

GATE as a GEANT4-based Monte Carlo platform for the evaluation of proton pencil beam scanning treatment plans

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2012 Phys. Med. Biol. 57 4223

(<http://iopscience.iop.org/0031-9155/57/13/4223>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 195.220.108.5

The article was downloaded on 11/06/2012 at 09:21

Please note that [terms and conditions apply](#).

GATE as a GEANT4-based Monte Carlo platform for the evaluation of proton pencil beam scanning treatment plans

L Grevillot^{1,2,3}, D Bertrand³, F Dessy³, N Freud^{1,2} and D Sarrut^{1,2}

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

² Centre Léon Bérard, Université de Lyon, Lyon, France

³ IBA, B-1348, Louvain-la-Neuve, Belgium

E-mail: loic.grevillot@creatis.insa-lyon.fr

Received 26 September 2011, in final form 23 April 2012

Published 8 June 2012

Online at stacks.iop.org/PMB/57/4223

Abstract

Active scanning delivery systems take full advantage of ion beams to best conform to the tumor and to spare surrounding healthy tissues; however, it is also a challenging technique for quality assurance. In this perspective, we upgraded the GATE/GEANT4 Monte Carlo platform in order to recalculate the treatment planning system (TPS) dose distributions for active scanning systems. A method that allows evaluating the TPS dose distributions with the GATE Monte Carlo platform has been developed and applied to the XiO TPS (Elekta), for the IBA proton pencil beam scanning (PBS) system. First, we evaluated the specificities of each dose engine. A dose-conversion scheme that allows one to convert dose to medium into dose to water was implemented within GATE. Specific test cases in homogeneous and heterogeneous configurations allowed for the estimation of the differences between the beam models implemented in XiO and GATE. Finally, dose distributions of a prostate treatment plan were compared. In homogeneous media, a satisfactory agreement was generally obtained between XiO and GATE. The maximum stopping power difference of 3% occurred in a human tissue of 0.9 g cm^{-3} density and led to a significant range shift. Comparisons in heterogeneous configurations pointed out the limits of the TPS dose calculation accuracy and the superiority of Monte Carlo simulations. The necessity of computing dose to water in our Monte Carlo code for comparisons with TPSs is also presented. Finally, the new capabilities of the platform are applied to a prostate treatment plan and dose differences between both dose engines are analyzed in detail. This work presents a generic method to compare TPS dose distributions with the GATE Monte Carlo platform. It is noteworthy that GATE is also a convenient tool for imaging applications, therefore opening new research possibilities for the PBS modality.

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

1. Introduction

A medical physicist's main task is to deliver the right dose to the right location. Therefore, quality assurance is of primary importance, especially in regard to the delivery system and the treatment planning system (TPS). In this report, we address the latter by comparing TPS pencil beam scanning (PBS) dose distributions with Monte Carlo simulations. Patient-specific quality assurance is routinely performed by comparing patient treatment plans delivered in water with measurements (Lomax *et al* 2004). The overall TPS validation procedure is however more complex (Jäkel *et al* 2001). Validating a treatment plan in water does not fully guarantee the dose distribution accuracy within the patient. Today, the Monte Carlo method is the most accurate possibility of recomputing TPS treatment plans based on patient CT images, with the advantage of providing a very detailed beam-interaction simulation in the patient. In the past decade, treatments have been mostly delivered using passive scattering delivery techniques and many studies using the Monte Carlo method have been performed with GEANT4 (Paganetti *et al* 2008, Bednarz *et al* 2010). This work is focused on active scanning, which is the most advanced delivery technique and allows higher dose conformity to the tumor while delivering a lower dose to surrounding healthy tissues (Lomax *et al* 1999). The GATE Monte Carlo simulations are not limited to dosimetric aspects and allow, e.g., β^+ or prompt- γ imaging investigations. Faster Monte Carlo codes for proton therapy applications exist (Fippel and Soukup 2004). The main interest of a slower and detailed Monte Carlo simulation, such as GATE/GEANT4, is its versatility: it can be used as a reference Monte Carlo code for validation purposes, and it allows combined imaging and dose simulations, micro-dosimetric applications, radiotherapy simulations using other particles, such as carbon ions, etc. Details about the GATE capabilities are presented elsewhere (Jan *et al* 2004, 2011). The active scanning technique is expected to be used more widely in the coming years due to technological improvements as well as the development of new proton and carbon therapy facilities worldwide (Dosanjh *et al* 2010). The combination of thousands of individual pencil beams allows for very complex dose distributions, especially when using intensity modulated proton therapy (Lomax *et al* 2004). Therefore, the use of a Monte Carlo code as quality assurance tool to benchmark TPSs for active scanning delivery becomes even more attractive. Since the GATE V6.0 release, the platform has allowed for radiation therapy and dosimetric applications (Jan *et al* 2011, Grevillot *et al* 2011a). The selection of the appropriate physics models and parameters leading to a reference *physics list* together with an optimized *parameters list* has been detailed in Grevillot *et al* (2010). A proton PBS model has been developed and validated for an IBA system (Grevillot *et al* 2011b). In this work, we used the following releases: GATE V6.1 and GEANT4.9.4p01 (Agostinelli *et al* 2003, Allison *et al* 2006). As TPSs calculate the so-called *dose to water* and Monte Carlo codes the *dose to medium* (Paganetti 2009), a method for converting dose to medium into dose to water is described in section 2.1. We investigate in section 2.2 the specificities of both XiO and GATE dose engines. Stopping power differences between GATE and XiO are presented in section 3.1. In section 3.2, the dose distributions of pristine Bragg peaks with different energies calculated using XiO and GATE in various homogeneous configurations are compared. A complex treatment plan in a homogeneous medium is used to validate both dose engines with appropriate measurements in section 3.3. In section 3.4, dose distributions in heterogeneous phantoms are analyzed and the limits of the TPS dose calculation algorithms on dose accuracy are investigated. Finally, a prostate treatment plan is presented and dose differences between XiO and GATE are analyzed in detail in section 3.5. This paper intends to present a generic method for PBS treatment plan evaluation with GATE. It points out various sources of differences between MC and TPS dose calculations, from simple cases in homogeneous

media to complex cases in heterogeneous configurations. The prostate treatment plan studied illustrates the possibilities of the GATE platform for clinical applications. The evaluation of various and more complex treatment plans will be the subject of further studies. This work is directed toward the IBA PBS system together with the XiO TPS; however, the overall method is expected to be useable with other TPSs and irradiation systems (Grevillot *et al* 2011b).

2. Materials and methods

Different definitions of range can be found in the literature, depending on physical or clinical constraints. In this paper, the physical and clinical ranges correspond to the 80% and 90% distal dose levels in the Bragg peak, respectively. Unless otherwise specified, beam ranges given in unit of g cm^{-2} correspond to ranges in water. The expression *spot size* will always refer to the Gaussian standard deviation of the size of the beam spot. Dose distributions calculated with GATE were converted into the DICOM format using a home-made toolkit, which is based on the ITK library⁴, in order to ensure compatibility with commercial dose analyzing tools. Simulations were performed either locally on a 4-CPU computer or on the EGEE grid, which provides about 40 000 CPUs through 250 resource centers worldwide (Camarasu-Pop *et al* 2010).

2.1. Dose to water versus dose to medium

Monte Carlo tools compute the so-called *dose to medium*. Historically, treatment planning systems have computed *dose to water* (or *water-equivalent dose*), by rescaling depth-dose distributions measured in water using a water-equivalent path length approximation. Differences between *dose to medium* and *dose to water* calculated with the Monte Carlo method and the TPS, respectively, were first investigated for photons (Siebers *et al* 2000) and later for protons (Paganetti 2009). Unlike photons, protons undergo non-elastic nuclear interactions. The impact of such interactions on the dose-conversion scheme was investigated by Palmans and Verhaegen (2005). In soft tissues, the water to medium dose ratio (D_w/D_m) is close to unity. For photons, as well as for protons, the largest difference in human tissues occurs in high-density bony structures, with differences between D_w and D_m of up to about 10% (Siebers *et al* 2000, Paganetti 2009). The ‘true’ dose delivered to each voxel in the patient is the dose to medium calculated by the Monte Carlo method. However, to evaluate the accuracy of TPS dose calculation algorithms, it becomes necessary to ‘degrade’ the dose to medium into dose to water. The question of whether reporting dose to medium or dose to water should be preferred was discussed some years ago (Liu 2002). In this work, we implemented a dose-conversion scheme similar to D_w^B from Paganetti (2009), which accounts on-the-fly for the energy- and particle-dependent relative mass stopping powers, as defined in equation (1). When all nuclear fragments are not explicitly tracked, an additional term accounting for relative non-elastic nuclear cross sections can be added (Palmans and Verhaegen 2005, Paganetti 2009). As in GATE every particle is tracked, this additional correction can be omitted:

$$D_w = [D_m \times S_{w,m}(E)]_{\text{proton}} + [D_m \times S_{w,m}(E)]_{\text{electron}} + [D_m \times S_{w,m}(E)]_{\text{others}}, \quad (1)$$

where E is the kinetic energy of the particle, ‘others’ refer to other secondary charged particles produced in the simulation and $S_{w,m}$ is the inverse of the relative mass stopping power ($S_{m,w}$), as defined in equation (2). The indices m and w refer to medium and water, respectively,

$$S_{m,w}(E) = \frac{1}{S_{w,m}(E)} = \frac{\frac{1}{\rho_m} \left(\frac{dE}{dx} \right)_m}{\frac{1}{\rho_w} \left(\frac{dE}{dx} \right)_w}, \quad (2)$$

