frame lengths chosen such that decay would be observed over several half-lives of ^{11}C, ^{10}C and ^{15}O (20 min, 19 s and 2 min respectively). PET images were then dynamically reconstructed frame-by-frame using the 3D ordinary Poisson ordered-subset-expectation-maximisation (3D-OP-OSEM) algorithm, with a voxel size of $1.5 \times 1.5 \times 1.5$ mm^3.

The absolute yields of each positron-emitting radionuclide were estimated by parametrically fitting a simple multi-exponential radioactive decay model to the observed time-activity curves (TACs), with no decay correction applied, via the Levenberg-Marquardt error minimisation algorithm [41]. Total activity as a function of time t in a volume with initial activities of ^{11}C, ^{10}C and ^{15}O of $A_{0,C11}$, $A_{0,C10}$ and



(a) Fit to final 10 minutes of TAC (^{11}C)   (b) Fit to full TAC (^{10}C, ^{15}O)

Figure 3: Example TACs used for fitting the parameters of Equation (1) (with ^{12}C beam).

$A_{0,O15}$ respectively, is given by

$$A_{total}(t) = A_{0,C11}e^{-\ln t/T_{C11}} + A_{0,C10}e^{-\ln t/T_{C10}} + A_{0,O15}e^{-\ln t/T_{O15}} \quad (1)$$

where T_{C11}, T_{C10}, and T_{O15} are the half-lives of ^{11}C, ^{10}C and ^{15}O, respectively.

The model described in Equation (1) was fitted to TACs corresponding to the average activity in each of a stacked series of small volumes along the path of the beam. Firstly, Equation (1) was fitted to the final 10 minutes of the TAC under the assumption that all ^{10}C and ^{15}O had decayed by this point in order to obtain the activity $A_{0,C11}$ of ^{11}C present immediately following irradiation (Figure 3(a)). Holding $A_{0,C11}$ constant, the remaining coefficients of Equation (1) were then fitted to the TAC spanning the entire time period (Figure 3(b)). The process was performed for each 1.5 mm-deep sample volume extended along the path of the beam. Two different transverse in-beam regions were chosen: the full width at half maximum (FWHM) and the full width at tenth maximum (FWTM) of the beam.

For quantitative analysis, three different regions were chosen: the entrance, build-up and Bragg peak, and tail regions (refer to Figure 4). The central build-up and Bragg peak region is defined as follows:

- The proximal edge in the z dimension (along the path of the beam) is defined as the first point at which activity along the central axis exceeds the entrance plateau activity by more than than 5% of the difference between peak activity and the entrance plateau activity; and
- The distal edge in z is defined as the last point at which activity is greater than 5% of the absolute peak value.

The entrance region is then defined as the region proximal to the build-up and Bragg peak region, while the tail region is defined as the region distal to the build-up and Bragg peak region. In each of the specified regions, different physical processes will dominate the production of positron-emitting radionuclides. In the entrance region, the signal is dominated by target fragmentation from the primary beam, in the build-up and Bragg peak region the signal is dominated by fragmentation of the primary beam, while in the tail region the signal is dominated by the fragmentation of the target by light fragments from the primary beam and target.

The yields of positron emitting nuclei are defined via Equation (2):

$$\text{Yield(Isotope)} = \frac{N(\text{Isotope})}{N(\text{Primary})} \tag{2}$$

where N(Isotope) is the yield of the isotope under study in that region and N(Primary) is the total number of incident particles. Yields were calculated in each voxel along the path of the beam.

For the corresponding simulation studies of fragmentation production, the beam was modelled as a series of 20 spills, with beam-on and beam-off intervals of 1.9 s and 1.4 s, respectively (to match the HIMAC beam used in the experiment). PMMA, gelatin and polyethylene target phantoms were used, with phantom geometries, beam energies, beam dimensions and all other simulation parameters matching the experimental configuration.

The same sets of beam parameters and phantoms were used in the Geant4 simulations as for the experimental study. The locations of positron annihilation (corresponding to the origin of the 511 keV photons) occurring during the 30 minute simulated image acquisition period following final irradiation were scored with a voxel size of 1.5×1.5×1.5 mm^3 and classified according to their parent radionuclide: either ^{10}C, ^{11}C and ^{15}O (other positron-emitting radionuclides were present only in negligible quantities). A total of 20 runs were simulated with the mean and standard deviation of the number of positron annihilations per incident particle calculated in each voxel with a total of 1.0×10^8 incident particles. The mean and standard deviation of the number of each type of parent positron-emitting radionuclide (^{10}C, ^{11}C and ^{15}O) was also calculated for each voxel.

The distributions of positron annihilations parent radionuclides (^{10}C, ^{11}C and ^{15}O) were then convolved with a 3D Gaussian kernel with 2.6 mm FWHM in all dimensions, to model the point spread function of the whole-body DOI-PET system as measured by a ^{18}F point source. The relative yields were then calculated using the multi-exponential model-fitting procedure as used for the experimental positron-emitting fragment yield analysis.

The metric chosen to evaluate the accuracy of the different Geant4 hadronic physics models relative to experimental data was the normalised mean squared error (NMSE). For each phantom (PMMA, gelatin and polyethylene), beam type and energy, the NMSE of annihilation photons as well as the parent isotopes (^{10}C, ^{11}C and ^{15}O) was calculated across the N_{reg} points in the entrance, build-up and Bragg peak, and tail regions.

$$NMSE = \frac{\sum_{i=1}^{N_{reg}} |S_i - E_i|^2}{\sum_{i=1}^{N_{reg}} |E_i|^2} \tag{3}$$

where S_i and E_i are the simulation and experimental yields in the ith voxel of the N_{reg} voxels in region reg.

4. Results and Discussion

In Figure4 and 5(a), the entrance, build-up/Bragg peak, and tail regions are denoted **A**, **B** and **C**, respectively.

4.1. Depth-Dose Relationship in Water

The experimentally-measured and simulated depth dose measurements in the water phantom irradiated with mono-energetic ^{12}C beams with energies of 148.5, 290.5 and 350 MeV/u, normalised to entrance values, are shown in Figures 4(a)-4(c). The minimum measurable depth in the water tank is 26.1 mm due to the dimensions of the water tank (the shallowest entrance-dose samples are omitted from Figure 4(b) due to very high levels of noise which occurred during those measurements which was only discovered after the experiments were completed).

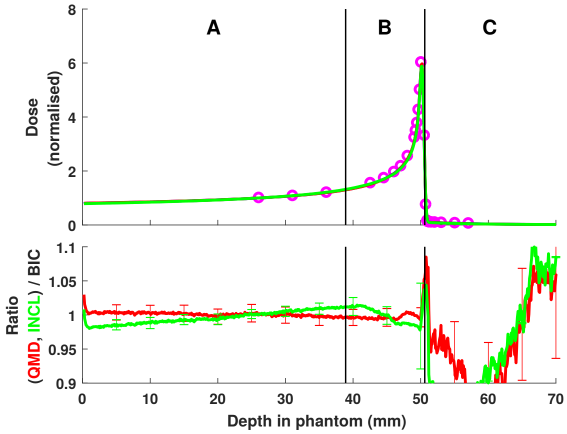
The experimentally-measured and simulated depth dose measurements in the water phantom irradiated with mono-energetic ^{16}O beams with energies of 148 and 290 MeV/u, normalised to the entrance value are shown in Figures 4(d)-4(e). The variation between the depth-dose curves obtained using each hadronic ion inelastic physics model was less than 5% in the entrance (**A**) and build-up/Bragg peak regions (**B**) (which is why the simulation depth-dose curves overlap to the point of obscuring each other in most cases). The large errorbars in the tail region of the QMD to BIC and INCL++ to BIC plots are a consequence of the very low dose recorded in this region (as can be seen in the upper sub-plots).

