


# Clinical Validation of a Ray-Casting Analytical Dose Engine for Spot Scanning Proton Delivery Systems

Technology in Cancer Research & Treatment  
Volume 18: 1-15  
© The Author(s) 2019  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/1533033819887182  
journals.sagepub.com/home/tct  


James E. Younkin, PhD<sup>1</sup> , Danairis Hernandez Morales, MS<sup>1</sup>,  
Jiajian Shen, PhD<sup>1</sup>, Jie Shan, MS<sup>1</sup>, Martin Bues, PhD<sup>1</sup>,  
Jarrod M. Lentz, MS<sup>1</sup>, Steven E. Schild, MD<sup>1</sup>, Joshua B. Stoker, PhD<sup>1</sup>,  
Xiaoning Ding, PhD<sup>1</sup>, and Wei Liu, PhD<sup>1</sup>

## Abstract

**Purpose:** To describe and validate the dose calculation algorithm of an independent second-dose check software for spot scanning proton delivery systems with full width at half maximum between 5 and 14 mm and with a negligible spray component. **Methods:** The analytical dose engine of our independent second-dose check software employs an altered pencil beam algorithm with 3 lateral Gaussian components. It was commissioned using Geant4 and validated by comparison to point dose measurements at several depths within spread-out Bragg peaks of varying ranges, modulations, and field sizes. Water equivalent distance was used to compensate for inhomogeneous geometry. Twelve patients representing different disease sites were selected for validation. Dose calculation results in water were compared to a fast Monte Carlo code and ionization chamber array measurements using dose planes and dose profiles as well as 2-dimensional–3-dimensional and 3-dimensional–3-dimensional  $\gamma$ -index analysis. Results in patient geometry were compared to Monte Carlo simulation using dose–volume histogram indices, 3-dimensional–3-dimensional  $\gamma$ -index analysis, and inpatient dose profiles. **Results:** Dose engine model parameters were tuned to achieve 1.5% agreement with measured point doses. The in-water  $\gamma$ -index passing rates for the 12 patients using 3%/2 mm criteria were 99.5%  $\pm$  0.5% compared to Monte Carlo. The average inpatient  $\gamma$ -index analysis passing rate compared to Monte Carlo was 95.8%  $\pm$  2.9%. The average difference in mean dose to the clinical target volume between the dose engine and Monte Carlo was  $-0.4\% \pm 1.0\%$ . For a typical plan, dose calculation time was 2 minutes on an inexpensive workstation. **Conclusions:** Following our commissioning process, the analytical dose engine was validated for all treatment sites except for the lung or for calculating dose–volume histogram indices involving point doses or critical structures immediately distal to target volumes. Monte Carlo simulations are recommended for these scenarios.

## Keywords

proton therapy, pencil beam, analytic, spot scanning, dose calculation

## Abbreviations

AP, anterior to posterior; CTV, clinical target volume; DE1, analytical dose engine; DVH, dose–volume histogram; ERS, extended range shifter; FWHM, full width at half maximum; IDD, integral depth dose; LPF, lateral profile function; LR, left to right; MC, Monte Carlo; MC2, fast Monte Carlo code; MX, MatriXX PT measurement; OAR, organ at risk; PBS, pencil beam scanning; RBE, relative biological effectiveness; RS, range shifter; SI, superior to inferior; SOBP, spread-out Bragg peak; TPS, treatment planning system; TPS3, Eclipse treatment planning system; VAC, vacuum; WED, water equivalent depth; WET, water equivalent thickness; 2D, 2-dimensional; 3D, 3-dimensional.

Received: December 5, 2018; Revised: August 2, 2019; Accepted: October 11, 2019.

<sup>1</sup> Department of Radiation Oncology, Mayo Clinic Arizona, Phoenix, AZ, USA

## Corresponding Author:

Wei Liu, Department of Radiation Oncology, 5881 E Mayo Blvd, Phoenix, AZ 85054, USA.  
Email: liu.wei@mayo.edu



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

## Introduction

Pencil beam scanning (PBS) is now the favored delivery technique for proton radiotherapy. Pencil beam scanning uses magnetic steering and energy selection to direct the Bragg peak of discrete or continuously delivered beam spots to preprogrammed positions throughout the target volume. The latest PBS proton delivery systems offered by major vendors including Hitachi (Tokyo, Japan) and IBA (Louvain-La-Neuve, Belgium) can deliver spots that are less than a centimeter in width, allowing highly conformal dose delivery without patient-specific hardware. Pencil beam scanning treatment fields contain a large number of spots with independent intensities, and the dose at 1 position can be affected by spatially distant spots. Among the obstacles that must be overcome to calculate PBS dose, 2 important problems are how to accurately model dose distribution of each spot and how to calculate dose in inhomogeneous patient geometry.

Pencil beam scanning spot dose distributions can be divided into a primary Gaussian and a low-dose tail. These tails have 2 sources: halo from inpatient scattering and spray from beam-line scattering.<sup>1</sup> Monte Carlo (MC) simulations<sup>2</sup> as well as film and scintillator charge-coupled device measurements<sup>3,4</sup> have shown that low-dose tails may extend several centimeters from a spot's centroid position. In a PBS field with many spots, these tails produce an increase in delivered dose with larger lateral field size known as the field size effect.<sup>5</sup> If the spot dose model does not incorporate these corrections to the primary Gaussian, the predicted dose may have up to 10% error.<sup>5-7</sup>

Patient anatomies near target volumes often include low-density materials such as air in lung tissue or high-density materials such as bones. Proton particle transport through these inhomogeneous structures is governed by energy loss, multiple Coulomb scattering, and nuclear interactions.<sup>1,8</sup> These processes are challenging to accurately represent in an analytic dose calculation engine,<sup>9</sup> which becomes especially problematic for some disease sites (eg, lung or head and neck) allocated in areas with high levels of tissue inhomogeneity.<sup>10-12</sup>

Monte Carlo simulation calculates the field size effect and inhomogeneity problems directly from the underlying physics, but even fast MC methods<sup>13,14</sup> are currently not widely available or fast enough for iterative optimization. Although MC dose calculations are suggested when treating highly heterogeneous sites,<sup>10-12</sup> comprehensively commissioned and validated analytic dose calculation models<sup>6,7</sup> provide higher computational speed required for many routine clinical tasks. For this reason, commercial treatment planning systems (TPSs) rely on analytic dose calculations. The most common analytical dose calculation models used in proton therapy are based on either fluence-dose or ray-casting pencil beam algorithms.<sup>9</sup> Both methods divide the dose model into an integral depth dose (IDD) component and a lateral profile function (LPF) component that describes the distribution of dose at each depth.<sup>15</sup>

Gaussian mixtures<sup>16</sup> are often used in the LPF to model the non-Gaussian low-dose tails, and a double Gaussian LPF has been used to correct gross dose errors due to field size effects.<sup>5</sup> A triple Gaussian LPF has been used to calculate the dose of

carbon-12 beams, for which nuclear fragmentation generates secondaries of varying mass.<sup>17-19</sup> With an accurate in-air fluence model, the use of a triple Gaussian LPF in the dose kernel of a fluence-dose algorithm was shown to correct for mid-range effects of nuclear interactions at high (>150 MeV) proton energies.<sup>20</sup> The triple Gaussian LPF has been shown to provide a superior fit to measured lateral dose profile of proton beams compared to the double Gaussian LPF.<sup>21</sup>