⁴ <http://www.kitware.com/>

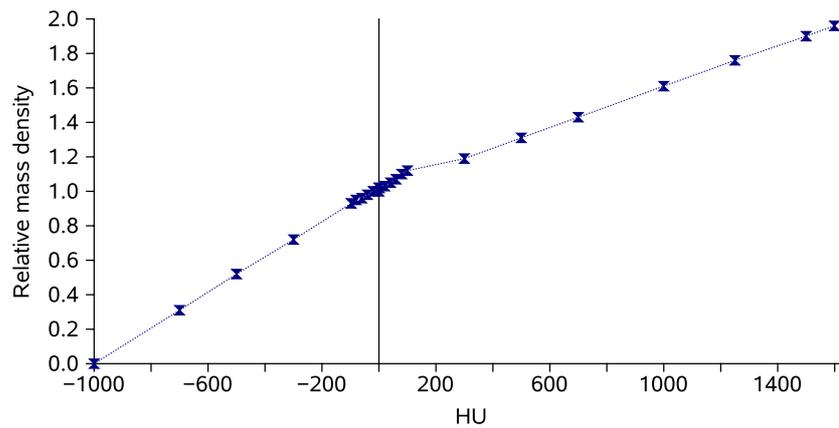


Figure 1. HU to relative mass density calibration curve used in XiO.

where $\frac{dE}{dx}$ is the stopping power and ρ is the mass density. The stopping power values of the different particles in the different media were calculated at each step using the ‘G4EmCalculator’ class from GEANT4. For some particles, these values are not accessible. In these cases, the mass stopping power ratio used for ‘others’ was set to be equal to that of a 100 MeV proton. It is noteworthy that the simplest conversion scheme is applied retroactively at the end of the calculation (Paganetti 2009), therefore neglecting the stopping power energy and particle dependences. The proposed dose-conversion tool presented will be publicly available in the next GATE release.

2.2. GATE and XiO dose engine specificities

2.2.1. Dealing with Hounsfield units.

For patient dose calculations, Hounsfield units (HUs) are converted into tissue compositions and mass densities for Monte Carlo codes and into relative stopping powers for TPSs. A reference method has shown how to interpolate tissue composition and density for every HU number and for a given scanner (Schneider *et al* 2000). In a second step, authors considered HU uncertainties and divided the HU scale $[-1000, +1600]$ into 24 different material compositions, assuming that each voxel can be associated with a mass density. In GATE/GEANT4, all materials must be defined prior to the simulation. Therefore, a *density tolerance parameter* expressed in g cm^{-3} was introduced into GATE, in order to specify the piecewise correspondence between the HUs and the densities (Jan *et al* 2011). We set a density tolerance parameter of 0.001 g cm^{-3} , which led to the definition of 2847 materials. In XiO, the HU to material-relative stopping power conversion is based on a HU to density calibration curve established by the user. The HU to density calibration curve has a resolution of 0.01 g cm^{-3} . An equation allows for converting mass density into relative mass stopping power (Fippel and Soukup 2004), as described later in section 2.2.2. The XiO calibration curve has been defined with 25 points, using the same HU to density correspondence as in GATE: 4 points in the HU interval $[-1000; -100]$, 12 points in HU $[-100; 100]$, 6 points in HU $[100; 1500]$ and 3 points in HU $[1500; 3000]$ (figure 1). As the HU to stopping power conversion methods used in GATE and XiO are different, the stopping powers associated with a given HU in both codes are different (but within the stopping power uncertainties).

2.2.2. TPS specificities. The pencil beam⁵ dose calculation algorithm implemented in XiO for PBS treatments is detailed in Soukup *et al* (2005) and Fippel and Soukup (2004), but a short summary is proposed in this subsection. A fit based on the ICRU'46, ICRU'49 and the computer tool PSTAR material data allows one to determine energy-dependent relative mass stopping powers as a function of material density (Fippel and Soukup 2004). Relative mass stopping powers allow scaling measured depth–dose profiles according to the radiological depth, as defined in Soukup *et al* (2005). Scaling methods used in TPSs provide water equivalent doses, as discussed in section 2.1. The multiple Coulomb scattering algorithm is based on the Rossi formula and allows computing the mean characteristic scattering angle (Gaussian standard deviation) (Soukup *et al* 2005). In addition to the initial *primary pencil beam*, the dose contribution from nuclear products is accounted for by means of a secondary *nuclear pencil beam* (Soukup *et al* 2005). Its contribution is described by a specific variance, which depends on the energy, spot size and radiological depth. It was modeled based on GEANT4 simulations. Spots are initially described via their energy, position, size, divergence and direction. They are considered to be elliptical Gaussians, with different parameters along the two lateral axes (Soukup *et al* 2005). The number of sub-spots (N) used in the implemented pencil beam algorithm is defined by the *precision parameter* (n), which varies between 0 and 5. The precisions $n = 0$, $n = 3$ and $n = 5$ correspond to numbers of sub-spots $N = 1$, $N = 49$ and $N = 121$, respectively. The larger the number of sub-spots, the better the heterogeneities can be accounted for (Soukup *et al* 2005). For homogeneous phantoms or heterogeneous sandwich configurations with slabs set perpendicularly to the beam axis, the precision parameter has no effect. In contrast, the precision parameter plays a major role when heterogeneities are adjacent to the beam axis. Unless otherwise specified, a precision $n = 0$, which corresponds to a pencil beam algorithm with a single ray tracing on the central axis was used by default. For heterogeneous geometries, different n values were tested. In addition to the CT matrix, a calculation grid with a user-defined resolution is created by re-sampling the CT matrix. The dose contribution from every pencil beam reaching a voxel is integrated over the entire voxel.

2.2.3. Monte Carlo code specificities. The GEANT4-based GATE Monte Carlo code allows for detailed simulations of particle interactions in matter (Apostolakis *et al* 2009). Details about the GEANT4 physics implementation are provided in Geant4 Collaboration (2009). The physics models and parameters selected for our application have been presented in Grevillot *et al* (2010) and the beam model of the IBA system has been presented in Grevillot *et al* (2011b). The overall description is briefly summarized in this subsection. The energy loss of protons and secondary ions is based on the Bethe–Bloch equation, except below 2 MeV, where parameterized data are used (Geant4 Collaboration 2009). Secondary particles, such as electrons, positrons and photons, are produced only if their range⁶ is larger than 1 mm and then tracked until they do not have any kinetic energy left (Grevillot *et al* 2010). The multiple Coulomb scattering algorithm implemented is based on the Lewis theory (Geant4 Collaboration 2009). It is a class II condensed algorithm, in which the global effect (lateral displacement, energy loss and secondary particle production) is computed at the end of a track segment. Non-elastic and elastic nuclear interactions are simulated in detail. The non-elastic nuclear model used is based on a compound nucleus theory with a pre-equilibrium stage. Pre-equilibrium and further equilibrium stages allow for producing and tracking charged and neutral secondaries. The beam model accounts for beam size, divergence and emittance

⁵ It is necessary to distinguish the expression *pencil beam* used to describe a dose calculation algorithm from the expression *PBS* used to describe a beam delivery technique.

⁶ For photons, the concept of range does not exist. Instead, an absorption length related to an absorption cross section is used in GEANT4.

parameters, with different values in each lateral direction, as described by Grevillot *et al* (2011b). The CT matrix is used as it is for particle-interaction simulation and the corresponding dose depositions are scored within a scoring grid with a resolution defined by the user. Dose contributions from each particle reaching a voxel are integrated over the entire voxel.

2.3. Detailed comparison method of GATE and XiO dose calculations

2.3.1. Stopping powers. Relative mass stopping powers calculated with XiO and GATE were compared at 100 MeV, using a set of materials within the density interval [0.0012, 1.96] (HU [-1000, +1600]). GATE stopping powers were calculated using the newly released *EmCalculatorActor* tool from GATE V6.1, which is based on the *G4EmCalculator* class from GEANT4 and allows extracting specific properties of all materials involved in a simulation.

2.3.2. Pristine Bragg peaks in homogeneous media. First, we compared pristine Bragg peaks obtained with GATE and XiO in water for three clinical ranges: 28.0, 18.0 and 7.7 g cm⁻². Second, we compared pristine Bragg peaks obtained with GATE and XiO in five different media with mass densities of 0.9, 1.0, 1.2, 1.5 and 1.8 g cm⁻³, using a fixed beam energy corresponding to a range of 25.0 g cm⁻² in water. Note that when defining a treatment plan, each beam is referred to by its range in water (in g cm⁻²). Water was defined in XiO by forcing the density of the phantom to unity. The water phantom surface was positioned at the isocenter. The resolution of the calculation grid was set to 2 mm in XiO and reproduced as such in GATE. Integral depth-dose and transverse profiles at three depths were extracted from the 3D dose maps. We evaluated the physical and clinical ranges, as well as the mean point-to-point dose differences between GATE and XiO. Depth-dose and transverse profiles were normalized relative to the integral dose. Transverse profiles were fitted with Gaussian functions in order to extract the standard deviations. Maximum spot sizes were evaluated at a depth corresponding to 98% of the proton physical range. The method used for such comparisons was presented by Grevillot *et al* (2010, 2011b).