From these results, it is clear that little variation is evident between the depth-dose curves produced using each of the three hadronic ion inelastic physics models. All of the evaluated models will provide an excellent prediction of the expected depth-dose relationship for the ion species and energies evaluated.

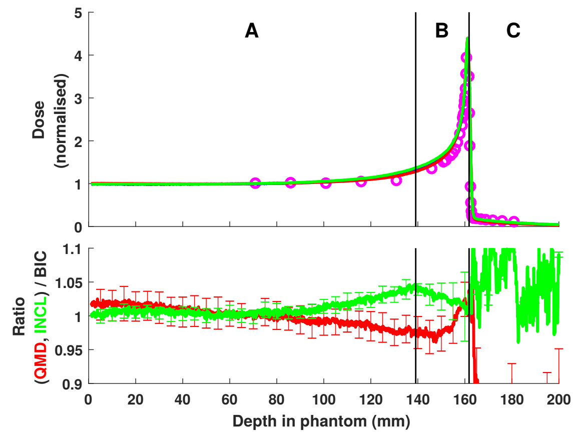
Figure 4 shows that the experimentally-measured and the simulated depth-dose curves are in good agreement for both ^{12}C and ^{16}O at all evaluated beam energies.

4.2. Positron-Emitting Fragment Yield

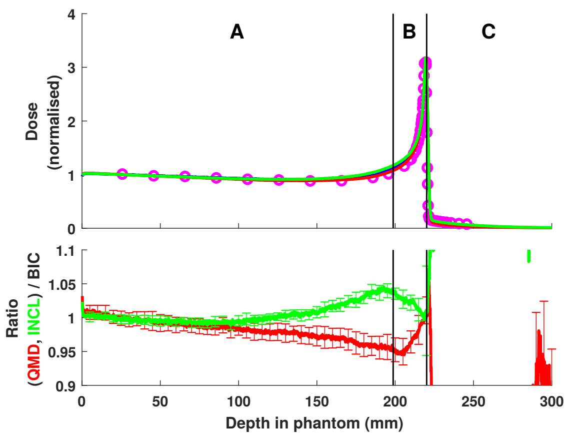
The validation of the Levenberg-Marquardt method for the fitting of the TAC is discussed in the supplementary material. On average, the algorithm estimates the



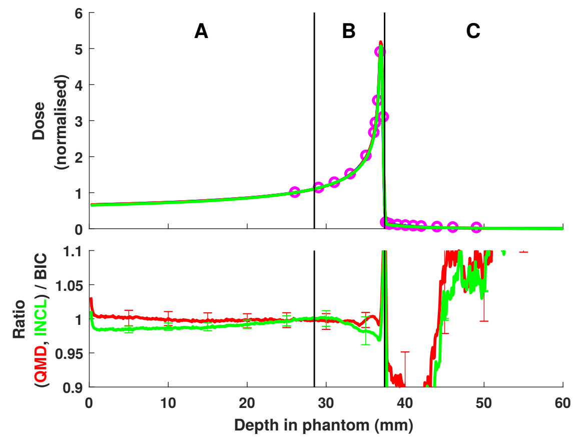
(a) ^{12}C, 148.5 MeV/u



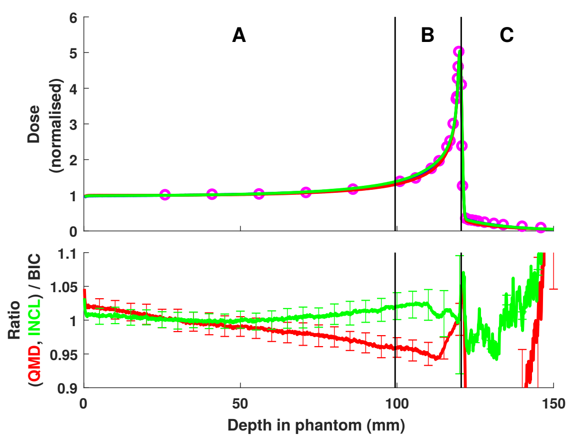
(b) ^{12}C, 290.5 MeV/u



(c) ^{12}C, 148.5 MeV/u

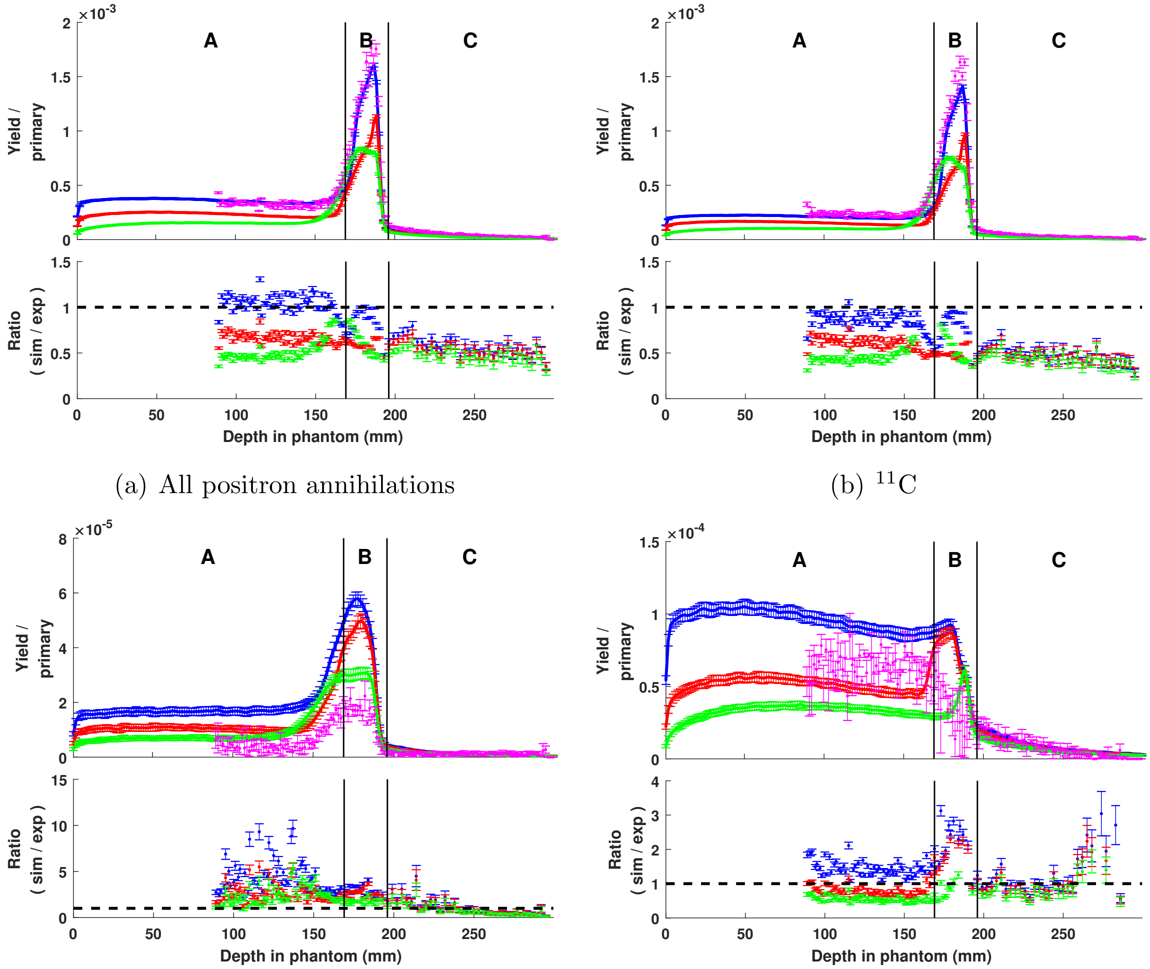


(d) ^{16}O, 148 MeV/u



(e) ^{16}O, 290 MeV/u

Figure 4: The upper sub-plots show experimental (magenta) and simulated (blue = BIC; red = QMD; green = INCL++) dose deposition as a function of depth for ^{12}C and ^{16}O ion beams, normalised to experimental entrance dose. The lower sub-plots show the ratios between the depth-dose simulation results for QMD to BIC (red) and INCL++ to BIC (green). 95% confidence intervals for dose measurements are $< \pm 2\%$ of the mean in all cases and are omitted from the upper sub-plots for clarity; the ratio sub-plots show 95% confidence intervals every 5 mm.