The fluence-dose and ray-casting algorithms use different models of the physical beam and different methods to compensate for inhomogeneous geometries. The fluence-dose algorithm uses a beam fluence model and a model of the dose deposited in water by an "elemental" monoenergetic pencil beam of infinitesimal size. Dose deposited by the physical beam at a point is equal to the convolution over the transverse plane of the beam fluence in air with the elemental pencil beam dose kernel evaluated at the appropriate lateral distance and water equivalent depth (WED) in the beam direction.<sup>9</sup> On the other hand, the ray-casting algorithm starts with a pencil beam model of the finite-size spot with realistic proton beam energy spectrum and therefore does not require a beam fluence model. In an inhomogeneous geometry, the dose delivered by a beam to a point is the spot dose evaluated at the lateral distance and at the WED of the point, which is computed by casting a ray in the beam direction from the calculation point to the body surface. By modifying the IDD to account for the range dilution caused by multiple Coulomb scattering at the interface of lateral inhomogeneities, Schaffner *et al* developed a modified ray-casting model, which more accurately predicted the MC generated dose distribution in half-slab geometries.<sup>9</sup>

The complimentary nature of fluence-dose and ray-casting algorithms in heterogeneous geometries was described by Schaffner *et al* in a study recommending the use of both algorithms in clinical workflows and MC simulation as a final dose check in highly complex geometries.<sup>9</sup> Comparing the fluence-dose and the modified ray-casting algorithms to MC simulation, the fluence-dose algorithm more accurately reproduced the MC simulated dose distribution when the half-slab was at a shallow depth and the modified ray-casting algorithm had better performance when the half-slab was closer to the Bragg peak. Actual patient geometries contain inhomogeneities at a variety of depths relative to delivered Bragg peaks. If both analytical algorithms are available, then they can provide an independent check of one another.

In this article, we present an analytical dose engine (DE1) based on our own modification of a ray-casting pencil beam algorithm. This dose engine builds upon previous ray-casting algorithms<sup>9,22,23</sup> but offers several features that make dose calculation fast and accurate for narrow-width beamlines. Just as Hitachi's PROBEAT spot scanning proton delivery system was later redesigned with a retractable beam profile monitor (which had been the major source of spray) and spot widths reduced from 5 to 15 mm to 2 to 6 mm,<sup>1,6</sup> narrow-width proton delivery systems designed to minimize spray are likely to become standard in the future.<sup>24</sup> The features of our dose engine include (1) a triple Gaussian LPF that is used to accurately model the low-dose tails of spots and (2) water equivalent distance is used in the beam and lateral direction to improve accuracy in

inhomogeneous geometries. This analytic dose engine was comprehensively validated by comparing the results of the DE1 calculations to 2-dimensional (2D) ionization chamber array measurements in water and to MC simulations both in water and in patient geometry. Although other studies have compared DE1s to either in-water measurements<sup>23</sup> or inpatient MC simulations,<sup>10,12</sup> to our knowledge this is the first study to combine these analyses to validate the dose engine.

## Materials and Methods

### Proton Delivery System Characteristics

The proton delivery system at Mayo Clinic Arizona (Hitachi ProbeatV5, Tokyo, Japan) is a contemporary update of the earlier version used at the University of Texas MD Anderson Cancer Center.<sup>25,26</sup> The new version of this proton system used at Mayo Clinic Arizona utilizes PBS exclusively and provides 97 discrete energies varying from 71.3 to 228.8 MeV. The in-air width of spots at isocenter varies from 2 to 6 mm (full width at half maximum [FWHM] between 5 and 14 mm) for the highest and lowest energies. The system design minimizes beamline scattering and features a retractable beam profile monitor.

A range shifter (RS) with a 45 mm water equivalent thickness (WET) is used to treat shallow tumors. The RS can be inserted at 2 fixed positions in the nozzle: either at 42.5 or 30 cm from isocenter. The RS at 30 cm has the advantage of smaller spot size when compared with the RS at 42.5 cm.<sup>27</sup> To distinguish the 2 positions, the RS at 30 cm is called the extended RS (ERS) throughout the article. Since scattering can cause the spot profile

to change significantly after passing the RS,<sup>27</sup> the beamlines with the RS or the ERS were modeled separately, as different machines. Hence, 3 beam models were commissioned for the DE1: (1) the vacuum (VAC) machine, which does not use a RS; (2) the RS45 machine, which places the RS of 45 mm WET at 42.5 cm from isocenter; and (3) the ERS45 machine, which places the RS of 45 mm WET at 30 cm from isocenter.

### Analytical Dose Calculation Engine

The DE1's pencil beam algorithm is based on a model of the total beamlet (spot) dose in water. The 2 components of this pencil beam model are the IDD, which includes all contributions from the core, spray, and halo,<sup>1</sup> and the LPF, which specifies the distribution of dose at each depth in the lateral direction. A model of the spot dose in water is transformed to dose in the patient geometry according to the WED of each voxel calculated in the beam direction. This algorithm is sometimes called "ray-casting" since the dose in water is scaled by the WED along rays through the patient geometry.<sup>9</sup> Our modification to this algorithm (calculating WED in the lateral direction) is an attempt to incorporate the influence of inhomogeneities without sacrificing computational efficiency appreciably (eg, by subdividing the spot<sup>28</sup>). The algorithm is computationally efficient, but additional speed was obtained by parallelizing the dose calculation using OpenMP ver. 5.0.<sup>29</sup>

The dose  $D$  to each voxel  $(x, y, z)$  is the sum over contributions from each spot  $j$ :

$$D_i(x, y, z) = \sum_{j=1}^{n_s} \frac{\Psi_j \text{IDD}_i(z_{\text{WED}}, E_j)}{\Delta x \Delta y} \int_{y-1/2\Delta y}^{y+1/2\Delta y} \int_{x-1/2\Delta x}^{x+1/2\Delta x} \text{LPF}_i(z_{\text{WED}}, r_{\text{LAT}}, E_j) dx' dy' \quad (1)$$

In this formula,  $n_s$  is the total number of spots while  $\Psi_j$  is the fluence of a spot and  $E_j$  is the nominal beam energy of spot  $j$ . The parameter  $z_{\text{WED}}$  is the WED calculated by casting a ray in the beam direction from the calculation position  $(x, y, z)$  to the body surface, and  $r_{\text{LAT}}$  is the lateral water equivalent distance from the calculation point  $(x, y, z)$  to the projected spot position at the same depth  $(x', y', z)$ . The lengths  $\Delta x$  and  $\Delta y$  are the dose calculation resolution in the  $x$  and  $y$  directions, which were both equal to 2.5 mm in this study. The subscript  $i$  indicates the 3 machines (ie, VAC, RS45, and ERS45) corresponding to the beamline configurations used at our clinic. Model parameters for the in-air spot width  $\sigma_{i,\text{air}}$  as well as  $\text{IDD}_i(z_{\text{WED}}, E_j)$  and  $\text{LPF}_i(z_{\text{WED}}, r_{\text{LAT}}, E_j)$  were determined during commissioning for each nominal beam energy  $E_j$  and for each beam model  $i$ . In the remainder of the article, the machine subscript  $i$  is suppressed for simplification.