2.3.3. Beam models validation in homogeneous media. The beam model implemented in GATE has already been partially validated in a previous study (Grevillot *et al* 2011b). Similar tests have been also performed at IBA to validate the beam model implemented in XiO. In this section, additional comparisons including measurements as well as XiO and GATE calculations in a homogeneous phantom are proposed. A test treatment plan was generated with XiO in a water-equivalent phantom of 30 × 30 × 32 cm³. The treatment plan was made up of a single field composed of seven iso-energy layers modulated between 27 and 32 cm, with 1924 spots and a spot iso-spacing of 5 mm. The isocenter was set at 30 cm depth. The modulated region had a triangular shape with a homogeneous dose distribution as shown in figure 2. The phantom was made up of water equivalent RW3 slabs of 0.5–1 cm thickness, with a density of 1.045 g cm⁻³ (Solid Phantom SP34[®], IBA-Dosimetry). The CT images of the phantom were acquired and used for the dose calculation. The I^mRT MatriXX[®] tool (IBA-Dosimetry) was used for the measurements. This tool consists of a 2D matrix of 1020 ionization chambers with an active volume of 0.08 cm³ for each chamber and a resolution of 7.62 mm between two measuring points. Measurements were carried out at 5.0, 14.0, 20.0, 27.2, 28.7 and 30.2 cm depths, by placing the MatriXX tool below the appropriate quantity of SP34 and by virtually reproducing the original setup. A 3D dose distribution was computed with the TPS using a 4 mm calculation grid resolution. 2D maps corresponding to the measurement depths were obtained using interpolation with a 1 mm resolution. The treatment plan and CT images exported from XiO were used for the Monte Carlo simulation, using the

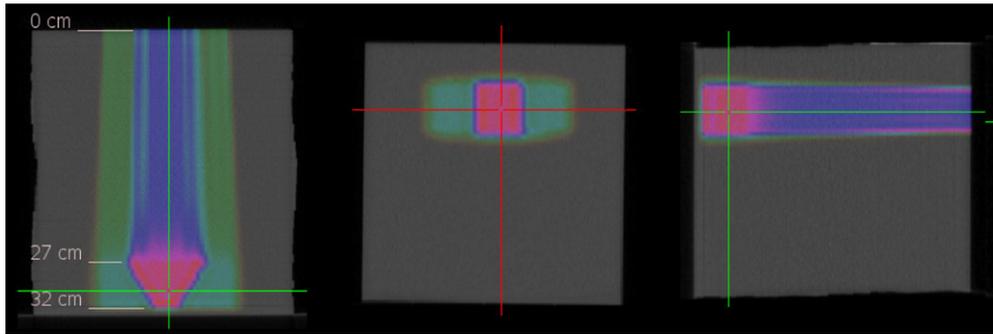


Figure 2. 3D dose distribution computed by XiO within the SP34 phantom. The coronal, transverse and sagittal views are represented on the left, middle and right parts of the picture, respectively.

same scoring grid resolution. Details about CT images and PBS treatment plan integration in GATE can be found elsewhere (Jan *et al* 2011, Grevillot *et al* 2011b). Additional 2D dose maps with a $2 \times 2 \times 2 \text{ mm}^3$ dosel (dose scoring voxel) resolution were scored in GATE at different depths in the phantom, for further comparisons with measurements and XiO. GATE and XiO dose maps were rescaled according to two normalization factors. Each factor was determined according to integral dose ratios calculated at 14 cm depth between measurement and calculations. At the time of this experiment, the CT calibration curve implemented in XiO was different from the one presented in section 2.2.1. Unfortunately, XiO does not allow recomputing dose distributions using a new calibration curve without re-optimizing the plan. We assumed that the only significant effect of using a different calibration curve is a shift of the dose distribution in depth (along the beam axis). This effect was partly compensated by comparing dose distributions at equivalent relative depths $z_{eq} = z/r_0$, with z_{eq} being a fraction of the beam range r_0 , at depth z . The main interest of this experiment was the evaluation of the 2D transverse dose distributions, based on gamma indices. Gamma indices were evaluated for all points receiving more than 0.2% of the maximum dose and compared between XiO, GATE and the measurements, using the OmniPro-I[®]mRT[®] software (IBA Dosimetry).

2.3.4. Pristine Bragg peaks in heterogeneous media. We compared GATE and XiO pristine Bragg peaks in heterogeneous configurations using sandwich phantoms. We differentiated *transverse sandwich configurations*, in which stacks of materials are placed perpendicularly to the beam axis, from *longitudinal sandwich configurations*, in which material interfaces are parallel to the beam axis, as illustrated in figure 3. For these tests, 2 mm scoring grid resolutions were selected.

We used two transverse sandwich phantoms (figure 3(a)). The first sandwich configuration was designed with small density variations, using four different media with densities of 1.08, 0.94, 1.09 and 1.22 g cm^{-3} . Thicknesses were 40, 5, 70 and 85 mm, respectively. We used a 18 g cm^{-2} proton beam (range in water). The second sandwich configuration was made up of six media, with densities of 1.09 (skin), 1.25 (sternum), 0.26 (lung), 1.00 (soft tissue), 1.90 (cortical bone) and 1.00 (soft tissue). Thicknesses were 10, 20, 50, 35, 10 and 20 mm, respectively. This second test case was designed to evaluate maximum D_w/D_m variations for the different human tissues. We used a 10 g cm^{-2} proton beam.

One longitudinal sandwich phantom was designed using a setup inspired from Soukup *et al* (2005). The phantom consisted of a 2 cm interface of adjacent bone ($d = 1.9$) and lung

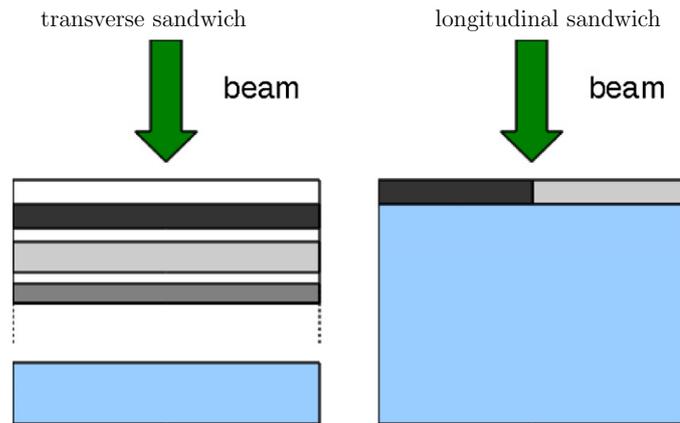


Figure 3. Schematic representation of the two types of sandwich configurations used. The different materials are represented by different shade of gray and water is in blue.



Figure 4. 3D dose distribution in the patient calculated with GATE and displayed with the visualization tool vv (Seroul and Sarrut 2008).

($d = 0.26$), parallel to the beam axis (figure 3(b)). A 25 g cm^{-2} pencil beam was used. In this test case, the precision parameter from XiO played an important role, as it determined the accuracy of the calculation algorithm in order to account for heterogeneities set parallel to the beam axis. Therefore, three levels of precision were tested: $n = 0$ (simple ray tracing), $n = 3$ (49 sub-spots) and $n = 5$ (121 sub-spots). In addition to the precision parameter, the sensitivity of the dose calculation to the scoring grid position was tested by shifting the calculation matrix perpendicularly to the beam axis.

2.4. Application to a prostate cancer treatment plan

We completed this study by evaluating dose distributions for a prostate treatment plan. CT images and structures were provided by Elekta. The gross tumor volume (GTV) corresponded to the prostate and the planned target volume (PTV) was defined as the volume enclosing the prostate and the seminal vesicle. A dose of 80 Gy was planned in the PTV using two laterally opposed fields. Maximum dose and dose–volume constraints were defined prior to the optimization for the following organs at risk (OARs): femurs (left and right), rectum and bladder. Each field was optimized separately, in order to deliver a homogeneous dose to the tumor (figure 4), using the so-called single-field uniform dose optimization technique (Lomax 1999).

A dose scoring grid resolution of 2 mm was used in XiO and reproduced in the GATE simulation. The simulation was run on the EGEE grid (Camarasu-Pop *et al* 2010) and very small statistical fluctuations were achieved: below 0.5% (standard deviation) in high-dose regions. Comparisons were carried out using a relative dose rather than the absolute dose, because the TPS was not calibrated for absolute dosimetry. Simulated dose distributions were normalized relative to the XiO integral dose. We evaluated the influence of the XiO precision parameter using three different values: $n = 0$, $n = 3$ and $n = 5$. The influence of the dose to medium and dose to water computations with GATE were also investigated. The ARTIVIEW software (AQUILAB) was used to compare the different dose matrices and to extract various clinical data:

- dose–volume histograms (DVHs),
- iso-dose volume index (for instance, the iso-dose volume index V_{20} corresponds to the fraction of the organ volume receiving at least 20% of the prescribed dose),
- near maximum dose (D_{nMax}): maximum dose received by 2% of an organ,
- average dose (D_{av}),
- near minimum dose (D_{nMin}): minimum dose received by 98% of an organ.

Comparisons were performed for each lateral field separately and for the complete treatment plan.

3. Results

3.1. Stopping power differences between GATE and XiO

Relative mass stopping powers calculated with XiO are lower than with GATE (table 1), except for mass densities within the interval [0.5, 0.9], for which larger discrepancies are observed. If we neglect the interval [0.5, 0.9], relative mass stopping powers calculated with XiO are systematically lower than with GATE, which could be partly explained by a lower water stopping power in GATE than in XiO. In ICRU'49 (ICRU 1993), stopping powers are stated to be accurate to within 1–2% for elements and 1–4% for compounds; therefore, the differences observed between XiO and GATE are acceptable. For a mass density of 1.00 g cm^{-3} , the relative stopping power calculated in XiO is 1.00, because the corresponding material is water. In GATE, a human tissue composition different from water is defined (section 2.2.1), for which the relative stopping power is not 1.00, but 1.02.