(a) All positron annihilations

(b) ^{11}C

(c) ^{10}C

(d) ^{15}O

Figure 5: Absolute yields of positron annihilations in a PMMA phantom irradiated by 350 MeV/u ^{12}C, both total (5(a)) and down by parent radionuclides (5(b), 5(c) and 5(d)) evaluated in the transverse central half-maximum region of the beam. The corresponding ratio of the simulation result to the experimental result is shown under the absolute yields. Blue = BIC, red = QMD, green = INCL++ and magenta = experiment. A dashed line is drawn at the ratio equal to one. 95% confidence intervals are shown. The 215 mm axial field of view of the whole-body DOI-PET scanner ranges from 85-300 mm.

relative yield of ^{11}C, ^{10}C and ^{15}O from the dynamic PET image with an error smaller 10% compared to the ground truth (the exact number of positron-emitting nuclei produced during the simulation, which is explicitly logged).

An example of total annihilation photon yield and the yield per primary particle of the positron-emitting nuclei ^{11}C, ^{10}C and ^{15}O within the transverse FWHM of the beam, are presented in Figure 5 for the specific case of a 350 MeV/u ^{12}C beam and a PMMA phantom. The ratio between the experiment and simulation results is displayed

Table 3: Entrance region normalised mean square errors for ^{12}C ion beams. Values shown in bold type denote the closest agreement to experimental measurements. "X" denotes measurements in which yields of that particular positron-emitting radionuclide were negligible.

Phantom	Energy (MeV/u)	Model	FWHM				FWTM			
			All e^+	^{11}C	^{10}C	^{15}O	All e^+	^{11}C	^{10}C	^{15}O
PMMA	148.5	BIC	**0.087**	**0.01**	16	**0.14**	**0.061**	**0.0036**	13	**0.078**
		QMD	0.13	0.12	6.7	0.19	0.12	0.11	5.6	0.17
		INCL++	0.41	0.41	**0.81**	0.55	0.69	0.68	**0.15**	0.79
	290.5	BIC	**0.034**	**0.0058**	10	0.19	**0.033**	**0.0033**	10	0.22
		QMD	0.059	0.091	3.4	**0.05**	0.043	0.076	3.6	**0.04**
		INCL++	0.2	0.23	**2.6**	0.21	0.17	0.19	**2.6**	0.18
	350	BIC	**0.016**	**0.058**	4.9	0.2	**0.027**	**0.11**	4.1	**0.042**
		QMD	0.12	0.19	1.4	**0.082**	0.15	0.22	1.3	0.089
		INCL++	0.2	0.21	**1.3**	0.24	0.22	0.23	**0.91**	0.25
Gelatin	148.5	BIC	0.29	**0.023**	11	0.19	0.22	**0.042**	11	**0.1**
		QMD	**0.16**	0.049	**4.7**	**0.17**	**0.13**	0.081	**4.5**	0.15
		INCL++	0.26	0.15	6.7	0.55	0.27	0.16	6.3	0.55
	290.5	BIC	0.14	**0.039**	12	0.17	0.14	**0.045**	20	0.16
		QMD	**0.032**	0.059	**5.8**	**0.068**	**0.021**	0.061	**11**	**0.045**
		INCL++	0.13	0.062	8.7	0.26	0.11	0.053	15	0.21
	350	BIC	**0.075**	0.25	5.4	0.22	**0.037**	0.31	5.2	**0.09**
		QMD	0.094	0.35	**2.4**	**0.086**	0.13	0.4	**2.7**	0.1
		INCL++	0.11	**0.03**	4.1	0.26	0.15	**0.064**	3.6	0.27
Polyethylene	148.5	BIC	**0.05**	**0.032**	21	X	**0.037**	**0.021**	18	X
		QMD	0.13	0.13	8.8	X	0.13	0.13	7.6	X
		INCL++	0.47	0.5	**0.41**	X	0.46	0.49	**0.34**	X
	290.5	BIC	**0.0068**	**0.0062**	4.7	X	**0.0051**	**0.0043**	4.7	X
		QMD	0.094	0.1	0.99	X	0.076	0.083	1.1	X
		INCL++	0.25	0.26	**0.57**	X	0.21	0.23	**0.58**	X
	350	BIC	**0.083**	**0.085**	5.1	X	**0.041**	**0.06**	2.6	X
		QMD	0.14	0.15	1	X	0.18	0.21	0.59	X
		INCL++	0.24	0.26	**0.81**	X	0.2	0.21	**0.34**	X

under each respective graph.

The following sections present detailed tabulated results comparing each simulation model with the experimental results in the transverse FWHM and FWTM sections of the entrance, build-up/Bragg peak and tail regions. In each table, the simulation results with the closest agreement to experimental results (i.e. where the NMSE is closest to 0) are shown in bold type.

4.2.1. Entrance region The normalised mean squared errors between simulation and experimental total annihilation photon yield and the yield of each of the parent radionuclides in the entrance region for ^{12}C and ^{16}O beams are listed in Tables 3 and Tables 4, respectively.

For a simulated ^{12}C beam (Table 3), simulations performed using the BIC hadronic physics model show the closest agreement to the observed experimental results in terms of total positron annihilations observed in the entrance region for PMMA

Table 4: Entrance region normalised mean square errors for ^{16}O ion beams. Values shown in bold type denote the closest agreement to experimental measurements. "X" denotes measurements in which yields of that particular positron-emitting radionuclide were negligible.