The IDD was parametrized using Bortfeld's analytical formula<sup>30</sup> with corrections introduced by Zhang *et al.*<sup>31</sup> The LPF incorporates both the in-air spread and the in-water scattering of the beam. The initial in-air profile was modeled downstream of the beamline using a single Gaussian of width  $\sigma_{\text{air}}$  for each energy  $E_j$ . The dependence of the in-air spot width  $\sigma_{\text{air}}(d, E_j)$  on the position  $d$  along the beam axis was determined by fitting to the well-known formula<sup>5,6,32</sup>:

$$\sigma_{\text{air}}(d, E_j) = \sqrt{\frac{A}{2} + Bd + \frac{C}{2}d^2} \quad (2)$$

A triple Gaussian LPF was used to model the in-water lateral dose profile of the spot:

$$\text{LPF}(z_{\text{WED}}, r_{\text{LAT}}, E_j) = \frac{1}{\pi} \left[ \left( \frac{1 - W_2 - W_3}{\sigma_1^2} \right) \exp\left(-\frac{r_{\text{LAT}}^2}{\sigma_1^2}\right) + \frac{W_2}{\sigma_2^2} \exp\left(-\frac{r_{\text{LAT}}^2}{\sigma_2^2}\right) + \frac{W_3}{\sigma_3^2} \exp\left(-\frac{r_{\text{LAT}}^2}{\sigma_3^2}\right) \right] \quad (3)$$

where the Gaussian widths were determined by the following formulae:

$$\begin{aligned}\sigma_1^2(d, z_{\text{WED}}, E_j) &= \sigma_{\text{air}}^2(d, E_j) + \sigma_{\text{core}}^2(z_{\text{WED}}, E_j) \\ \sigma_2^2(d, z_{\text{WED}}, E_j) &= \sigma_1^2(d, z_{\text{WED}}, E_j) + \sigma_{\text{LAS}}^2(z_{\text{WED}}, E_j) \\ \sigma_3^2(d, z_{\text{WED}}, E_j) &= \sigma_2^2(d, z_{\text{WED}}, E_j) + \sigma_{\text{NI}}^2(z_{\text{WED}}, E_j)\end{aligned}\quad (4)$$

and the LPF was normalized to make the integral over the transverse plane equal to one. Here,  $\sigma_{\text{core}}(z_{\text{WED}}; E_j)$  is the increase in core width of the beam with nominal energy  $E_j$  at a depth  $z_{\text{WED}}$  in water caused by multiple Coulomb scattering, while  $\sigma_{\text{LAS}}(z_{\text{WED}}; E_j)$  and  $\sigma_{\text{NI}}(z_{\text{WED}}; E_j)$  are the additional widths ascribed to large-angle Coulomb scattering and nuclear interactions, respectively. The Highland approximation,<sup>33</sup> which has been shown to be a good approximation of the root mean square scattering angle obtained from the Molière angular distribution,<sup>8</sup> was used to determine the core Gaussian width  $\sigma_{\text{core}}(z_{\text{WED}}; E_j)$ . The Gaussian weights  $W_2$  and  $W_3$  and the excess widths associated with large-angle scattering and nuclear interactions  $\sigma_{\text{LAS}}$  and  $\sigma_{\text{NI}}$  were obtained in units of the weight  $W_{\text{nuc}}$  and width  $\sigma_{\text{nuc}}$  of the Soukup nuclear pencil beam<sup>28</sup>:

$$\begin{aligned}W_2(z_{\text{WED}}, E_j) &= W_{\text{nuc}} f_{W_2} \left( \frac{z_{\text{WED}}}{R_0}, E_j \right) \\ W_3(z_{\text{WED}}, E_j) &= W_{\text{nuc}} f_{W_3} \left( \frac{z_{\text{WED}}}{R_0}, E_j \right) \\ \sigma_{\text{LAS}}(z, E_j) &= \sigma_{\text{nuc}} g_{\sigma_{\text{LAS}}} \left( \frac{z_{\text{WED}}}{R_0}, E_j \right) \\ \sigma_{\text{NI}}(z, E_j) &= \sigma_{\text{nuc}} g_{\sigma_{\text{NI}}} \left( \frac{z}{R_0}, E_j \right)\end{aligned}\quad (5)$$

where each nominal beam energy has power law functions  $f$  and polynomial functions  $g$  of the ratio of WED to fitted beam range  $z_{\text{WED}}/R_0$ .<sup>8</sup> Coefficients of these functions were themselves quadratic functions of nominal beam energy  $E_j$  to permit simultaneous optimization.<sup>31</sup>

### Commissioning the DE1

The DE1 was commissioned using data generated by an MC code<sup>34</sup> based on the Geant4 10.0 release.<sup>35-37</sup> The geometry of the PBS nozzle in the MC simulation was matched to the design configuration provided by the vendor and the vendor-provided data. In the course of commissioning our proton delivery system, the source model in the MC code was adjusted to match IDD measurements to within 0.2 mm in range and to match field size factor measurements to within 2%. Additional details concerning this MC code can be found in the literature.<sup>34</sup> A physics list for Geant4 medical applications, ‘‘QGSP\_BIC\_EMY,’’ provided the model for interactions between particles and materials. This physics list has been

validated by the authors of the Hadrontherapy module, from which our MC code was developed.<sup>36</sup> The water target consisted of voxels 1 mm<sup>3</sup> in size; dose and fluence were scored on this calculation grid. The MC simulation parameters were chosen to balance calculation accuracy against simulation time. In all simulations, the cut was 0.01 mm and the maximum allowed step was 0.01 mm.

This MC code was used to calculate both IDD curves and in-air lateral profiles at isocenter,  $\pm 10$  cm, and  $\pm 20$  cm to obtain the in-air width  $\sigma_{\text{air}}$  for all nominal beam energies. The target volume had 40 cm lateral extent to ensure that the entire IDD was measured. For a selected subset of 14 nominal beam energies, the MC code was used to calculate lateral profiles in water using the radial projection method.<sup>34</sup> Lateral profile function weights and widths in Equation 3 were fit to simulated lateral profiles by simultaneous optimization.<sup>31</sup> These parameters were interpolated to obtain LPFs for all 97 beam energies.

The DE1 parameters were finely adjusted to match field size factor measurements. A PTW 34045 advanced Markus electron chamber (PTW, Freiburg, Germany) was used to measure central axis dose at various depths on the flattop of the depth-dose profile of spread-out Bragg peaks (SOBPs) with various source-to-surface distances, ranges, modulations, and field sizes (from 4  $\times$  4 cm to 20  $\times$  20 cm). The selected set of SOBPs ranges and modulations included all nominal beam energies used by each machine in clinical practice. Modulations were representative of the typical extent of treatment fields in depth. The isocenter was positioned at the center of the flattop of the SObP ( $R - M/2$ ) unless this would place the water tank too close to the nozzle. The in-air width, IDD, and LPF parameters were required to achieve an agreement between the measured and calculated point doses within  $\pm 2\%$ . The difference between measured and calculated point dose after fine tuning was used as a quantitative measure of the beam model quality.

### Clinic Patient Cohort Statistics

A cohort of 12 patients treated at Mayo Clinic Arizona was used to validate the DE1. Patients were selected to represent the range of disease sites and machines at our institution. This study was approved by the institutional review board of Mayo Clinic Arizona. For each site, a set of dose–volume histogram (DVH) indices for organs at risk (OARs) used in treatment planning was selected. Table 1 contains information about each patient plan including the number of fields, the treatment machine, and the selected OARs for each disease site.

### Performance of the DE1 in Water

A fast Graphic-Processing Unit (GPU)-accelerated MC dose calculation engine developed at Mayo Clinic, Rochester, Minnesota<sup>13,14</sup> and a commercial TPS, Eclipse version 13.7 (Varian Medical Systems, Palo Alto, California),<sup>38</sup> were used to assess the relative accuracy of the DE1. Henceforth, for simplicity, the abbreviation DE1 refers to the analytical dose engine, MC2

**Table 1.** Characteristics of the 12 Patient Plans Used to Validate the Dose Engine.

Patient	Disease Site	Number of Fields	Machine <sup>a</sup>	OARs <sup>b</sup>	
1	Prostate	2	VAC	Bladder, femoral head, rectum	
2		2			
3		2			
4		2			
5	Head and neck	5	ERS45	Cochleae, optic nerve, eye, spinal cord	
6		4			Cochleae, brain stem, spinal cord, optic cavity
7		4			
8	Lung	2	VAC	Lungs, spinal cord, heart	
9		2			
10	Brain	2	ERS45	Brain stem, optic chiasm, cochleae	
11	Breast	2	RS45	Spinal cord, lungs, esophagus, thyroid	
12	Craniospinal	3	RS45	Brain stem, spinal cord, optic chiasm, eyes	

Abbreviations: DVH, dose–volume histogram; ERS, extended range shifter; OARs, organs at risk; RS, range shifter; VAC, vacuum.