3.2. Evaluation of pristine peaks in homogeneous media

3.2.1. *Pristine Bragg peaks in water.* GATE ranges are systematically larger than XiO ranges by about 0.5–0.6% (table 2), confirming the previous hypothesis of lower water stopping power in GATE than in XiO (section 3.1). Mean point-to-point dose differences are less than 2%. Transverse profiles computed at the isocenter were compared to the beam data library measurements of the system. Spot sizes (standard deviation) at the isocenter calculated with GATE and XiO agreed with measurements within 0.1 mm in the x - and y -directions for the three energies tested. Transverse dose profile standard deviations computed with GATE and XiO at 40%, 90% and 98% of the Bragg peak physical range were in agreement with less than 0.2 mm difference. Depth–dose and transverse profiles computed for the 28 g cm^{-2} beam are illustrated in figure 5. Differences on the Bragg peak shapes (figure 5(a)) can be due to different initial energy spread and to the energy straggling model in GEANT4.

Table 1. The GATE mass stopping powers (column 2), GATE and XiO relative mass stopping powers (column 3 and 4) and differences between GATE and XiO relative mass stopping powers (column 5), as a function of material mass density (column 1). Stopping power ratios are rounded to 1% and calculated for 100 MeV protons.

Mass density (g cm ⁻³)	$S_{m\text{GATE}}$ (MeV cm ² g ⁻¹)	$\frac{S_m}{S_w\text{GATE}}$	$\frac{S_m}{S_w\text{XiO}}$	$1 - \left(\frac{S_m}{S_w\text{GATE}} \times \frac{S_w}{S_m\text{XiO}} \right)$ (%)
0.0012	6.50	0.89	0.88	-1.4
0.26	7.30	1.00	0.99	-1.3
0.40	7.30	1.00	1.00	-0.3
0.50	7.30	1.00	1.01	+0.4
0.60	7.30	1.00	1.02	+1.0
0.70	7.30	1.00	1.02	+1.7
0.80	7.30	1.00	1.03	+2.3
0.90	7.30	1.00	1.04	+3.0
0.95	7.52	1.03	1.02	-1.3
1.00	7.42	1.02	1.00	-2.1
1.05	7.31	1.00	1.00	-0.9
1.10	7.25	1.00	0.99	-1.2
1.20	7.12	0.98	0.97	-1.4
1.30	7.05	0.97	0.95	-2.1
1.40	6.92	0.95	0.94	-1.8
1.50	6.79	0.93	0.92	-1.3
1.60	6.74	0.93	0.91	-1.9
1.70	6.65	0.92	0.90	-1.7
1.80	6.57	0.90	0.89	-1.6
1.90	6.53	0.90	0.88	-2.0
1.96	6.51	0.90	0.88	-1.6

Table 2. The integral depth-dose profile differences between GATE and XiO. r_{80} and r_{90} stand for physical and clinical ranges, respectively.

	r_{80} (mm)	r_{90} (mm)	Dose differences (%)	r_{90} differences (%)/(mm)
28.0 g cm ⁻²				
GATE	279.9	278.8	-	-
XiO	278.4	277.4	1.8	0.5 / 1.5
18.0 g cm ⁻²				
GATE	182.1	181.4	-	-
XiO	181.2	180.5	1.3	0.5 / 0.9
7.7 g cm ⁻²				
GATE	79.1	78.3	-	-
XiO	78.2	77.8	1.9	0.6 / 0.5

3.2.2. Pristine Bragg peaks in five different media. Range differences are presented in table 3. They are consistent with relative stopping power differences (table 1), when taking into account the difference in water stopping powers between XiO and GATE of 0.5% (table 2). However, at this energy, this led to range differences of 3.7 mm (1.5%) and 9.6 mm (3.5%) for human tissues having mass densities of 1.00 and 0.90 g cm⁻³, respectively. Obviously, the differences shown are the largest possible for a single tissue. Such range differences are very unlikely to happen in a human body, which is made up of numerous tissues. At the phantom entrance,

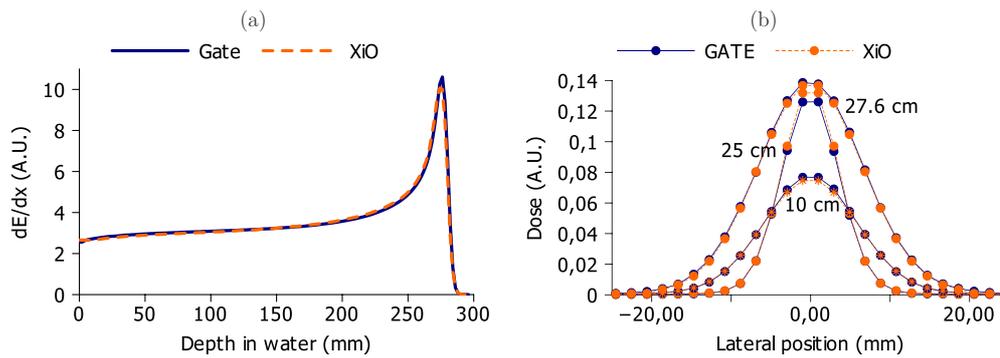


Figure 5. Comparison of GATE and XiO dose profiles in water for a 28.0 g cm^{-2} proton beam: (a) depth-dose profiles and (b) transverse dose profiles at 10.0, 25.0 and 27.6 cm depth. Lines are displayed to guide the eyes only.

Table 3. Comparison of depth-dose and transverse profiles for a 25.0 g cm^{-2} proton beam in various human media using GATE and XiO.

Density (g cm^{-3})	0.9	1	1.2	1.5	1.8
Ranges					
XiO (mm)	267.5	249.3	215.1	180.5	155.7
GATE (mm)	277.1	245.6	213.2	179.1	154.2
Difference (mm)	-9.6	3.7	1.9	1.4	1.5
Difference (%)	-3.5	1.5	0.9	0.8	1.0
Maximum spot sizes					
XiO (mm)	6.5	6.6	6.2	5.8	5.4
GATE (mm)	7.0	6.3	6.2	6.0	5.7

Table 4. Gamma index comparisons between measurements, XiO and GATE, for 2D dose maps calculated at different depths.

Depth of measurement (mm)	302	287	272	200	140	50
3%/3 mm gamma						
GATE versus measurement (%)	98.2	98.3	97.6	98.7	98.4	96.7
XiO versus measurement (%)	96.6	99.5	99.2	99.1	98.8	98
2%/2 mm gamma						
XiO versus GATE (%)	95.4	98.9	98.8	99.2	98.9	98.2

the spot sizes are in reasonable agreement, with less than 0.2 mm variations. Maximum spot size discrepancies in the Bragg peak can be attributed to range differences, the use of different multiple Coulomb scattering algorithms and the differences in the beam models.

3.3. Validation of the GATE and XiO beam models against measurements

Gamma index comparisons with measurements using GATE and XiO are satisfactory, with more than 96% of the points passing the test with a 3%/3 mm criterion (table 4), therefore validating both calculation tools. More than 95% of the points passed the gamma comparison

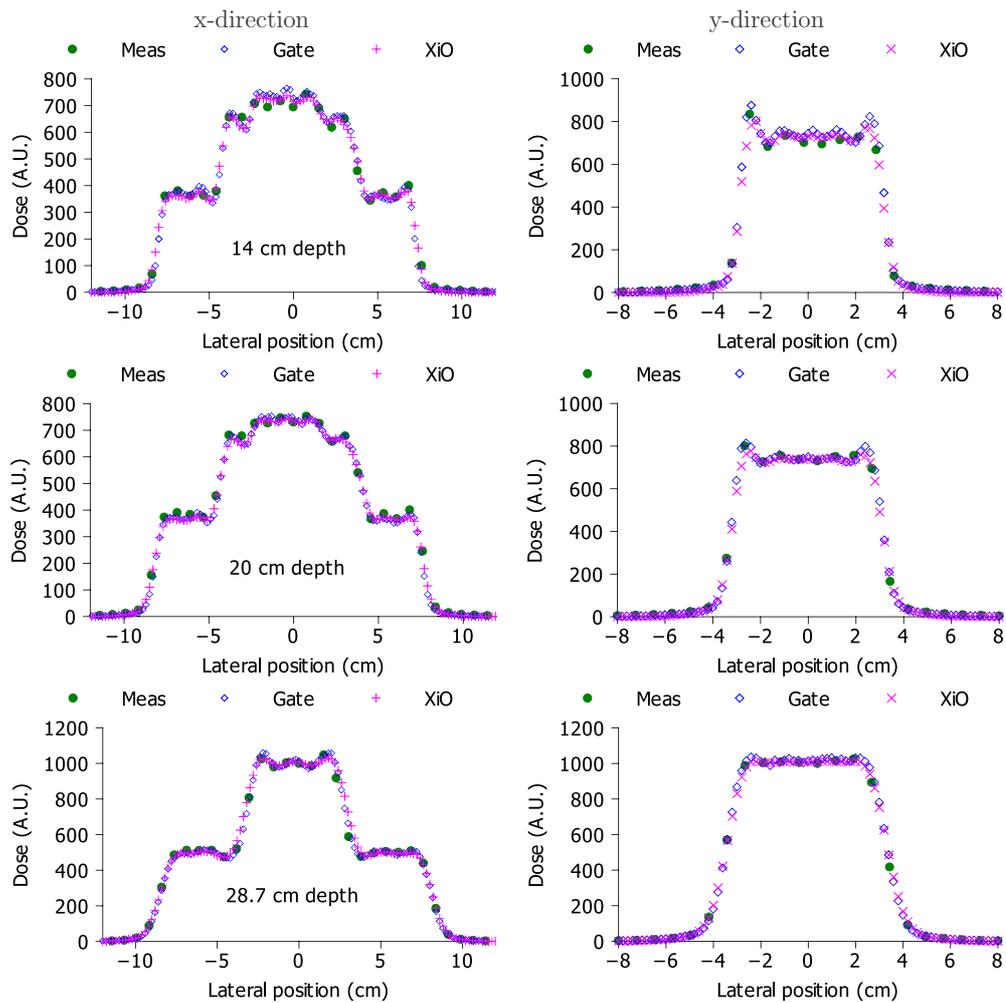


Figure 6. Transverse profile comparisons between measurements, GATE and XiO calculations in x - and y -directions for three depths.

between GATE and XiO using a 2%/2 mm criterion. XiO seems to agree slightly better than GATE with measurements, except for the 2D dose map at 30.2 cm depth. However, due to statistical fluctuations of about 1% and the fact that comparisons were performed at relative depths, the differences observed can be disregarded. A sample of transverse dose profiles is presented in figure 6.