Phantom	Energy (MeV/u)	Model	FWHM				FWTM			
			All e^+	^{11}C	^{10}C	^{15}O	All e^+	^{11}C	^{10}C	^{15}O
PMMA	148	BIC	**0.0063**	**0.047**	2.5	**0.15**	**0.0047**	**0.046**	3.4	**0.12**
		QMD	0.11	0.14	1.1	**0.15**	0.1	0.14	1.5	0.15
		INCL++	0.48	0.48	**0.29**	0.55	0.46	0.47	**0.37**	0.55
	290	BIC	**0.015**	**0.067**	2	**0.078**	**0.034**	**0.1**	2.2	**0.011**
		QMD	0.14	0.16	0.85	0.11	0.17	0.19	1	0.15
		INCL++	0.38	0.39	**0.22**	0.33	0.38	0.4	**0.26**	0.36
Gelatin	148	BIC	0.45	**0.0073**	16	0.32	0.34	**0.013**	25	0.21
		QMD	**0.19**	0.084	5.9	**0.18**	**0.17**	0.095	9.7	**0.15**
		INCL++	0.31	0.25	**3.9**	0.52	0.33	0.26	**6.7**	0.53
	290	BIC	**0.048**	**0.059**	12	**0.062**	0.085	**0.04**	20	0.1
		QMD	0.075	0.074	3.3	0.099	**0.037**	0.049	6.5	**0.059**
		INCL++	0.25	0.12	**3**	0.33	0.19	0.086	**6**	0.27
Polyethylene	148	BIC	**0.1**	**0.093**	1.1	X	**0.11**	**0.11**	0.81	X
		QMD	0.23	0.23	0.68	X	0.24	0.24	0.45	X
		INCL++	0.62	0.63	**0.099**	X	0.62	0.63	**0.08**	X
	290	BIC	**0.049**	**0.046**	2.8	X	**0.021**	**0.019**	3	X
		QMD	0.15	0.15	2	X	0.098	0.095	2.2	X
		INCL++	0.42	0.43	**0.29**	X	0.35	0.36	**0.35**	X

and polyethylene phantoms. For gelatin, the QMD model provides the best match to the experimental measurements in the entrance region for energies of 148.5 and 290.5 MeV/u, while BIC provides the best match at 350 MeV/u. In the case of the ^{16}O beam (Table 4), the BIC implementation provides the best fit for total positron annihilations.

In all models, the production of ^{10}C tends to be overestimated compared to the experimental estimates. However, this positron-emitting radioisotope is produced in relatively small quantities compared to the others, and small errors in the fitting of the multi-exponential radioactive decay model to the experimental data may have resulted in a underestimation of the true production of ^{10}C (the small proportion of ^{10}C in the observed PET signal does not significantly constrain the behaviour of the optimiser in these cases).

For the carbon beam in the entrance region, BIC was the most accurate in 49% of energy and target combinations, QMD in 30% and INCL++ in 21%. For oxygen, BIC was most accurate in 61% of cases, QMD in 13% and INCL++ in 26%. BIC was therefore the most accurate model for both ion species in the entrance region; QMD was next best for carbon followed by INCL++ while these results were reversed for oxygen.

4.2.2. Build-up and Bragg peak region The normalised mean squared errors between simulation and experimental total annihilation photon yield and the yield of each of the

Table 5: Build-up and Bragg peak region normalised mean square errors for ^{12}C ion beams. Values shown in bold type denote the closest agreement to experimental measurements. "X" denotes measurements in which yields of that particular positron-emitting radionuclide were negligible.

Phantom	**Energy (MeV/u)**	**Model**	**FWHM**				**FWTM**			
			All e$^+$	**^{11}C**	**^{10}C**	**^{15}O**	**All e$^+$**	**^{11}C**	**^{10}C**	**^{15}O**
PMMA	148.5	BIC	**0.03**	**0.042**	1.6	0.2	**0.034**	**0.05**	1.2	**0.17**
		QMD	0.076	0.11	1.6	**0.19**	0.079	0.12	1.2	0.18
		INCL++	0.22	0.27	**0.34**	0.27	0.55	0.59	**0.22**	0.56
	290.5	BIC	**0.06**	**0.063**	5.7	0.5	**0.062**	**0.065**	6.1	0.54
		QMD	0.076	0.13	3.9	0.39	0.068	0.12	4	0.43
		INCL++	0.17	0.21	**0.84**	**0.21**	0.17	0.21	**0.76**	**0.13**
	350	BIC	**0.02**	**0.04**	3.9	1.5	**0.039**	**0.065**	3.6	0.58
		QMD	0.17	0.24	2.4	1.2	0.19	0.26	2	0.42
		INCL++	0.2	0.24	**0.5**	**0.34**	0.27	0.31	**0.27**	**0.16**
Gelatin	148.5	BIC	**0.07**	**0.093**	0.72	**0.26**	**0.076**	**0.11**	0.54	**0.15**
		QMD	0.094	0.099	0.77	**0.26**	0.11	0.15	0.57	**0.15**
		INCL++	0.18	0.16	**0.64**	0.35	0.21	0.22	**0.41**	0.38
	290.5	BIC	**0.017**	**0.036**	3.1	0.38	**0.018**	**0.036**	3.6	0.4
		QMD	0.091	0.2	2.1	0.29	0.084	0.2	2.3	0.3
		INCL++	0.14	0.15	**1.2**	**0.19**	0.15	0.17	**1.1**	**0.15**
	350	BIC	**0.033**	**0.075**	2.6	1.1	**0.05**	**0.097**	3.4	0.44
		QMD	0.22	0.37	1.5	0.81	0.23	0.37	1.9	0.31
		INCL++	0.19	0.18	**1.2**	**0.21**	0.26	0.27	**1.1**	**0.17**
Polyethylene	148.5	BIC	**0.049**	**0.062**	1.8	X	**0.046**	**0.06**	1.5	X
		QMD	0.11	0.14	1.8	X	0.11	0.14	1.5	X
		INCL++	0.17	0.2	**0.21**	X	0.19	0.22	**0.13**	X
	290.5	BIC	**0.044**	**0.043**	3.6	X	**0.041**	**0.04**	4.2	X
		QMD	0.14	0.17	2.3	X	0.12	0.15	2.5	X
		INCL++	0.12	0.14	**0.34**	X	0.12	0.14	**0.34**	X
	350	BIC	**0.066**	**0.071**	2.7	X	**0.024**	**0.032**	2.4	X
		QMD	0.19	0.22	1.5	X	0.21	0.24	1.2	X
		INCL++	0.17	0.19	**0.25**	X	0.21	0.23	**0.11**	X

parent radionuclides in the build-up and Bragg peak region for ^{12}C and ^{16}O beams are listed in Tables 5 and Tables 6, respectively.

The results of the comparison are slightly different in the build-up and Bragg peak region compared to the entrance. For a simulated ^{12}C beam (Table 5), BIC outperforms all other hadronic physics models in all phantoms and at all energies in terms of both total positron annihilations and ^{11}C production. It achieves very good agreement with the experimental data in most cases. The discrepancy between the simulated and experimental estimates of ^{10}C production is still large, but smaller than in the entrance region.

With the ^{16}O beam (Table 6), BIC produces the overall best match for positron production (performing best in 5 of the 6 combinations of energy and phantom). The production of ^{15}O is best modelled by BIC in most cases; again, ^{10}C production is overestimated by all models compared to the fitted experimental data.

For the carbon beam in the build up/Bragg peak region, BIC was the most accurate

Table 6: Build-up and Bragg peak region normalised mean square errors for ^{16}O ion beams. Values shown in bold type denote the closest agreement to experimental measurements. “X” denotes measurements in which yields of that particular positron-emitting radionuclide were negligible.