<sup>a</sup>VAC: no RS; ERS45: 45 mm RS placed 30 cm upstream of isocenter; and RS45: 45 mm RS placed 42.5 cm upstream of isocenter.

<sup>b</sup>Due to the variety of tumor locations, DVH objectives and OARs differed between patients for the head and neck disease site.

refers to the fast MC code, and TPS3 refers to the commercial TPS. Dose distributions for each treatment field were calculated in a water phantom using all 3 dose engines. For each treatment field, dose planes at depths proximal to, distal to, and within the target volume were measured with a MatriXX PT 2D ionization chamber array (IBA Dosimetry GmbH, Schwarzenbruck, Germany).<sup>39</sup> The measurements of MatriXX PT detectors (with 7.62 mm spacing and 4.5 mm chamber diameter) were compared to DE1- and MC2-calculated lateral and depth-dose profiles, with the measurement providing the ground truth in the comparison.

As a more comprehensive evaluation, the measured dose planes were compared to both the DE1-calculated and the TPS3-calculated doses using 2-dimensional–3-dimensional (2D-3D)  $\gamma$ -index analysis.<sup>40</sup> The DE1 and TPS3 dose planes were aligned to the measured dose plane using the image registration algorithm employed by our institution’s patient-specific quality assurance program.<sup>39</sup> The entire DE1 dose distribution in water was also compared to the MC2 dose distribution using 3-dimensional–3-dimensional (3D-3D)  $\gamma$ -index analysis to determine the overall quality of the spot dose model. In this 3D-3D comparison, the MC2 dose was accepted as the ground truth. Unless specifically stated otherwise, all  $\gamma$ -index analyses in this article used 3% of the maximum dose and 2 mm distance criteria with a 10% dose threshold relative to the maximum dose in the reference volume. These criteria were selected to match recommendations of American Association

**Table 2.** Tests Used to Validate the DE1 and Their Associated Validation Criteria.

Test	Basis for Comparison	Validation Criteria
2D-3D $\gamma$ -index analysis in-water	MatrixxPT (MX) measurement	$\geq 90\%$ passing rate <sup>a</sup>
3D-3D $\gamma$ -index analysis in-water	MC2 calculation	$\geq 90\%$ passing rate <sup>a</sup>
3D-3D $\gamma$ -index analysis in patient geometry	MC2 calculation	$\geq 90\%$ passing rate <sup>a</sup>
CTV mean dose/ $D_{95\%}$	MC2 calculation	$< 3\%$ relative dose difference
OAR DVH indices	MC2 calculation	$< 10\%$ difference

Abbreviations: CTV, clinical target volume; DVH, dose volume histogram; MC2, fast Monte Carlo code; OAR, organ at risk; 2D, 2-dimensional; 3D, 3-dimensional.

<sup>a</sup>Universal action limit set by American Association of Physicists in Medicine Task Group (AAPM TG) report 218 for measurement-based Intensity-Modulated Radiation Therapy Quality Assurance (IMRT QA).<sup>41</sup>

of Physicists in Medicine (AAPM) Task Group 218 for patient-specific Intensity-Modulated Radiation Therapy (IMRT) quality assurance<sup>41</sup> since no comparable standards currently exist for proton therapy.

### Performance of the DE1 in Patient Geometry

For each treatment plan, the dose distribution in the patient geometry was calculated using DE1, MC2, and TPS3 to evaluate DE1 performance in inhomogeneous geometries. A 3D-3D  $\gamma$ -index analysis was performed to compare DE1 and TPS3 results to MC2 within the volume enclosed by an external structure defined as BODY, which encloses the whole region of interests. Then, DVH indices were computed from DE1 and MC2 dose distributions for the clinical target volume (CTV) and for selected DVH objectives for OARs to determine the effectiveness of DE1 as a second-dose check. The MC2 dose was treated as ground truth for the 3D-3D  $\gamma$ -index analysis in the patient geometry as well as the DVH index comparisons. In this study, a constant relative biological effectiveness (RBE) value of 1.1 was used to report absolute proton dose. Finally, color wash dose distributions and dose profiles were compared to identify discrepancies between DE1 and MC2 within the patient geometry. Table 2 contains a list of all tests used to validate DE1.

### Computational Efficiency of the DE1

Analytical dose engine calculations were executed on a modest workstation equipped with dual E5-2680 v3 processors (Intel Corporation, Santa Clara, California) and 64 GB (2133 MHz) RAM. For each patient, the amount of computation time required to complete the DE1 calculation was recorded along with the 2 major factors contributing to differences in plan dose calculation time, the number of spots in the treatment plan and the volume of the external structure BODY.