3.4. Evaluation of pristine peaks in heterogeneous media

3.4.1. Pristine Bragg peaks in transverse sandwich configurations. In the first sandwich (figure 7(a)), the range computed by XiO was 2 mm (1.1%) larger than with GATE. In the media with densities of 0.94 and 1.22 g cm⁻³, dose differences of -3% and +2%, respectively, are seen between the dose to water and the dose to medium calculated by GATE. These differences can be explained using water to medium mass stopping power ratios, corresponding to the inverse of column 3 from table 1.

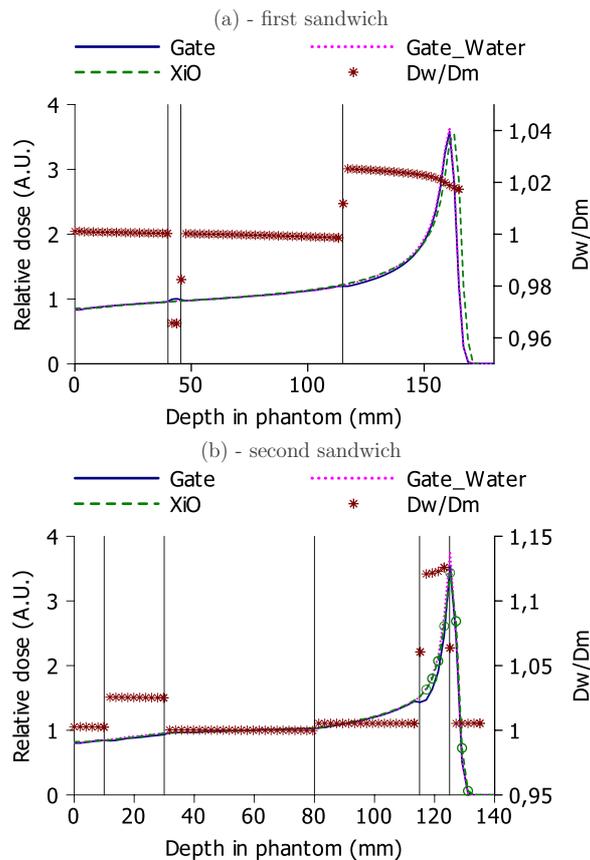


Figure 7. Depth-dose profiles computed in two sandwich configurations: (a) first sandwich and (b) second sandwich. The 2 mm scoring grid resolution is too large to allow a precise description of the Bragg peak, as illustrated by the additional green circles displayed in (b) in the Bragg peak and corresponding to the XiO calculated points. The D_w/D_m dose ratio computed with GATE is also displayed and corresponds to the right axis.

In the second sandwich (figure 7(b)), no significant range differences were noted. The largest difference arises in the high-density bony structure, with a D_w/D_m value of approximately 1.12. In contrast, in the low-density lung structure, $D_w/D_m = 1.00$. Water to medium dose differences are well marked in the second and fourth media of the first sandwich and in the second and fifth media of the second sandwich. The voxel resolution of 2 mm does not allow one to accurately calculate the dose deposition in the Bragg peak. This is more pronounced in the second phantom, where the Bragg peak stops at the distal bone/water interface, which makes the dose calculation uncertain. Mean point-to-point dose differences were lower than 2% for both sandwich configurations, when comparing GATE dose to water and XiO. Therefore, the differences between the two dose engines are clinically acceptable. Differences obtained between dose to water and dose to medium demonstrate the necessity of the conversion to evaluate TPS dose calculation algorithms with Monte Carlo codes.

3.4.2. Pristine Bragg peaks in longitudinal sandwich configurations. For this test case, two peaks are expected, corresponding to the fractions of beam crossing mainly bone or lung tissues (figure 8). With the precision $n = 0$, only one peak is produced, because the radiological depth

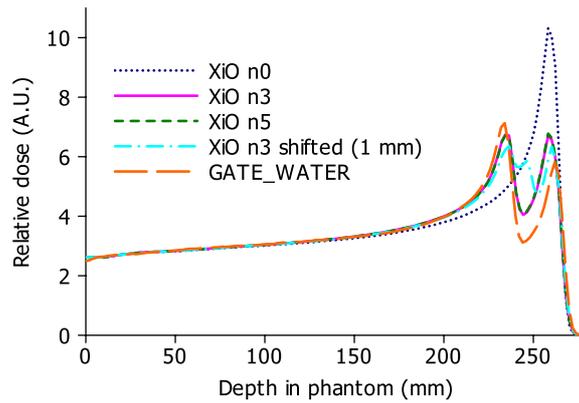


Figure 8. Depth–dose profiles obtained in the longitudinal sandwich configuration, using GATE and XiO with different levels of precision.

Table 5. Differences between XiO and GATE dose to water in the longitudinal phantom configuration.

	Bone peak (%)	Trough (%)	Lung peak (%)
XiO ($n = 3$) versus GATE	−6.0	+32.3	+15.5
XiO-shifted versus GATE	−12.7	+90.3	+8.6

is calculated through the lung tissue and the bony structure is completely neglected. Increasing the precision to $n = 3$ and $n = 5$ makes it possible to better account for the bone/lung interface and two peaks are calculated. No significant difference was shown between $n = 3$ and $n = 5$, suggesting no important improvement in the dose calculation. In such a case, Monte Carlo simulation is considered as the reference (Soukup *et al* 2005) and the limits of the pencil beam algorithm are pointed out. This corroborates similar studies performed by Soukup *et al* (2005) and Stankovskiy *et al* (2009). However, the differences shown are not only due to the pencil beam algorithm. Indeed, the CT re-sampling performed by the TPS to generate the calculation grid is also responsible for part of the difference. This is illustrated in figure 8, where the scoring grid has been shifted laterally by 1 mm (half the scoring grid resolution). With this shift, a third peak appears in between the lung and the bone peaks. It corresponds to a sampling artifact, which produced voxels of intermediate densities in between bone and lung densities. Such artifacts do not occur in our Monte Carlo implementation, because the CT image is not resampled. Differences between GATE dose to water and XiO in the ‘bone peak’ at 235 mm depth, in the ‘lung peak’ at 260 mm depth and in the ‘trough’ at 245 mm depth, are summarized in table 5. The tests performed in the longitudinal sandwich configuration clearly demonstrate the limits of the XiO analytical algorithms.

3.5. Evaluation of a prostate treatment plan

A detailed analysis is proposed for the right lateral field and only complementary information for the left field and for the complete treatment is presented.

3.5.1. Right lateral field. As a first verification, we evaluated the integral doses delivered along the left/right (depth–dose), anterior/posterior and head/feet axes (figure 9). These comparisons show a good overall agreement for the three axes, with a range shift of about

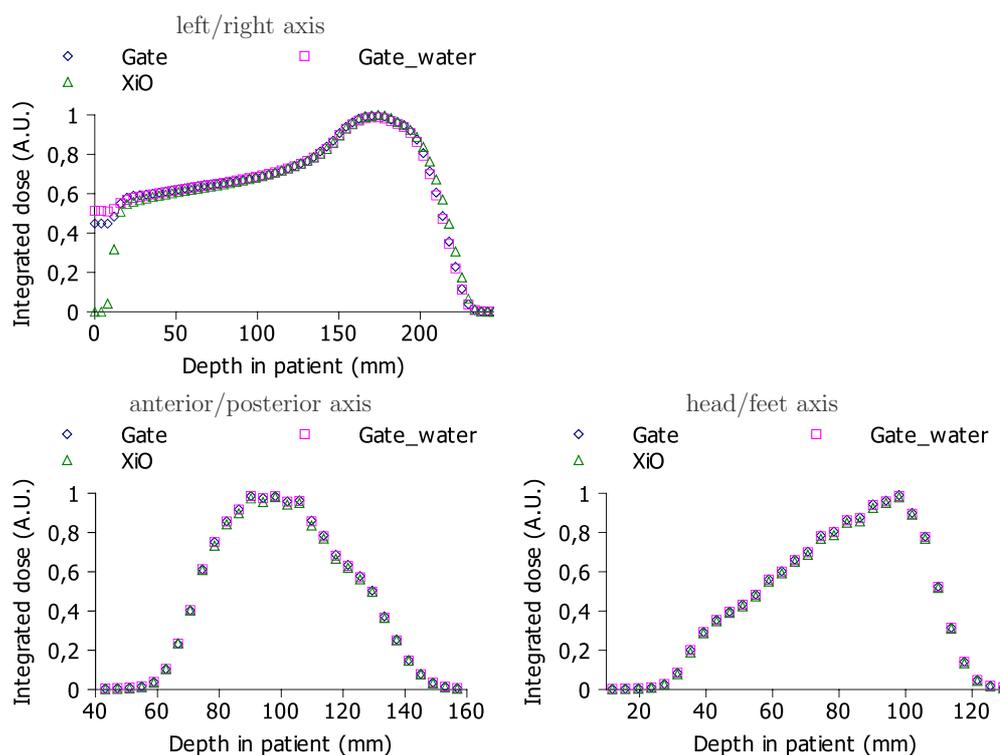


Figure 9. Integral dose profiles computed along the three axes using XiO (with $n = 0$) and GATE, for the right lateral field. In (a), dose differences within the first 20 mm are due to dose scoring artifacts in air: dose is set to 0 in XiO, but not in GATE.