Phantom	Energy (MeV/u)	Model	FWHM				FWTM			
			All e^+	^{11}C	^{10}C	^{15}O	All e^+	^{11}C	^{10}C	^{15}O
PMMA	148	BIC	0.13	0.17	3.5	**0.12**	0.14	0.16	5.6	**0.14**
		QMD	**0.096**	**0.14**	3.9	0.15	**0.098**	**0.13**	6	0.16
		INCL++	0.2	0.21	**2.3**	0.3	0.21	0.22	**3.3**	0.31
	290	BIC	**0.087**	0.16	16	**0.078**	**0.078**	**0.17**	19	**0.075**
		QMD	0.089	**0.14**	16	0.21	0.12	**0.17**	16	0.25
		INCL++	0.14	0.17	**13**	0.24	0.17	0.21	**11**	0.27
Gelatin	148	BIC	**0.024**	0.099	4.9	**0.023**	**0.019**	0.11	7.5	**0.025**
		QMD	0.032	0.074	5.7	0.096	0.035	0.084	8.4	0.11
		INCL++	0.091	**0.048**	**3.3**	0.21	0.11	**0.077**	**4.8**	0.23
	290	BIC	**0.012**	0.28	84	**0.02**	**0.031**	0.21	170	**0.047**
		QMD	0.17	0.23	92	0.26	0.12	0.16	180	0.2
		INCL++	0.12	**0.12**	**81**	0.19	0.078	**0.071**	**150**	0.14
Polyethylene	148	BIC	**0.032**	0.16	2.6	**0.047**	**0.036**	0.16	3.6	**0.051**
		QMD	0.065	**0.13**	3.2	0.16	0.07	**0.14**	4.1	0.17
		INCL++	0.16	0.19	**1.5**	0.26	0.18	0.21	**1.9**	0.28
	290	BIC	**0.026**	**0.12**	13	**0.04**	0.085	**0.071**	22	**0.06**
		QMD	0.18	0.14	13	0.36	0.13	0.098	20	0.31
		INCL++	0.098	0.15	**9.8**	0.17	**0.064**	0.12	**14**	0.12

in 56% of energy and target combinations, QMD in 6% and INCL++ in 38%. For oxygen, BIC was most accurate in 50% of cases, QMD in 17% and INCL++ in 33%. BIC was therefore the most accurate model for both ion species in the build up/Bragg peak region, followed by INCL++ and QMD.

4.2.3. Tail region The normalised mean squared errors between simulation and experimental total annihilation photon yield and the yield of each of the parent radionuclides in the entrance region for ^{12}C and ^{16}O beams are listed in Tables 7 and Tables 8, respectively.

For the ^{12}C beam (Table 7), none of the models provided a particularly good fit to the experimental positron annihilation distribution; however, INCL++ was consistently the worst performer. For most phantoms and energies, the estimated ^{10}C production was closer to the experimentally-measured values than was the case in the entrance or build-up/Bragg peak region.

With ^{16}O (Table 8), none of the models significantly out performed the others. BIC provided the best match to the experimental positron annihilation distributions in gelatin, while QMD provided the best match in PMMA.

For the carbon beam in the tail region, BIC was the most accurate in 43% of energy and target combinations, QMD in 34% and INCL++ in 23%. For oxygen, BIC was most accurate in 23% of cases, QMD in 50% and INCL++ in 27%. QMD was therefore the

Table 7: Tail region normalised mean square errors for ^{12}C ion beams. Values shown in bold type denote the closest agreement to experimental measurements. "X" denotes measurements in which yields of that particular positron-emitting radionuclide were negligible.

Phantom	Energy (MeV/u)	Model	FWHM				FWTM			
			All e^+	^{11}C	^{10}C	^{15}O	All e^+	^{11}C	^{10}C	^{15}O
PMMA	148.5	BIC	0.15	0.28	**0.14**	0.054	0.13	0.26	**0.16**	0.078
		QMD	**0.12**	**0.25**	**0.14**	**0.04**	**0.096**	**0.22**	0.21	**0.046**
		INCL++	0.3	0.36	0.21	0.16	0.7	0.73	0.53	0.66
	290.5	BIC	**0.073**	0.15	0.96	**0.022**	0.066	0.15	2.6	**0.013**
		QMD	**0.073**	**0.14**	0.75	0.027	**0.059**	**0.12**	2.3	0.018
		INCL++	0.19	0.24	**0.27**	0.096	0.19	0.24	**0.69**	0.11
	350	BIC	**0.14**	**0.24**	0.58	**0.058**	**0.14**	**0.24**	1.9	**0.037**
		QMD	0.19	0.27	0.38	0.069	0.16	0.25	1.3	0.071
		INCL++	0.26	0.31	**0.27**	0.11	0.27	0.33	**0.44**	0.17
Gelatin	148.5	BIC	0.17	0.61	0.52	0.086	0.14	0.61	0.34	0.12
		QMD	**0.15**	0.59	**0.48**	**0.057**	**0.11**	**0.58**	**0.3**	**0.079**
		INCL++	0.32	**0.58**	0.5	0.24	0.31	**0.58**	0.34	0.29
	290.5	BIC	**0.067**	0.47	**0.6**	**0.02**	**0.047**	0.41	2.3	**0.015**
		QMD	0.08	0.44	0.67	0.036	0.049	0.36	2.7	0.022
		INCL++	0.18	**0.39**	0.63	0.14	0.16	**0.35**	**2**	0.14
	350	BIC	**0.071**	0.34	0.56	**0.027**	**0.068**	0.34	**2.1**	**0.033**
		QMD	0.13	0.33	**0.54**	0.073	0.1	0.3	2.2	0.075
		INCL++	0.17	**0.18**	0.59	0.15	0.18	**0.22**	2.2	0.18
Polyethylene	148.5	BIC	0.075	0.15	0.9	X	0.066	0.13	1	X
		QMD	**0.063**	**0.12**	0.81	X	**0.047**	**0.093**	1.1	X
		INCL++	0.18	0.23	**0.23**	X	0.18	0.22	**0.15**	X
	290.5	BIC	**0.04**	**0.064**	4.7	X	**0.04**	**0.065**	4.9	X
		QMD	0.052	0.076	3.3	X	0.044	0.066	3.6	X
		INCL++	0.14	0.16	**0.74**	X	0.15	0.17	**0.73**	X
	350	BIC	**0.2**	**0.25**	2.4	X	**0.12**	**0.15**	3.2	X
		QMD	0.27	0.31	1.1	X	0.16	0.19	1.8	X
		INCL++	0.32	0.35	**0.4**	X	0.24	0.27	**0.43**	X

most accurate model for carbon in the tail region, followed by QMD and INCL++, while for oxygen the best performing model is QMD, followed by INCL++ and QMD.

4.2.4. Overall performance In summary, the hadronic the inelastic physics model which was most consistently able to match experimental results obtained with a ^{12}C or ^{16}O beam across the widest range of phantoms and energies was BIC. INCL++ was rarely the best or worst-performing model, most frequently achieving a middle ranking. QMD varied between good and poor performance depending on the region, incident ion, target and the positron-emitting fragment analysed. While excellent agreement was obtained for depth-dose curves, and (for BIC in most cases) for positron annihilation distributions, the accuracy of the predicted level of production of individual positron-emitting radionuclides varied substantially. In most cases, the distribution of the dominant radionuclide could be predicted with a good degree of reliability.

For both beam types, results obtained when positron activity and positron-emitting

Table 8: Tail region normalised mean square errors for ^{16}O ion beams. Values shown in bold type denote the closest agreement to experimental measurements. "X" denotes measurements in which yields of that particular positron-emitting radionuclide were negligible.