## Results

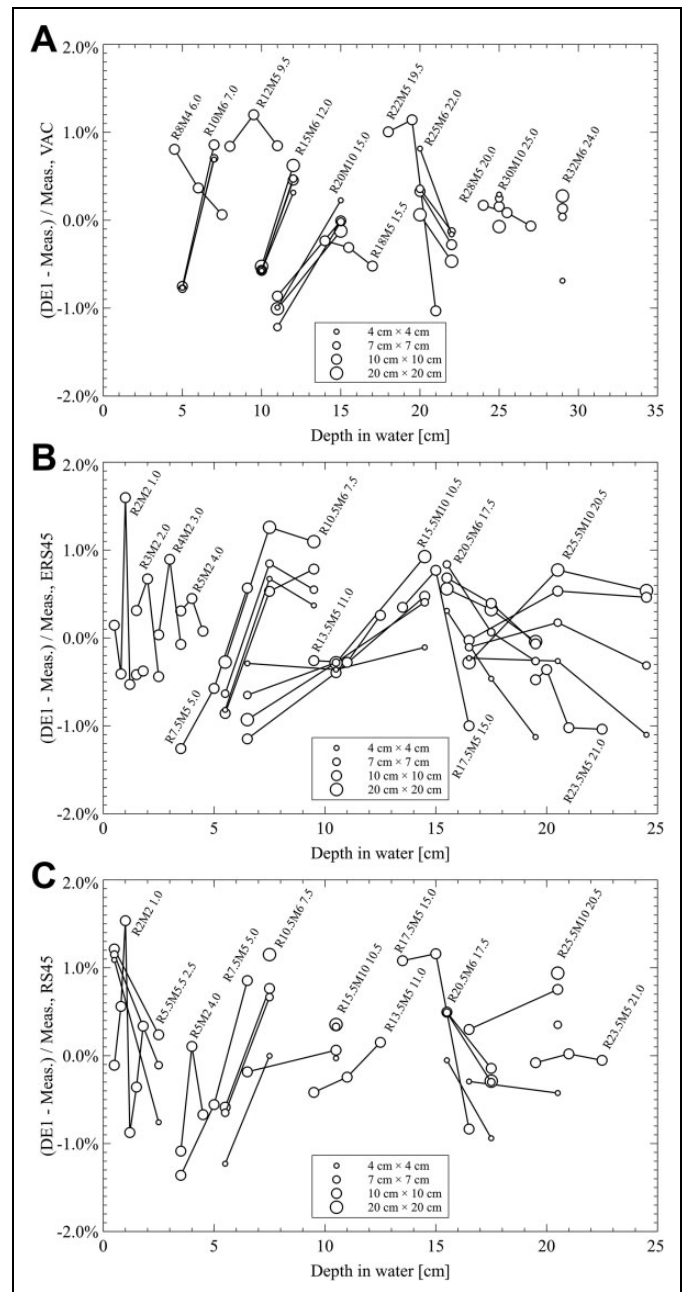
### Field Size Effects

Figure 1 shows the percent deviation of the DE1-calculated central axis point doses from measured dose. Each plot contains results for 1 machine: (1) VAC, (2) ERS45, and (3) RS45. Circle size denotes the dimensions of the measured field, and connected circles correspond to identical SOBPs and field sizes measured at different depths. The labels placed near sets of circles list the SOBP range, width, and depth of isocenter below the water surface in a water tank. For example, “R25.5M10 20.5” on the ERS45 graph means that the circles at 16.5, 20.5, and 24.5 cm depth in water were measurements of a field with 25.5 cm range (after the RS with 4.5 cm WET), 10 cm modulation (ie, SOBP width), and the isocenter positioned at 20.5 cm below the water surface. Additionally, the label “R32M6 24.0” means that the circles at 29 cm depth in water were measurements of a field with 32 cm range, 6 cm modulation, and the isocenter positioned at 24 cm below the water surface with 4 different field sizes (from 4 cm × 4 cm to 20 cm × 20 cm). All DE1 calculated central axis point doses were within 1.5% of measured values. Mean percent deviations (the average overall measurements of the ratio of the dose difference DE1 measurement to the measured dose) for each machine were 0.02% ± 0.61% for VAC, -0.03% ± 0.68% for ERS45, and 0.1% ± 0.66% for RS45.

### Accuracy of DE1 Calculation in Water

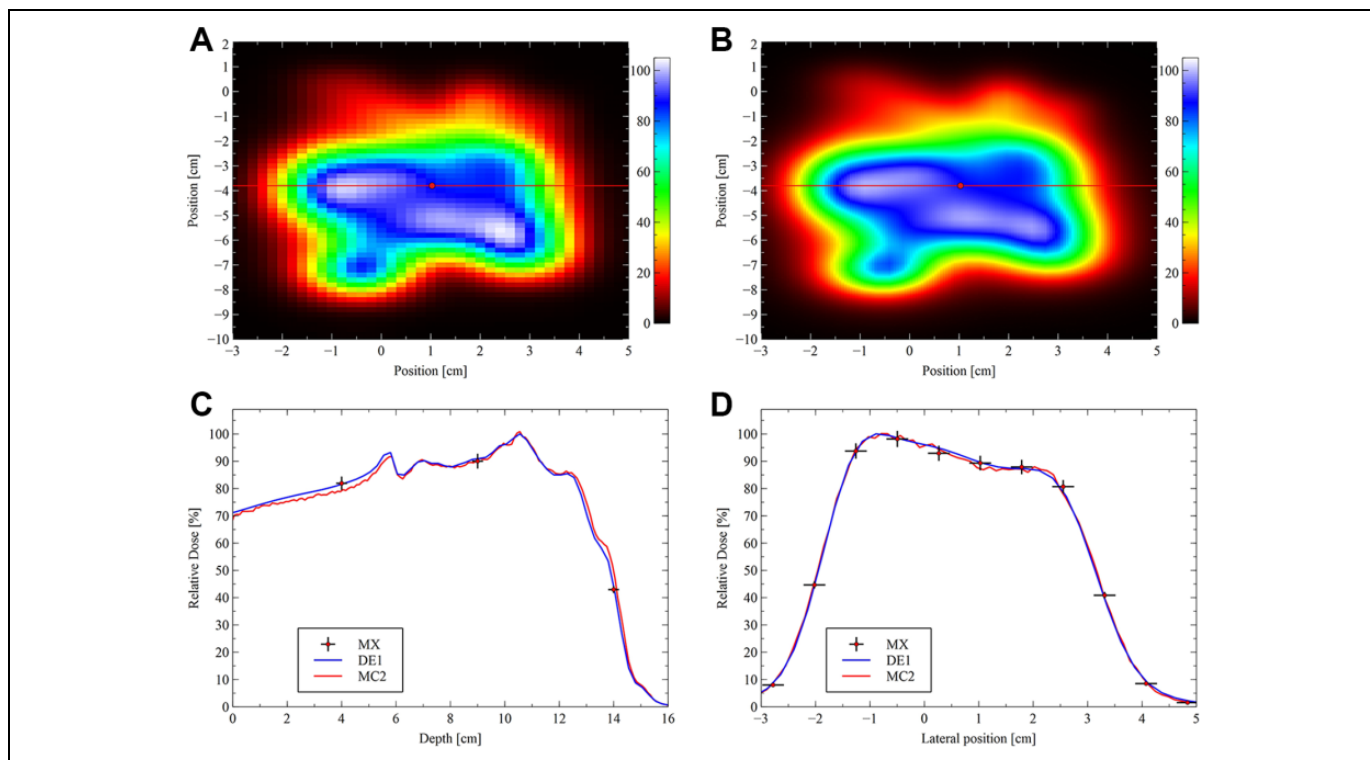
Figure 2 is an example of the in-water dose plane and dose profile comparisons between MatriXX PT measurement (MX), DE1, and MC2 dose calculation for field 4 of patient 6 at a depth of 9.0 cm within the target. The top row shows (1) color map images of the measured dose plane and (2) the corresponding plane from the DE1 calculation. Cubic spline interpolation was used to increase the image resolution in Figure 2A for ease of comparison to Figure 2B. In both images, 100% relative dose corresponds to the maximum dose in the DE1 plane. The bottom row has (3) a depth-dose profile along the beam direction and (4) a lateral profile in the same plane shown as a red point and a red line, respectively, in Figure 2A and B. In both plots, the maximum DE1 dose defined 100% relative dose for all profiles and measurements. All ionization chamber measurements from MX were within 3% of the measured dose and 2 mm in depth or position from the DE1-calculated results, as shown by the error bars that are provided as a visual aid for this comparison.

The combined box plots and strip plots in Figure 3 show the in-water  $\gamma$ -index analysis results for all treatment fields. Each box extends vertically from the first to the third quartile, with the median value indicated by the horizontal line within the box. The lengths of the whiskers (error bars) are 1.5 times the interquartile range (third quartile – first quartile). Strip plots of  $\gamma$ -index analysis passing rates for each field are overlaid on the corresponding box plot. Strip plot symbols indicate both the disease site treated and the machine used by each field (see

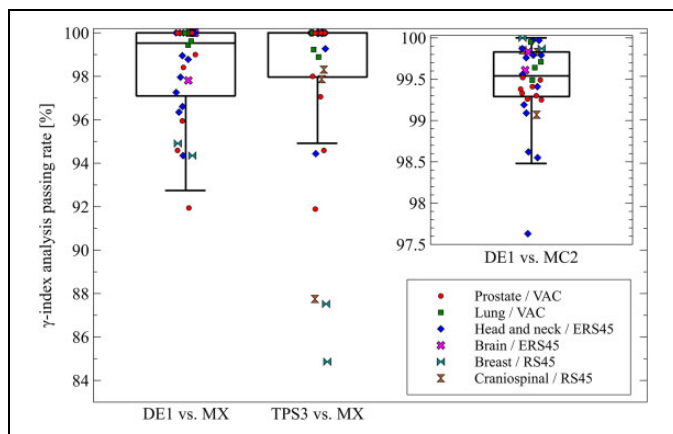


**Figure 1.** Percent deviation of the DE1-calculated central axis point dose from measurements taken on the flat tops of SOBPs with various field sizes using each of the 3 machines: (A) VAC, (B) ERS45, and (C) RS45. Circles indicate percent deviation from measurement; the size of the circle corresponds to the field size and lines connect circles that have field size and SOBP in common but different measurement depths. Labels indicate SOBP range, width, and isocenter depth below the water surface in a water tank. The horizontal axis is the detector depth in water. DE1 indicates analytical dose engine; ERS, extended range shifter; RS, range shifter; SOBP, spread-out Bragg peak; VAC, vacuum.