2 mm (along the left/right axis), that will be analyzed later in this section. In a second step, we compared 2D dose distributions slice by slice. Some discrepancies were noted, with maximum differences visible in the CT slice 70 (figure 10). In this figure, the influence of the XiO precision parameter is evidenced: when n is increased from 0 to 5, a dose discrepancy (range shift) due to gas in the rectum becomes visible, but is less marked than with GATE. Using $n = 3$ produces similar iso-doses as with $n = 5$, suggesting no improvement on the dose calculation, as already noted in section 3.4.2. The dose difference between XiO ($n = 3$) and GATE is presented in figure 10(d). It shows an over dosage of about +10 Gy at the distal edge of the SOBP, because of the larger ranges computed in XiO than in GATE, except behind the rectum, where it shows an under dosage of about -10 Gy. The differences between GATE and XiO might also be due to the CT re-sampling necessary in XiO to create the calculation grid (section 3.4.2). In a third step, we focused on 1D dose distributions parallel and perpendicular to the field axis. Range shifts of about 1% (2 mm) are observed, but they vary slightly with the tissues crossed (figure 11(a)). In regard to the transverse profiles, a satisfactory agreement is obtained in the central area of the tumor: the mean dose difference calculated between 47 and 147 mm for the upper curve presented in figure 11(b) is 1.9%, with a maximum dose difference of 5% at the central point. Larger differences occur at the edge of the tumor, due to a combined effect of dose gradient and range shift. Differences between GATE dose to medium and dose to water reached 4% for some of the voxels compared.

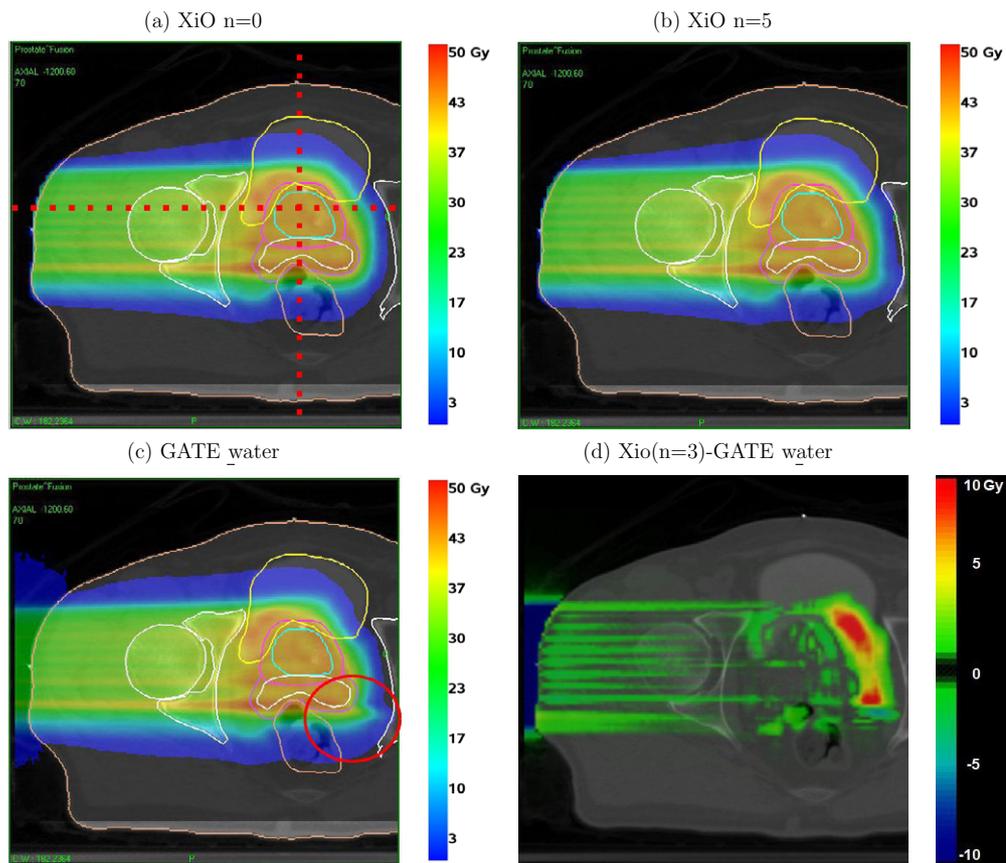


Figure 10. 2D dose distributions from GATE and XiO using $n = 0$ and $n = 5$. The red dotted lines in (a) represent two axes used to compare 1D dose distributions between XiO and GATE in figure 11. The dose discrepancy area due to gas in the rectum is circled in red in the GATE calculation in (c). A dose difference plot between XiO using $n = 3$ and GATE_water is presented in (d). For visualization purposes, a transparency scale is added to the color scale, so that dose differences close to 0 are transparent.

3.5.2. Left lateral field. Similar conclusions were drawn for this field, using the same test procedure. As for the right lateral field, a dose discrepancy occurred in the rectum, with maximum differences visible in the CT slice 70 (figure 12). A smaller range difference between GATE and XiO of about 1 mm (0.5%) was noted, instead of 2 mm for the right lateral field.

3.5.3. Complete treatment plan. The shorter ranges predicted by GATE for the two lateral fields led to tumor dose coverage shorter by about 1–2 mm on each side of the tumor, as illustrated in figure 13. The dose gradient between the tumor and the healthy tissues reaches the PTV border (pink contour, figure 13) in the GATE calculation, but the GTV (blue contour, figure 13) is still correctly covered. The dose discrepancies presented for both fields are less apparent when the two fields are combined; however, they are still clearly evidenced when computing the dose difference between XiO and GATE (figure 13(c)). Similarly, a recent evaluation of small proton fields delivered with a passive scattering system indicated

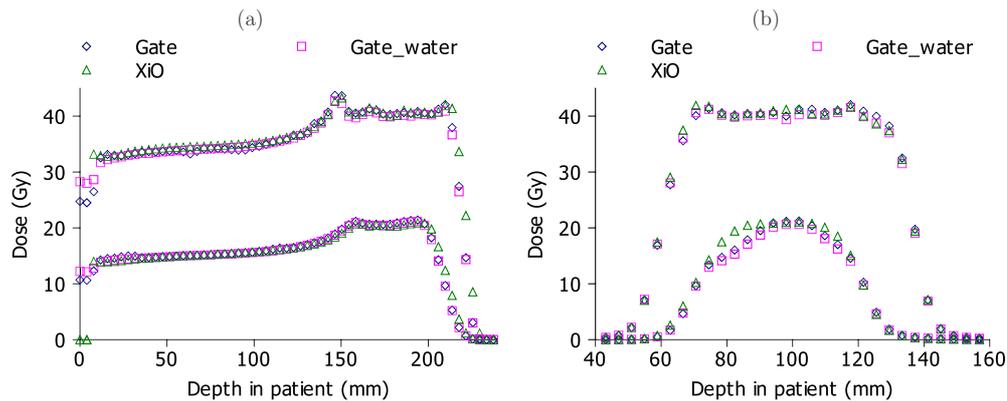


Figure 11. 1D dose distributions calculated with GATE and XiO ($n = 0$) for the right lateral field: (a) two depth–dose profiles and (b) two transverse profiles. In figures (a) and (b) the lower curves were rescaled by a factor of 1/2 for visualization purpose. Dose distributions were extracted from different CT slices; however, the approximate positions of two axes used to compare depth–dose and transverse profiles are presented in figure 10(a). In (a), dose differences within the first 10 mm are due to dose scoring artifacts in air, as previously explained in figure 9.

Table 6. Near-minimum, near-maximum and average doses calculated with GATE and XiO in the PTV, GTV and seminal vesicle. The V_{100} and V_{95} iso-dose volume indices are also presented. The results are calculated for the complete treatment plan (i.e. with both lateral fields).