Phantom	**Energy (MeV/u)**	**Model**	**FWHM**				**FWTM**			
			All e$^+$	**^{11}C**	**^{10}C**	**^{15}O**	**All e$^+$**	**^{11}C**	**^{10}C**	**^{15}O**
PMMA	148	BIC	0.38	0.33	0.085	0.55	0.36	0.33	0.12	0.57
		QMD	**0.34**	**0.31**	**0.052**	**0.47**	**0.31**	**0.28**	0.15	**0.48**
		INCL++	0.41	0.35	0.1	0.56	0.4	0.35	**0.05**	0.58
	290	BIC	0.19	0.24	2	0.25	0.2	0.24	2.4	0.26
		QMD	**0.13**	**0.15**	2.3	**0.19**	**0.14**	**0.16**	2.5	**0.18**
		INCL++	0.22	0.22	**0.91**	0.31	0.26	0.26	**0.73**	0.32
Gelatin	148	BIC	**0.097**	**0.035**	2.9	0.33	**0.12**	**0.063**	5.1	0.35
		QMD	0.13	0.083	1.9	**0.28**	0.14	0.12	3.2	**0.28**
		INCL++	0.22	0.11	**0.84**	0.42	0.26	0.17	**1.2**	0.43
	290	BIC	**0.024**	**0.049**	15	0.068	0.022	**0.027**	28	0.064
		QMD	0.044	0.067	9.4	**0.052**	**0.018**	0.04	17	**0.031**
		INCL++	0.091	0.068	**5.5**	0.17	0.077	0.06	**9.2**	0.17
Polyethylene	148	BIC	0.22	0.19	**1.7**	X	0.25	0.22	**1.7**	X
		QMD	0.19	0.15	2.8	X	**0.21**	**0.16**	2.9	X
		INCL++	**0.18**	**0.14**	2	X	**0.21**	0.17	1.9	X
	290	BIC	0.063	0.12	7.8	X	0.051	0.087	9.6	X
		QMD	**0.025**	**0.044**	12	X	**0.015**	**0.022**	14	X
		INCL++	0.037	0.061	**7.4**	X	0.037	0.055	**8**	X

radionuclide production were evaluated over the transverse FWTM of the beam rather than FWHM were essentially equivalent to the FWHM case.

Despite the overall underestimation of ^{10}C production, it may be noted from Figure 5 that both edges of the Bragg peak region in the ^{10}C signal are still clearly defined and are in good agreement with experimental data for the case of INCL++, in shape if not in magnitude; therefore, in modelling on-line range verification systems which rely on the production of ^{10}C, INCL++ may be worth considering (although the other models nevertheless provide a fair estimate of the position of the distal edge and a fair estimate of the proximal edge).

5. Conclusion

The performance of three Geant4 hadronic inelastic ion physics models - Binary Ion Cascade (BIC), Quantum Molecular Dynamics (QMD) and Liege Intranuclear Cascade model (INCL++) - were evaluated according to their ability to accurately predict the depth-dose curve, overall positron annihilation distribution and the distributions of individual positron-emitting fragmentation products produced during heavy ion therapy, with both ^{12}C and ^{16}O beams, in three different homogeneous phantoms in Geant4 version 10.2.p03. The yield of positron-emitting radionuclides predicted by each of these models depends strongly on both the phantom composition and region of interest inside the phantom, with the BIC model outperforming the other two models for the overall prediction of the in-beam positron annihilation and dominant positron-emitting fragment distribution profiles for both ^{12}C and ^{16}O beams. Therefore the adoption of the BIC hadronic inelastic ion physics model is recommended as the best model for fragmentation processes observable using in-beam, in-vivo PET imaging in heavy ion therapy, although for modelling real-time intra-spill imaging, INCL++ may provide a better estimate of the ^{10}C-dominated proximal edge of the Bragg peak.

6. Acknowledgements

The authors would like to acknowledge the following organisations for providing access to their high performance computing resources: University of Wollongong "Centaur 2" Cluster, the Multi-modal Australian Sciences Imaging and Visualisation Environment (MASSIVE) "M3" cluster and Australia's Nuclear Science and Technology Organisation (ANSTO) "Commodore" cluster. This research has been conducted with the support of the Australian government research training program scholarship. The authors acknowledge the scientific and technical assistance of the National Imaging Facility, a National Collaborative Research Infrastructure Strategy (NCRIS) capability, at the Australian Nuclear Science and Technology Organisation, ANSTO.

References

[1] D. Schardt, T. Elsässer, D. Schulz-Ertner, Heavy-ion tumor therapy: Physical and radiobiological benefits, Reviews of Modern Physics 82 (1) (2010) 383–425. doi:10.1103/revmodphys.82.383.

[2] M. Durante, J. Loeffler, Charged particles in radiation oncology, Nature Reviews Clinical Oncology 7 (1) (2009) 37–43. doi:10.1038/nrclinonc.2009.183.
URL `https://doi.org/10.1038/nrclinonc.2009.183`

[3] U. Schneider, E. Pedroni, A. Lomax, The calibration of CT Hounsfield units for radiotherapy treatment planning, Physics in Medicine and Biology 41 (1) (1996) 111–124. doi:10.1088/0031-9155/41/1/009.
URL `https://doi.org/10.1088/0031-9155/41/1/009`

[4] H. D. Maccabee, U. Madhvanath, M. R. Raju, Tissue activation studies with alpha-particle beams, Physics in Medicine and Biology 14 (2) (1969) 213–224. doi:10.1088/0031-9155/14/2/304.
URL `https://doi.org/10.1088/0031-9155/14/2/304`

[5] K. Parodi, W. Enghardt, T. Haberer, In-beam PET measurements of β^+ radioactivity induced by proton beams, Physics in Medicine and Biology 47 (1) (2001) 21–36. doi:10.1088/0031-9155/47/1/302.
URL `https://doi.org/10.1088/0031-9155/47/1/302`

[6] H. Paganetti, G. El Fakhri, Monitoring proton therapy with PET, The British Journal of Radiology 88 (1051) (2015) 20150173.

[7] Y. Iseki, H. Mizuno, Y. Futami, T. Tomitani, T. Kanai, M. Kanazawa, A. Kitagawa, T. Murakami, T. Nishio, M. Suda, E. Urakabe, A. Yunoki, H. Sakai, Positron camera for range verification of heavy-ion radiotherapy, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 515 (3) (2003) 840–849. doi:10.1016/j.nima.2003.07.005.
URL `https://doi.org/10.1016/j.nima.2003.07.005`

[8] G. Kraft, U. Arndt, W. Becher, D. Schardt, H. Stelzer, U. Weber, T. Archinal, Heavy ion therapy at GSI, Nuclear Instruments and Methods in Physics Research Section A: Accelerators,

Spectrometers, Detectors and Associated Equipment 367 (1-3) (1995) 66–70. doi:10.1016/0168-9002(95)00735-0.
URL https://doi.org/10.1016/0168-9002(95)00735-0

[9] E. Yoshida, H. Tashima, T. Shinaji, K. Shimizu, H. Wakizaka, A. Mohammadi, F. Nishikido, T. Yamaya, Development of a Whole-Body Dual Ring OpenPET for in-Beam PET, IEEE Transactions on Radiation and Plasma Medical Sciences 1 (4) (2017) 293–300. doi:10.1109/trpms.2017.2703823.
URL https://doi.org/10.1109/trpms.2017.2703823

[10] V. Ferrero, E. Fiorina, M. Morrocchi, F. Pennazio, G. Baroni, G. Battistoni, N. Belcari, N. Camarlinghi, M. Ciocca, A. Del Guerra, M. Donetti, S. Giordanengo, G. Giraudo, V. Patera, C. Peroni, A. Rivetti, M. Rolo, S. Rossi, V. Rosso, G. Sportelli, S. Tampellini, F. Valvo, R. Wheadon, P. Cerello, M. Bisogni, Online proton therapy monitoring: clinical test of a Silicon-photodetector-based in-beam PET, Scientific Reports 8 (4100) (2018). doi:10.1038/s41598-018-22325-6.