Table 1). Fields that had passing rates below the whiskers were particularly challenging for the respective dose engine and are considered outliers; however, all fields including these outliers are included in our analysis.



**Figure 2.** Comparison of dose planes in water between DE1 and MX for patient 6 (head and neck) field at a depth that intersects the target volume. A, MatriXX PT measurement dose plane measured at 9 cm depth in water is compared to (B) the corresponding dose plane calculated by DE1. C, Depth dose profiles through the target volume calculated by DE1 and MC2 were compared to the point-dose measurements at 3 depths. Error bars on MX correspond to 2 mm depth/position and 3% point dose uncertainty. D, Lateral dose profiles at 9 cm depth through the target volume from DE1 and MC2 were compared to MX point doses. Both profiles are normalized to the maximum DE1 dose. Error bars are included as a visual aid for the comparison between profiles and measurements; they indicate either a 3% difference in dose or 2 mm difference in position or depth. DE1 indicates analytical dose engine; MC2, fast Monte Carlo code; MX, MatriXX PT measurement.



**Figure 3.** The  $\gamma$ -index analyses for in-water cohort dose calculations using 3%/2 mm and 10% relative dose criteria. From left to right, the 3 box plots correspond to the 2D-3D comparison of DE1 and MX, the 2D-3D comparison of TPS3 and MX, and the 3D-3D comparison of DE1 and MC2. The median and third quartile passing rates for the TPS3 versus MX box plot were both equal to 100%. Strip plots of field passing rates are overlaid on the box plots, and symbols for each field indicate both treatment site and treatment machine. DE1 indicates analytical dose engine; MC2, fast Monte Carlo code; MX, MatriXX PT measurement; TPS, treatment planning system; 2D, 2-dimensional; 3D, 3-dimensional.

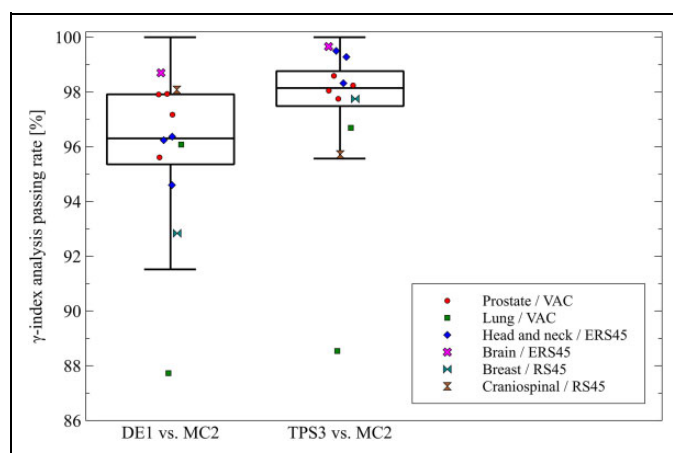
The 2 box plots on the left side of Figure 3 show the 2D-3D  $\gamma$ -index analysis passing rates between either DE1 or TPS3 and MX. Median passing rates were 99.5% for DE1 and 100% for TPS3, whereas average passing rates were  $98.3\% \pm 2.2\%$  for DE1 and  $97.8\% \pm 4.1\%$  for TPS3. All DE1 passing rates were at least 91.9% (which was greater than our validation criteria of 90%), whereas the lowest TPS3 passing rate was 84.9%. The outlier for DE1 was a prostate field (circle), whereas the outliers for TPS3 were prostate, head and neck (diamonds), breast (horizontal ties), and craniospinal fields (vertical ties). In particular, the TPS3 breast fields had the lowest passing rates.

The inset box plot on the right side of Figure 3 shows the 3D-3D  $\gamma$ -index analysis passing rates comparing the full DE1 and MC2 dose volumes in water. The average passing rate for the DE1 volumes was  $99.5\% \pm 0.5\%$ , the median passing rate was 99.5%, and the minimum passing rate was 97.6%. A head and neck treatment field using the ERS45 machine was the outlier; head and neck fields exhibited slightly larger variation in passing rates compared to other types of fields.

### Accuracy of DE1 Calculation in Patient Geometry

The 3D-3D  $\gamma$ -index analysis results comparing doses calculated in the patient geometry by DE1 and TPS3 to the dose





**Figure 4.** Three-dimensional (3D)  $\gamma$ -index analyses for inpatient dose calculations using 3%/2 mm and 10% relative dose criteria. Box plots correspond to the comparison of DE1 and MC2 and the comparison of TPS3 and MC2. Overlaid strip plots show passing rates for each treatment plan, with symbols for each plan indicating the combination of treatment site and treatment machine. DE1 indicates analytical dose engine; MC2, fast Monte Carlo code; TPS, treatment planning system; 3D, 3-dimensional.

calculated by MC2 are shown in the combined box and strip plots in Figure 4. Median  $\gamma$ -index analysis passing rates were 96.3% for DE1 and 98.1% for TPS3. The average DE1 passing rate was  $95.8\% \pm 2.9\%$  and the average TPS3 passing rate was  $97.2\% \pm 2.9\%$ . Patient 9 (lung/VAC) was the common outlier for both DE1 and TPS3. This was also the only patient whose  $\gamma$ -index analysis passing rate (87.7%) was below the 90% passing rate required for validation. However, it is worth noting that the TPS3 passing rate (88.5%) was also less than 90% for this patient. The average per-patient difference between DE1 and TPS3 passing rates (DE1 passing rate – TPS3 passing rate) was  $-1.4\% \pm 1.7\%$ .

Table 3 lists DVH indices derived from the DE1- and MC2-calculated CTV dose distributions. Pairs of columns compare DE1 results to MC2 results for the following DVH indices: mean dose,  $D_{95\%}$ ,  $D_{5\%}$ , and  $D_{95\%} - D_{5\%}$ . The average difference in CTV dose (DE1 – MC2) was  $-0.43\% \pm 1.02\%$  over the entire cohort. The average differences in  $D_{95\%}$  and  $D_{5\%}$  indices were  $-0.43\% \pm 1.27\%$  and  $-0.41\% \pm 1.05\%$ , respectively. The average difference in  $D_{95\%} - D_{5\%}$  was  $0.02\% \pm 1.12\%$ . For all patients, the difference between DE1 and MC2 in CTV mean dose and  $D_{95\%}$  was less than 3% of the prescription dose. The largest difference was found in patient 11 (breast), for which the DE1 CTV mean dose and  $D_{95\%}$  were both 2.1% lower than the corresponding MC2 indices.

Table 4 contains additional DVH indices for the selected OARs. Overall, 94.4% of the DE1 DVH indices were within 10% of the corresponding MC2 DVH indices. In cases where DE1 indices were different from MC2, the TPS3 DVH indices were either consistent with or worse than DE1 when compared with MC2. These differences were observed in particular DVH indices used for the prostate, lung, and craniospinal sites.