Organ	GATE	GATE_water	XiO $n = 0$	XiO $n = 3$	XiO $n = 5$
PTV V_{100} (%)	85.3	68.5	82.3	80.1	80.0
PTV V_{95} (%)	96.9	94.9	98.0	97.8	97.8
PTV D_{nMax} (Gy)	85.2	83.3	82.9	82.8	82.9
PTV D_{av} (Gy)	81.5	80.2	80.7	80.6	80.6
PTV D_{nMin} (Gy)	74.9	73.8	76.1	75.8	75.8
GTV V_{100} (%)	96.4	82.5	88.7	88.6	88.8
GTV V_{95} (%)	100	100	100	100	100
GTV D_{nMax} (Gy)	83.9	82.9	82.4	82.4	82.5
GTV D_{av} (Gy)	81.7	80.8	80.8	80.8	80.9
GTV D_{nMin} (Gy)	79.7	78.9	79.2	79.2	79.2
Seminal vesicle V_{100} (%)	91.7	68.1	74.7	72.7	73.0
Seminal vesicle V_{95} (%)	99.7	98.7	99.4	100	100
Seminal vesicle D_{nMax} (Gy)	84.5	82.9	83.1	82.8	82.7
Seminal vesicle D_{av} (Gy)	81.7	80.5	80.7	80.6	80.6
Seminal vesicle D_{nMin} (Gy)	77.4	76.3	77.2	77.4	77.5

discrepancies for single fields, with hot and cold spots, but it was noted that the physical limits of the pencil beam algorithm appeared to cancel out with multiple fields (Bednarz *et al* 2010).

4. Discussion

We compared XiO and GATE DVHs for all contoured organs from the prostate treatment plan (figure 14(a)). The corresponding data presenting the main differences are summarized in table 6.

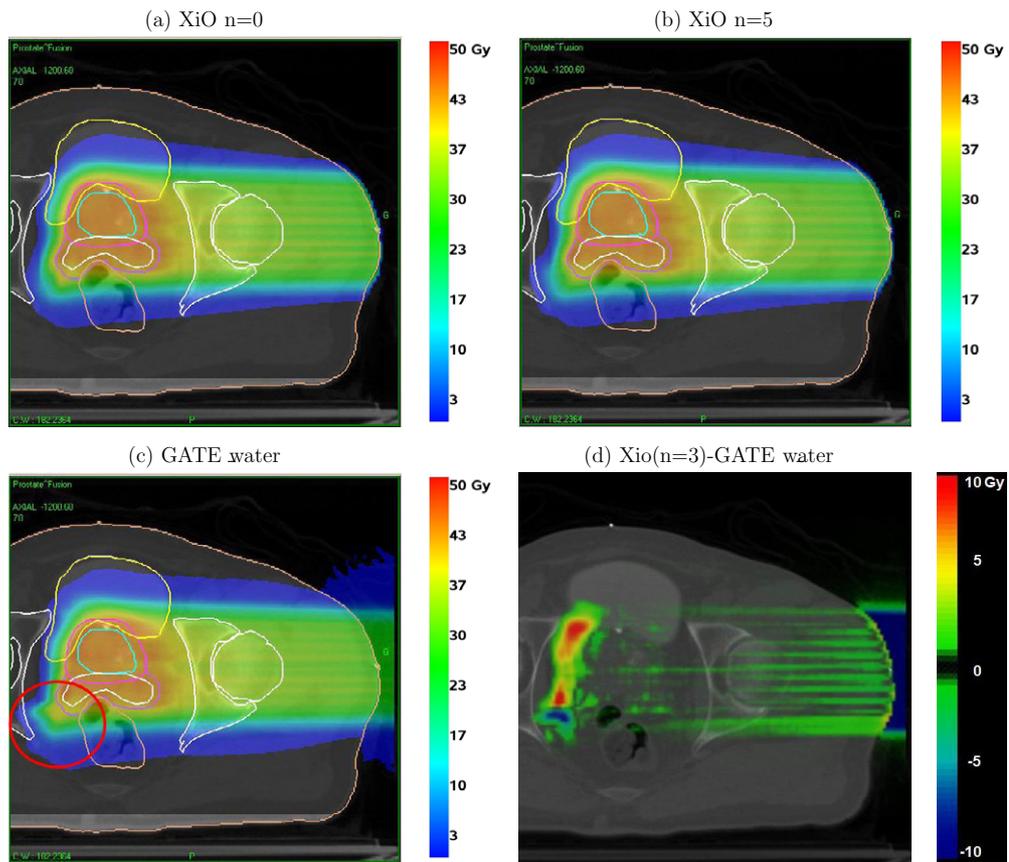


Figure 12. 2D dose distributions from GATE and XiO using $n = 0$ and $n = 5$. The dose discrepancy area due to gas in the rectum is circled in red in the GATE calculation in (c). A dose difference plot between XiO using $n = 3$ and GATE_water is presented in (d).

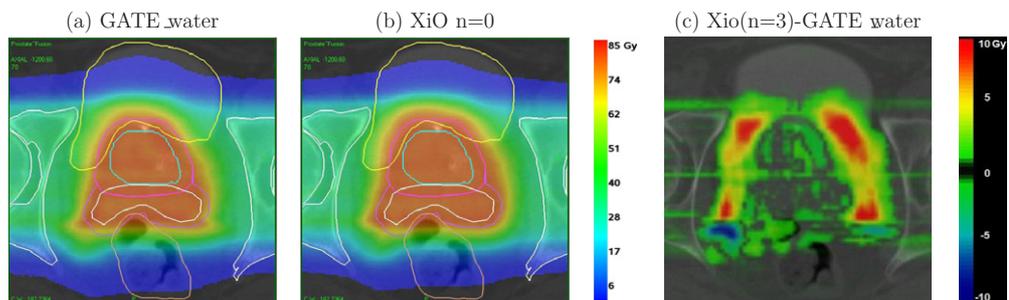


Figure 13. Close-up of the dose distributions delivered in the target with XiO and GATE, including organ contours. In (a) GATE, the distance between the prostate PTV contour (pink solid line) and the dose distal fall-off along the beam directions is shorter than in (b) XiO. (c) A dose difference plot between XiO using $n = 3$ and GATE_water.

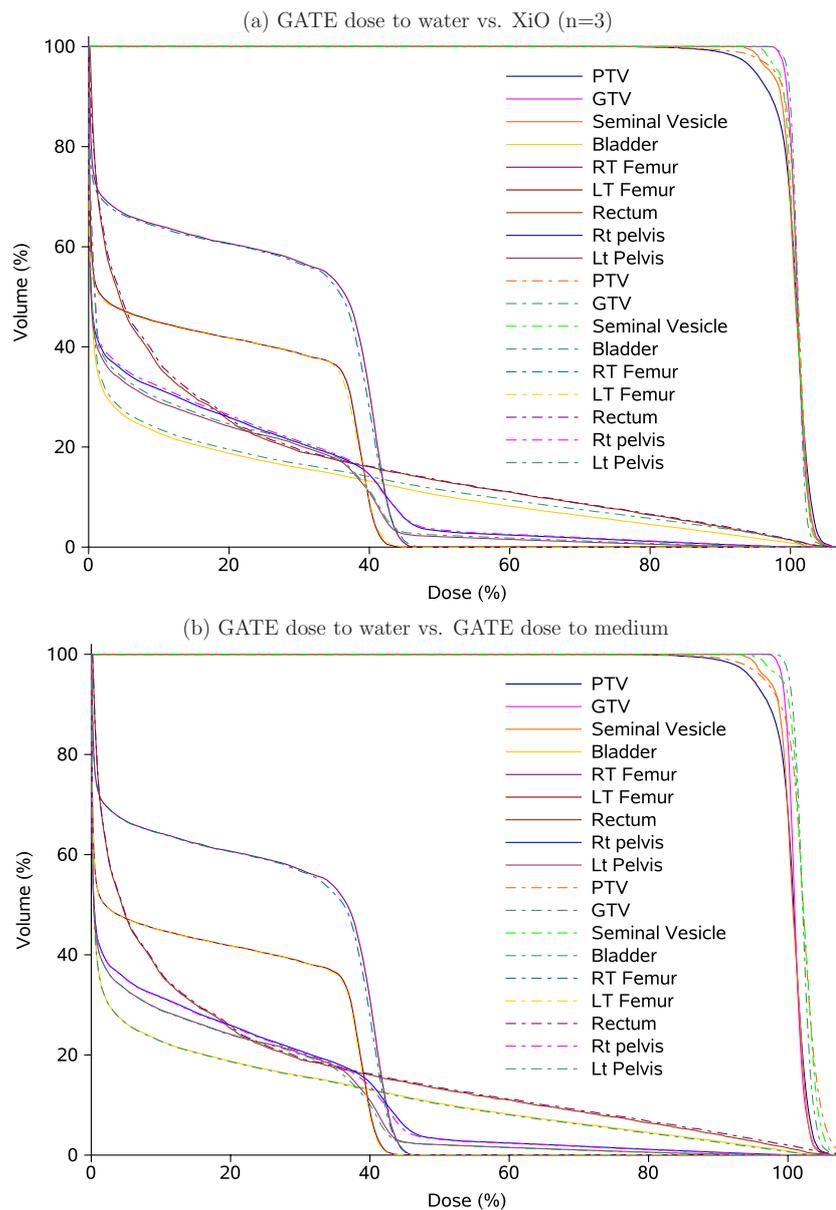


Figure 14. (a) DVHs computed with XiO using a precision $n = 3$ (dotted lines) and GATE dose to water (solid line). (b) DVHs computed with GATE using dose to water (solid line) and dose to medium (dotted line). Doses and volumes are expressed relatively to the prescribed dose and organ volumes, respectively.

4.1. Differences between dose to water and dose to medium

The main differences between dose to medium and dose to water is illustrated by a shift of the DVHs (figure 14(b)) for the target volumes, while only small differences are visible for the OARs. Similar effects have been presented by Paganetti (2009). For the organs presented,

Table 7. XiO calculation times for the left lateral field. Calculations were run on a 3.6 GHz Intel Xeon bi-processor.