[11] Y. Hirano, E. Yoshida, S. Kinouchi, F. Nishikido, N. Inadma, H. Murayama, T. Yamaya, Monte Carlo simulation of small OpenPET prototype with11C beam irradiation: effects of secondary particles on in-beam imaging, Physics in Medicine and Biology 59 (7) (2014) 1623–1640. doi:10.1088/0031-9155/59/7/1623.
URL https://doi.org/10.1088/0031-9155/59/7/1623

[12] K. Parodi, H. Paganetti, H. Shih, S. Michaud, J. Loeffler, T. DeLaney, N. Liebsch, J. Munzenrider, A. Fischman, A. Knopf, T. Bortfeld, Patient Study of In Vivo Verification of Beam Delivery and Range, Using Positron Emission Tomography and Computed Tomography Imaging After Proton Therapy, International Journal of Radiation Oncology*Biology*Physics 68 (3) (2007) 920–934. doi:10.1016/j.ijrobp.2007.01.063.
URL https://doi.org/10.1016/j.ijrobp.2007.01.063

[13] S. Combs, J. Bauer, D. Unholtz, C. Kurz, T. Welzel, D. Habermehl, T. Haberer, J. Debus, K. Parodi, Monitoring of patients treated with particle therapy using positron-emission-tomography (PET): the MIRANDA study, BMC Cancer 12 (1) (Apr. 2012). doi:10.1186/1471-2407-12-133.
URL https://doi.org/10.1186/1471-2407-12-133

[14] A. Rahmim, J. Qi, V. Sossi, Resolution modeling in PET imaging: Theory, practice, benefits, and pitfalls, Medical Physics 40 (6Part1) (2013) 064301. doi:10.1118/1.4800806.
URL https://doi.org/10.1118/1.4800806

[15] O. Bertolli, A. Eleftheriou, M. Cecchetti, N. Camarlinghi, N. Belcari, C. Tsoumpas, PET iterative reconstruction incorporating an efficient positron range correction method, Physica Medica 32 (2) (2016) 323–330. doi:10.1016/j.ejmp.2015.11.005.
URL https://doi.org/10.1016/j.ejmp.2015.11.005

[16] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, F. Behner, L. Bellagamba, J. Boudreau, L. Broglia, A. Brunengo, H. Burkhardt, S. Chauvie, J. Chuma, R. Chytracek, G. Cooperman, G. Cosmo, P. Degtyarenko, A. Dell'Acqua, G. Depaola, D. Dietrich, R. Enami, A. Feliciello, C. Ferguson, H. Fesefeldt, G. Folger, F. Foppiano, A. Forti, S. Garelli, S. Giani, R. Giannitrapani, D. Gibin, J. G. Cadenas, I. González, G. G. Abril, G. Greeniaus, W. Greiner, V. Grichine, A. Grossheim, S. Guatelli, P. Gumplinger, R. Hamatsu, K. Hashimoto, H. Hasui, A. Heikkinen, A. Howard, V. Ivanchenko, A. Johnson, F. Jones, J. Kallenbach, N. Kanaya, M. Kawabata, Y. Kawabata, M. Kawaguti, S. Kelner, P. Kent, A. Kimura, T. Kodama, R. Kokoulin, M. Kossov, H. Kurashige, E. Lamanna, T. Lampén, V. Lara, V. Lefebure, F. Lei, M. Liendl, W. Lockman, F. Longo, S. Magni, M. Maire, E. Medernach, K. Minamimoto, P. M. de Freitas, Y. Morita, K. Murakami, M. Nagamatu, R. Nartallo, P. Nieminen, T. Nishimura, K. Ohtsubo, M. Okamura, S. O'Neale, Y. Oohata, K. Paech, J. Perl, A. Pfeiffer, M. Pia, F. Ranjard, A. Rybin, S. Sadilov, E. D. Salvo, G. Santin, T. Sasaki, N. Savvas, Y. Sawada, S. Scherer, S. Sei, V. Sirotenko, D. Smith, N. Starkov,

H. Stoecker, J. Sulkimo, M. Takahata, S. Tanaka, E. Tcherniaev, E. S. Tehrani, M. Tropeano, P. Truscott, H. Uno, L. Urban, P. Urban, M. Verderi, A. Walkden, W. Wander, H. Weber, J. Wellisch, T. Wenaus, D. Williams, D. Wright, T. Yamada, H. Yoshida, D. Zschiesche, Geant4—a simulation toolkit, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 506 (3) (2003) 250–303. doi:10.1016/s0168-9002(03)01368-8.
URL https://doi.org/10.1016/s0168-9002(03)01368-8

[17] T. Goorley, M. James, T. Booth, F. Brown, J. Bull, L. J. Cox, J. Durkee, J. Elson, M. Fensin, R. A. Forster, J. Hendricks, H. G. Hughes, R. Johns, B. Kiedrowski, R. Martz, S. Mashnik, G. McKinney, D. Pelowitz, R. Prael, J. Sweezy, L. Waters, T. Wilcox, T. Zukaitis, Initial MCNP6 release overview, Nuclear Technology 180 (3) (2012) 298–315. doi:10.13182/nt11-135.
URL https://doi.org/10.13182/nt11-135

[18] T. Bohlen, F. Cerutti, M. Chin, A. Fasso, A. Ferrari, P. Ortega, A. Mairani, P. Sala, G. Smirnov, V. Vlachoudis, The FLUKA Code: Developments and Challenges for High Energy and Medical Applications, Nuclear Data Sheets 120 (2014) 211 – 214. doi:10.1016/j.nds.2014.07.049.

[19] A. Ferrari, P. Sala, A. Fasso, J. Ranft, FLUKA: a multi-particle transport code, CERN-2005-10 INFN/TC05/11 SLAC-R-773 (Oct. 2005).

[20] Z. Li, Y. Fan, M. Dong, L. Tong, L. Zhao, Y. Yin, X. Chen, In-Beam PET Imaging in Carbon Therapy for Dose Verification, IEEE Transactions on Radiation and Plasma Medical Sciences 2 (1) (2018) 61–67. doi:10.1109/trpms.2017.2769109.
URL https://doi.org/10.1109/trpms.2017.2769109

[21] K. Parodi, On the feasibility of dose quantification with in-beam PET data in radiotherapy with 12C and proton beams , PhD Dissertation Technische Universitat Dresden (2004).

[22] G. Akamatsu, H. Tashima, Y. Iwao, H. Wakizaka, T. Maeda, A. Mohammadi, S. Takyu, M. Nitta, F. Nishikido, H. Rutherford, A. Chacon, M. Safavi-Naeini, E. Yoshida, T. Yamaya, Performance evaluation of a whole-body prototype PET scanner with four-layer DOI detectors, Physics in Medicine and Biology (Apr. 2019). doi:10.1088/1361-6560/ab18b2.
URL https://doi.org/10.1088/1361-6560/ab18b2

[23] A. Lechner, V. Ivanchenko, J. Knobloch, Validation of recent Geant4 physics models for application in carbon ion therapy, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms 268 (14) (2010) 2343–2354. doi:10.1016/j.nimb.2010.04.008.
URL https://doi.org/10.1016/j.nimb.2010.04.008

[24] M. D. Napoli, C. Agodi, G. Battistoni, A. A. Blancato, G. A. P. Cirrone, G. Cuttone, F. Giacoppo, M. C. Morone, D. Nicolosi, L. Pandola, V. Patera, G. Raciti, E. Rapisarda, F. Romano, D. Sardina, A. Sarti, A. Sciubba, V. Scuderi, C. Sfienti, S. Tropea, Carbon fragmentation measurements and validation of the Geant4 nuclear reaction models for hadrontherapy, Physics in Medicine and Biology 57 (22) (2012) 7651–7671. doi:10.1088/0031-9155/57/22/7651.
URL https://doi.org/10.1088/0031-9155/57/22/7651