**Table 3.** DVH Indices Computed From CTV Dose Distributions.<sup>a</sup>

Patient	Mean Dose (%)		$D_{95\%}$ (%)		$D_{5\%}$ (%)		$D_{95\%} - D_{5\%}$ (%)	
	DE1	MC2	DE1	MC2	DE1	MC2	DE1	MC2
1	102.1	101.6	100.5	99.9	104.1	103.5	3.6	3.6
2	102.0	101.3	99.5	99.2	104.6	103.7	5.1	4.5
3	102.8	101.7	101.3	99.6	104.3	104.1	3.0	4.5
4	102.0	101.3	100.6	98.9	103.5	103.3	2.9	4.4
5	103.0	104.1	97.7	100.0	107.6	108.1	9.9	8.1
6	104.7	105.2	100.5	101.4	108.8	109.2	8.3	7.8
7	103.8	104.0	99.3	100.0	107.5	108.0	8.2	8.0
8	103.2	105.3	100.9	102.6	107.3	108.2	6.4	5.6
9	101.6	102.5	99.3	100.1	104.5	105.6	5.2	5.5
10	100.8	101.6	99.0	99.1	102.6	104.2	3.6	5.1
11	105.0	107.2	99.9	102.0	109.0	111.8	9.1	9.8
12	103.3	103.7	98.3	99.1	110.0	109.0	11.7	9.9

Abbreviations: CTV, clinical target volume; DE1, analytical dose engine; DVH, dose–volume histogram; MC2, fast Monte Carlo code.

<sup>a</sup>From left to right: columns show the mean CTV dose,  $D_{95\%}$ ,  $D_{5\%}$ , and  $D_{95\%} - D_{5\%}$  for DE1 and MC2 CTV dose. All doses are relative to the prescription dose.

**Table 4.** Comparison of DVH Indices for OARs Calculated in DE1 and MC2 for Each Patient Treatment Plan.<sup>a</sup>

DVH Index	Objective	DE1	MC2
Patient 1 (prostate)			
Bladder $D_{2cc}$ , Gy (RBE)	<81	78.9	81.7
Bladder $V_{40Gy}$ , %	$\leq 33$	3.8	3.7
Femoral head mean dose (L/R), Gy (RBE)	$\leq 26.6$	20.5/20.5	20.7/20.8
Femoral head $V_{15Gy}$ (L/R), %	$\leq 90$	65.7/65.7	65.1/65.0
Rectum $D_{2cc}$ , Gy (RBE)	<79.5	79.5	81.7
Rectum $V_{50Gy}$ , %	$\leq 24$	8.1	8.3
Patient 2 (prostate)			
Bladder $D_{2cc}$ , Gy (RBE)	<81	73.1	74.3
Bladder $V_{40Gy}$ , %	$\leq 33$	15.2	15.3
Femoral head mean dose (L/R), Gy (RBE)	$\leq 26.6$	17.6/17.5	17.9/17.7
Femoral head $V_{15Gy}$ (L/R), %	$\leq 90$	63.2/60.1	62.8/59.6
Rectum $D_{2cc}$ , Gy (RBE)	<79.5	72.0	73.2
Rectum $V_{50Gy}$ , %	$\leq 24$	7.8	8.3
Patient 3 (prostate)			
Bladder $D_{2cc}$ , Gy (RBE)	<81	79.4	81.1
Bladder $V_{40Gy}$ , %	$\leq 33$	7.3	7.4
Femoral head mean dose (L/R), Gy (RBE)	$\leq 26.6$	14.3/14.5	14.5/14.7
Femoral head $V_{15Gy}$ (L/R), %	$\leq 90$	46.7/47.7	46.3/47.3
Rectum $D_{2cc}$ , Gy (RBE)	<79.5	70.0	75.7
Rectum $V_{50Gy}$ , %	$\leq 24$	2.6	2.7
Patient 4 (prostate)			
Bladder $D_{2cc}$ , Gy (RBE)	<81	80.0	82.0
Bladder $V_{40Gy}$ , %	$\leq 33$	6.9	6.8
Femoral head mean dose (L/R), Gy (RBE)	$\leq 26.6$	20.5/21.3	20.7/21.5
Femoral head $V_{15Gy}$ (L/R), %	$\leq 90$	67.4/69.1	66.9/68.8
Rectum $D_{2cc}$ , Gy (RBE)	<79.5	58.6	61.8
Rectum $V_{50Gy}$ , %	$\leq 24$	1.8	1.9

(continued)



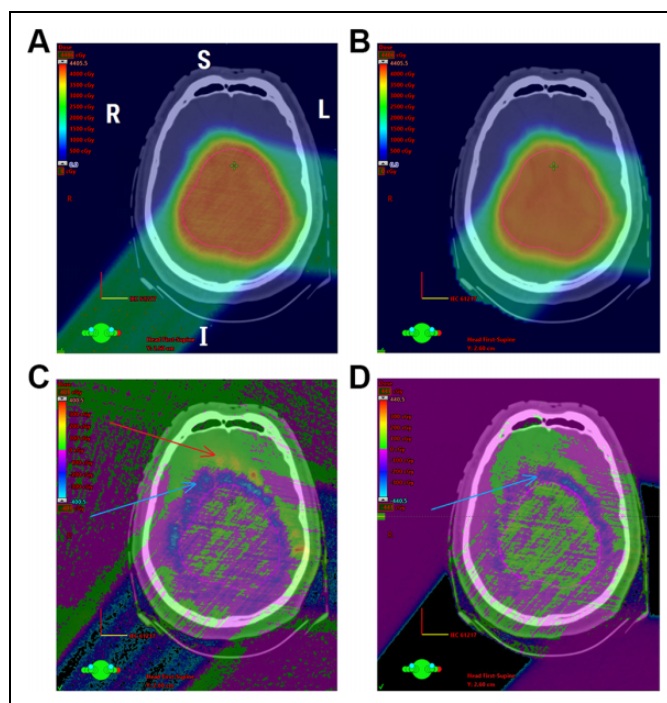
**Table 4.** (continued)

DVH Index	Objective	DE1	MC2
<b>Patient 5 (head and neck)</b>			
Cochlea $D_{90\%}$ (L/R), Gy (RBE)	<35	20.0/21.1	19.5/21.0
L optic nerve <sup>b</sup> max dose ( $D_{1\%}$ ), Gy (RBE)	≤50	42.6	41.4
Eye mean dose (L/R), Gy (RBE)	≤20	1.0/18.7	1.0/18.9
Spinal cord max dose ( $D_{1\%}$ ), Gy (RBE)	≤45	5.9	5.6
<b>Patient 6 (head and neck)</b>			
Cochlea $D_{90\%}$ (L/R), Gy (RBE)	≤35	10.0/0.1	10.0/0.1
Brain stem max dose ( $D_{1\%}$ ), Gy (RBE)	≤50	2.5	2.3
Spinal cord max dose ( $D_{1\%}$ ), Gy (RBE)	≤45	31.3	30.8
Oral cavity mean dose, Gy (RBE)	≤50	12.1	12.0
<b>Patient 7 (head and neck)</b>			
Cochlea $D_{90\%}$ (L/R), Gy (RBE)	<35	22.8/25.2	21.9/25.5
Optic chiasm $D_{10\%}$ , Gy (RBE)	<58	25.8	25.0
Eye mean dose (L/R), Gy (RBE)	≤20	13.9/18.9	13.4/19.3
Spinal cord max dose ( $D_{1\%}$ ), Gy (RBE)	≤45	27.3	26.7
<b>Patient 8 (lung)</b>			
Mean lung dose, Gy (RBE)	<20	7.6	7.9
Lung $V_{20\%}$ , %	<35	15.6	15.9
Spinal cord max dose ( $D_{1\%}$ ), Gy (RBE)	<50	23.6	24.3
Heart mean dose, Gy (RBE)	<26	1.1	1.2
<b>Patient 9 (lung)</b>			
Mean lung dose, Gy (RBE)	<20	6.6	6.9
Lung $V_{20\%}$ , %	<35	14.5	15.1
Spinal cord max dose ( $D_{1\%}$ ), Gy (RBE)	<50	31.7	32.4
Heart mean dose, Gy (RBE)	<26	0.0	0.0
<b>Patient 10 (brain)</b>			
Brain stem max dose ( $D_{1\%}$ ), Gy (RBE)	<40	34.4	34.7
Optic chiasm max dose ( $D_{1\%}$ ), Gy (RBE)	<40	34.5	36.6
Cochlea max dose (L/R) ( $D_{1\%}$ ), Gy (RBE)	<45	7.6/0.3	8.0/0.3
Cochlea mean dose (L/R), Gy (RBE)	<35	4.7/0.2	5.1/0.2
<b>Patient 11 (breast)</b>			
Spinal cord $D_{0.01cc}$ , Gy (RBE)	≤36	2.5	1.6
Right lung $V_{40\%}$ , %	≤20	14.7	13.0
Esophagus $D_{0.01cc}$ , Gy (RBE)	≤45	52.5	51.8
Thyroid $D_{1cc}$ , Gy (RBE)	≤50	51.8	50.7
<b>Patient 12 (craniospinal)</b>			
Brain stem $D_{10\%}$ , Gy (RBE)	<54	36.0	36.5
Spinal cord $D_{10\%}$ , Gy (RBE)	<59	36.7	37.7
Optic chiasm $D_{50\%}$ , Gy (RBE)	<56	37.7	37.6
Eye $D_{50\%}$ (L/R), Gy (RBE)	<20	11.8/10.8	10.6/9.8