Resolution	$n = 0$	$n = 3$	$n = 5$
4 mm	20 s	3 min 10 s	6 min 10 s
2 mm	2 min 50 s	14 min 30 s	25 min 30 s

iso-dose volume indices are always higher by a few percents using the GATE dose to medium instead of the GATE dose to water. The largest differences occur for V_{100} , with up to 24% difference in the seminal vesicle. The average, near-minimum and near-maximum doses are also higher by about 1–2% using the dose to medium, instead of the dose to water.

4.2. Differences between GATE dose to water and XiO

There were only minor differences when using different precision values with XiO, suggesting that for this patient the precision parameter has almost no influence on the clinical validation of the treatment plan. From the different results obtained, we considered the precision parameter $n = 3$ as a good compromise between calculation time and accuracy. The main differences between XiO and GATE dose to water occur for the PTV, for which GATE predicts lower contributions of high doses (figure 14(a)). Large differences also occur for V_{100} in the PTV, GTV and seminal vesicle. These differences can be partly explained by the shorter ranges predicted in GATE. As the PTV is the most external volume, discrepancies also occur for its V_{95} , which is about 3% lower in GATE than in XiO. It reached 94.9% in GATE and 97.8–98% in XiO, while a clinical treatment plan is generally validated for $V_{95} > 95\%$. Some slight differences can be noted in the OARs, as for instance in the bladder, for which GATE predicts lower doses.

4.3. Calculation time in XiO and GATE

The TPS calculation time is strongly related to the calculation grid resolution and to the precision parameter, as summarized in table 7 for the left field. It took 15 h with GATE/GEANT4 to recalculate the complete treatment plan on a 3.0 GHz Intel Xeon CPU, using 2×10^6 primary protons. The statistical uncertainty in voxels of $4 \times 4 \times 4 \text{ mm}^3$ was of about 1.8% in the PTV. The calculation time could be significantly reduced by using clusters of computers: for instance, using a 30 CPU cluster could reduce the calculation of the prostate plan down to about 30 min.

5. Conclusion

Active scanning is the most advanced delivery technique for proton and carbon ion therapy, but also the most challenging technique for quality assurance. The Monte Carlo simulation is currently the most accurate possibility of reviewing clinical treatment plans based on patient CT data, with an expected higher accuracy than TPSs.

In this work, we have used the open-source GATE/GEANT4 Monte Carlo simulation platform to evaluate dose distributions computed with the XiO TPS for an IBA proton PBS system. In a first step, the beam models implemented in the two dose engines were compared and found to be in satisfactory agreement. However, relative stopping powers calculated with GATE were lower than with XiO by about 1–2%, except for low density media within the

density interval [0.5, 0.9], for which XiO relative stopping powers were higher by up to 3.0%. A complex treatment plan based on the CT images of a homogeneous phantom was used to validate both dose engines: more than 96% of the points passed a 3%/3 mm gamma evaluation, when comparing GATE and XiO results with measurements. More than 95% of the points passed a 2%/2 mm gamma evaluation, when comparing GATE with XiO for the same treatment plan. In a second step, pristine Bragg peak dose distributions calculated in heterogeneous configurations were compared and showed the limits of the TPS dose calculation accuracy. Differences were attributed to the pencil beam algorithm and to the CT re-sampling artifacts. The necessity of calculating dose to water with Monte Carlo simulations for TPS evaluation has also been presented. A prostate cancer treatment plan was evaluated and dose differences between TPS and Monte Carlo calculations were analyzed in detail. Dose scoring discrepancies due to heterogeneities (gas in the rectum of the patient) were evidenced. Range differences were of about 0.5–1% (1–2 mm). DVHs were evaluated for the various organs and discrepancies were partly attributed to the stopping power differences.

At this stage, a detailed validation phase of the platform is necessary. It should include simple and complex treatment plans, heterogeneous phantoms and absolute dosimetry. This work is the first study demonstrating the capabilities of GATE to evaluate clinical treatment plans for proton active scanning delivery. The GATE platform can therefore contribute to the evaluation, benchmarking and improvement of TPS dose calculation algorithms in hadron therapy. It is also a convenient tool for imaging studies, such as PET, or for the investigation of new imaging modalities, such as the detection of prompt radiations (prompt gamma-rays or charged nuclear fragments) toward online treatment monitoring. The GATE possibilities foretell exciting interdisciplinary research studies in the field of particle therapy.

Acknowledgments

The authors of this paper would like to thank Dr Martin Soukup and Dr Gustav Meedt from the Elekta Company for fruitful discussions and support in regard to the XiO TPS. They also thank Dr Rômulo Pinho from the Creatis laboratory for his help in the processing of the patient dose difference images. This work was conducted via collaboration between the IBA Company and the Creatis laboratory. The research leading to these results has received funding from the European Community's 7th Framework Program [FP7/2007-2013] under the grant agreement number 215840-2.

References

- Agostinelli S *et al* 2003 Geant4—a simulation toolkit *Nucl. Instrum. Methods Phys. Res. A* **506** 250–303
- Allison J *et al* 2006 Geant4 developments and applications *IEEE Trans. Nucl. Sci.* **53** 270–8
- Apostolakis J *et al* 2009 Geometry and physics of the Geant4 toolkit for high and medium energy applications *Radiat. Phys. Chem.* **78** 859–73
- Bednarz B, Daartz J and Paganetti H 2010 Dosimetric accuracy of planning and delivering small proton therapy fields *Phys. Med. Biol.* **55** 7425–38
- Camarasu-Pop S, Glatard T, Mościcki J, Benoit-Cattin H and Sarrut D 2010 Dynamic partitioning of GATE Monte Carlo simulations on EGEE *J. Grid Comput.* **8** 241–59
- Dosanjh M, Jones B and Meyer R 2010 ENLIGHT and other EU-funded projects in hadron therapy *Br. J. Radiol.* **83** 811–3
- Fippel M and Soukup M 2004 A Monte Carlo dose calculation algorithm for proton therapy *Med. Phys.* **31** 2263–73
- Geant4 Collaboration 2009 *Physics Reference Manual for Geant4* CERN
- Grevillot L, Bertrand D, Dessy F, Freud N and Sarrut D 2011b A Monte Carlo pencil beam scanning model for proton treatment plan simulation using GATE/GEANT4 *Phys. Med. Biol.* **56** 5203–19

- Grevillot L, Frisson T, Maneval D, Zahra N, Badel J-N and Sarrut D 2011a Simulation of a 6 MV Elekta Precise Linac photon beam using GATE/GEANT4 *Phys. Med. Biol.* **56** 903
- Grevillot L, Frisson T, Zahra N, Bertrand D, Stichelbaut F, Freud N and Sarrut D 2010 Optimization of GEANT4 settings for proton pencil beam scanning simulations using GATE *Nucl. Instrum. Methods Phys. Res. B* **268** 3295–305
- ICRU 1993 Stopping powers and ranges for protons and alpha particles *International Commission on Radiation Units and Measurements Report 49*
- Jäkel O, Jacob C, Schardt D, Karger C P and Hartmann G H 2001 Relation between carbon ion ranges and x-ray CT numbers *Med. Phys.* **28** 701–3
- Jan S *et al* 2004 GATE: a simulation toolkit for PET and SPECT *Phys. Med. Biol.* **49** 4543
- Jan S *et al* 2011 GATE V6: a major enhancement of the GATE simulation platform enabling modelling of CT and radiotherapy *Phys. Med. Biol.* **56** 881
- Liu H H 2002 D_m rather than D_w should be used in Monte Carlo treatment planning *Med. Phys.* **29** 922–3
- Lomax A 1999 Intensity modulation methods for proton radiotherapy *Phys. Med. Biol.* **44** 185–205
- Lomax A J, Bortfeld T, Goitein G, Debus J, Dykstra C, Tercier P A, Coucke P A and Mirimanoff R O 1999 A treatment planning inter-comparison of proton and intensity modulated photon radiotherapy *Radiother. Oncol.* **51** 257–71
- Lomax A J *et al* 2004 Treatment planning and verification of proton therapy using spot scanning: initial experiences *Med. Phys.* **31** 3150–7
- Paganetti H 2009 Dose to water versus dose to medium in proton beam therapy *Phys. Med. Biol.* **54** 4399–421
- Paganetti H, Jiang H, Parodi K, Slopesma R and Engelsman M 2008 Clinical implementation of full Monte Carlo dose calculation in proton beam therapy *Phys. Med. Biol.* **53** 4825–53
- Palmans H and Verhaegen F 2005 Assigning nonelastic nuclear interaction cross sections to Hounsfield units for Monte Carlo treatment planning of proton beams *Phys. Med. Biol.* **50** 991–1000
- Schneider W, Bortfeld T and Schlegel W 2000 Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions *Phys. Med. Biol.* **45** 459–78
- Seroul P and Sarrut D 2008 VV: Viewer for the evaluation of 4D image registration *Int. Conf. Medical Image Computing and Computer-Assisted Intervention (New York)*
- Siebers J V, Keall P J, Nahum A E and Mohan R 2000 Converting absorbed dose to medium to absorbed dose to water for Monte Carlo based photon beam dose calculations *Phys. Med. Biol.* **45** 983–95
- Soukup M, Fippel M and Alber M 2005 A pencil beam algorithm for intensity modulated proton therapy derived from Monte Carlo simulations *Phys. Med. Biol.* **50** 5089–104
- Stankovskiy A, Kerhoas-Cavata S, Ferrand R, Nauraye C and Demarzi L 2009 Monte Carlo modelling of the treatment line of the Proton Therapy Center in Orsay *Phys. Med. Biol.* **54** 2377–94