[25] E. Haettner, H. Iwase, M. Krämer, G. Kraft, D. Schardt, Experimental study of nuclear fragmentation of 200 and 400 MeV/u12C ions in water for applications in particle therapy, Physics in Medicine and Biology 58 (23) (2013) 8265–8279. doi:10.1088/0031-9155/58/23/8265.
URL https://doi.org/10.1088/0031-9155/58/23/8265

[26] T. T. Böhlen, F. Cerutti, M. Dosanjh, A. Ferrari, I. Gudowska, A. Mairani, J. M. Quesada, Benchmarking nuclear models of FLUKA and GEANT4 for carbon ion therapy, Physics in Medicine and Biology 55 (19) (2010) 5833–5847. doi:10.1088/0031-9155/55/19/014.
URL https://doi.org/10.1088/0031-9155/55/19/014

[27] J. Dudouet, D. Cussol, D. Durand, M. Labalme, Benchmarking geant4 nuclear models for hadron therapy with 95 MeV/nucleon carbon ions, Physical Review C 89 (5) (May 2014). doi:10.1103/physrevc.89.054616.
URL https://doi.org/10.1103/physrevc.89.054616

[28] D. Bolst, G. A. Cirrone, G. Cuttone, G. Folger, S. Incerti, V. Ivanchenko, T. Koi, D. Mancusi, L. Pandola, F. Romano, A. Rosenfeld, S. Guatelli, Validation of Geant4 fragmentation for Heavy Ion Therapy, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 869 (2017) 68–75. doi:10.1016/j.nima.2017.06.046.
URL `https://doi.org/10.1016/j.nima.2017.06.046`

[29] K. Parodi, F. Ponisch, W. Enghardt, Experimental study on the feasibility of in-beam PET for accurate monitoring of proton therapy, IEEE Transactions on Nuclear Science 52 (3) (2005) 778–786. doi:10.1109/tns.2005.850950.
URL `https://doi.org/10.1109/tns.2005.850950`

[30] F. Pönisch, K. Parodi, B. G. Hasch, W. Enghardt, The modelling of positron emitter production and PET imaging during carbon ion therapy, Physics in Medicine and Biology 49 (23) (2004) 5217–5232. doi:10.1088/0031-9155/49/23/002.
URL `https://doi.org/10.1088/0031-9155/49/23/002`

[31] M. Priegnitz, F. Fiedler, D. Kunath, K. Laube, K. Parodi, F. Sommerer, W. Enghardt, A novel approach for predicting the positron emitter distributions produced during therapeutic ion irradiation, in: 2008 IEEE Nuclear Science Symposium Conference Record, IEEE, 2008. doi:10.1109/nssmic.2008.4774294.
URL `https://doi.org/10.1109/nssmic.2008.4774294`

[32] M. Priegnitz, F. Fiedler, D. Kunath, K. Laube, W. Enghardt, An Experiment-Based Approach for Predicting Positron Emitter Distributions Produced During Therapeutic Ion Irradiation, IEEE Transactions on Nuclear Science 59 (1) (2012) 77–87. doi:10.1109/tns.2011.2172629.
URL `https://doi.org/10.1109/tns.2011.2172629`

[33] C. Robert, G. Dedes, G. Battistoni, T. T. Böhlen, I. Buvat, F. Cerutti, M. P. W. Chin, A. Ferrari, P. Gueth, C. Kurz, L. Lestand, A. Mairani, G. Montarou, R. Nicolini, P. G. Ortega, K. Parodi, Y. Prezado, P. R. Sala, D. Sarrut, E. Testa, Distributions of secondary particles in proton and carbon-ion therapy: a comparison between GATE/Geant4 and FLUKA Monte Carlo codes, Physics in Medicine and Biology 58 (9) (2013) 2879–2899. doi:10.1088/0031-9155/58/9/2879.
URL `https://doi.org/10.1088/0031-9155/58/9/2879`

[34] I. Pshenichnov, I. Mishustin, W. Greiner, Distributions of positron-emitting nuclei in proton and carbon-ion therapy studied with GEANT4, Physics in Medicine and Biology 51 (23) (2006) 6099–6112. doi:10.1088/0031-9155/51/23/011.
URL `https://doi.org/10.1088/0031-9155/51/23/011`

[35] I. Pshenichnov, I. Mishustin, W. Greiner, MCHIT - Monte Carlo model for proton and Heavy-Ion Therapy, in: ND2007, EDP Sciences, 2007. doi:10.1051/ndata:07214.
URL `https://doi.org/10.1051/ndata:07214`

[36] F. Sommerer, F. Cerutti, K. Parodi, A. Ferrari, W. Enghardt, H. Aiginger, In-beam PET monitoring of mono-energetic16O and12C beams: experiments and FLUKA simulations for homogeneous targets, Physics in Medicine and Biology 54 (13) (2009) 3979–3996. doi:10.1088/0031-9155/54/13/003.
URL `https://doi.org/10.1088/0031-9155/54/13/003`

[37] L. Andy, C. Yong, A. Salahuddin, Yields of positron and positron emitting nuclei for proton and carbon ion radiation therapy: A simulation study with GEANT4, Journal of X-Ray Science and Technology 20 (3) (2012) 317–329. doi:10.3233/XST-2012-0340.
URL `http://doi.org/10.3233/XST-2012-0340`

[38] G. Collaboration, Physics reference manual for geant4, Tech. rep. (2018).

[39] D. Mancusi, A. Boudard, J. Cugnon, J.-C. David, P. Kaitaniemi, S. Leray, Extension of the Liège intranuclear-cascade model to reactions induced by light nuclei, Phys. Rev. C 90 (2014) 054602. doi:10.1103/PhysRevC.90.054602.
URL `https://link.aps.org/doi/10.1103/PhysRevC.90.054602`

[40] A. Mohammadi, E. Yoshida, H. Tashima, F. Nishikido, T. Inaniwa, A. Kitagawa, T. Yamaya,

Production of an 15 o beam using a stable oxygen ion beam for in-beam PET imaging, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 849 (2017) 76–82. doi:10.1016/j.nima.2016.12.028. URL `https://doi.org/10.1016/j.nima.2016.12.028`

[41] D. Marquardt, An Algorithm for Least-Squares Estimation of Nonlinear Parameters, Journal of the Society for Industrial and Applied Mathematics 11 (2) (1963) 431–441. doi:10.1137/0111030. URL `https://doi.org/10.1137/0111030`

7. Supplementary material

7.1. Validation of TAC fitting method

In order to evaluate the Levenberg-Marquardt error minimisation algorithm for the fitting of Equation (1), 1000 time activity curves (TAC) were generated with initial weights randomly generated using ^{11}C ^{10}C and ^{15}O half lives. An additional component, with a half life of 5 seconds, was generated to approximately account for short lived positron emitters. This additional component was not used in the fitting but was used when the random TAC was generated. The timing sampling points were chosen to be the same as the experimental values (refer to Section 3.2). The initial weights were generated in order to achieve a total weight of 100 and according to the following conditions:

- ^{11}C had an initial weight between 30 and 80.
- The additional component of half life of 5 seconds had an initial weight of less than 1.
- ^{10}C had an initial weight between 1 and 5.
- ^{15}O had the remaining weight to add up to 100.

The fitting of the TACs followed the same procedure as detailed in Section 3.2). On average, ^{11}C had a fitting error of 2% , ^{10}C had a fitting error of 8%, ^{15}O had a fitting error of 1.5% of the initial weight value.