Abbreviations: DE1, analytical dose engine; DVH, dose–volume histogram; L, left; MC2, fast Monte Carlo code; OARs, organs at risk; R, right; RBE, relative biological effectiveness.

<sup>a</sup>Site-specific DVH indices are listed for each patient, and clinical objective for the index, the DE1-calculated index, and the MC2-calculated index are listed in the columns to the right.

<sup>b</sup>Loss of vision in the right eye was expected due to tumor location.



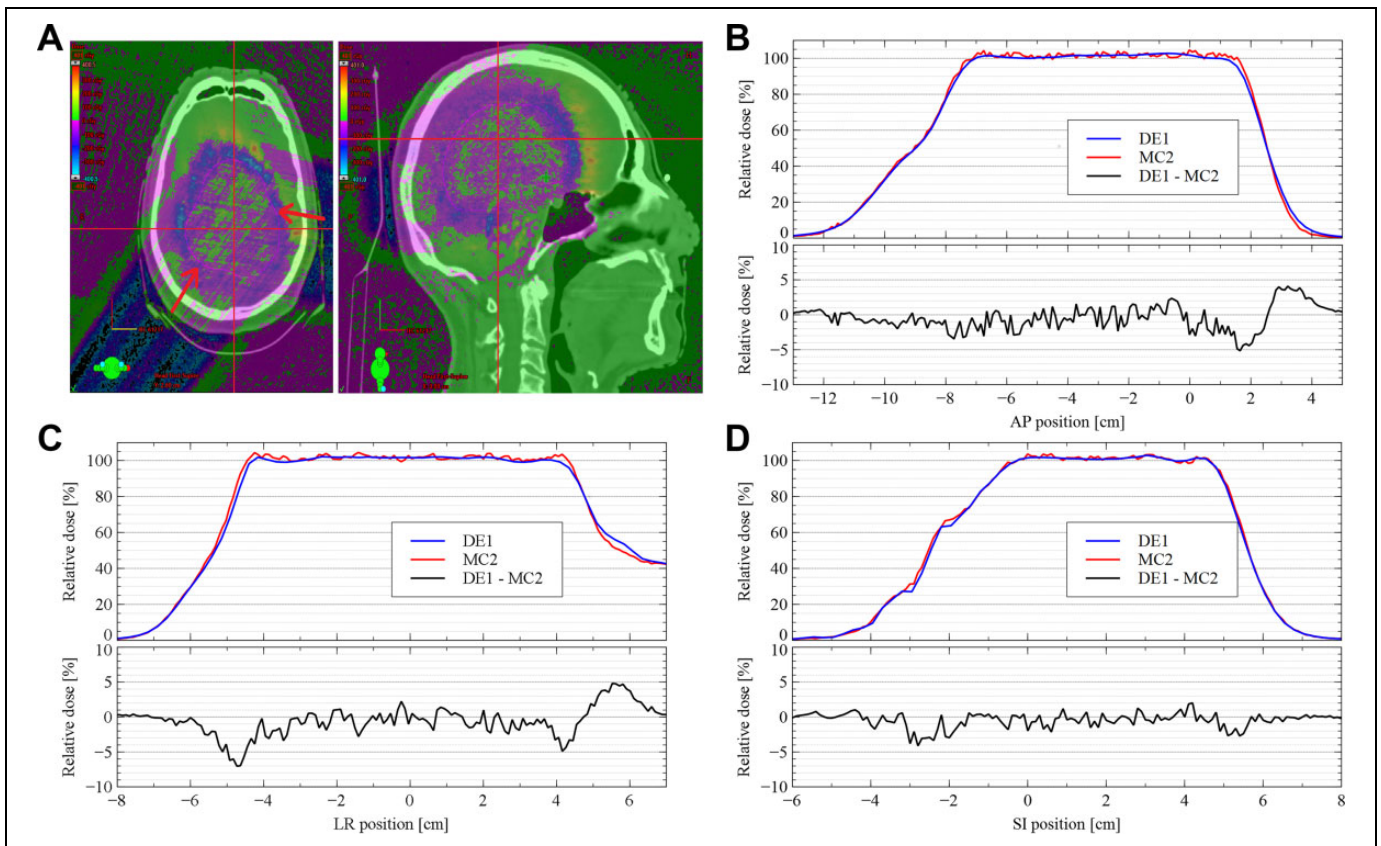
**Figure 5.** Comparison of dose distributions for patient 10 using the ERS45 machine calculated by (A) MC2 and (B) DE1. C, The difference between the MC2 and DE1 dose distributions ( $DE1 - MC2$ ). The color wash range is  $\pm 10\%$  of prescription dose. D, The difference between the MC2 and TPS3 dose distributions ( $TPS3 - MC2$ ). Arrows point to dose differences that are referenced in the text. DE1 indicates analytical dose engine; ERS, extended range shifter; MC2, fast Monte Carlo code; TPS, treatment planning system.

In each of the prostate patients, bladder and rectum  $D_{2cc}$  was underestimated by DE1. For patient 1, the difference in the bladder and rectum was about 3%; as a result, DE1 incorrectly reported that  $D_{2cc}$  met the objective for these OARs. However, TPS3 also underestimated these indices for each prostate patient; the patient 1 bladder  $D_{2cc}$  calculated by TPS3 was 78.5 Gy (RBE) and the rectum  $D_{2cc}$  was 78.7 Gy (RBE).

For patients 8 and 9, the mean lung dose calculated by DE1 was 3.6% and 4.6% lower than MC2, respectively. On the other hand, the TPS3 mean lung doses were 1.7 Gy (RBE) for patient 8 and 6.6 Gy (RBE) for patient 9, both of which were less than the corresponding DE1 and MC2 indices. In the case of patient 8, TPS3 underestimated the dose from a posterior-anterior (PA) field distal to the CTV in the normal lung tissue.

The left and right eyes of patient 12 had  $D_{50\%}$  values that are 10% higher in DE1 than in MC2. In this treatment plan, the patient's eyes were situated at the distal edge of the 2 oblique PA fields used for the craniospinal irradiation. The TPS3  $D_{50\%}$  for the left and right eye of patient 12 was 8.8 Gy (RBE) and 8.2 Gy (RBE), respectively, which was more than 15% lower than MC2.

In 3 of the 4 DE1 DVH indices that failed the validation criteria (patient 11 spinal cord  $D_{0.01cc}$  and right lung  $V_{40\%}$  and patient 12 left and right eye  $D_{50\%}$ ), the DE1 index was closer to



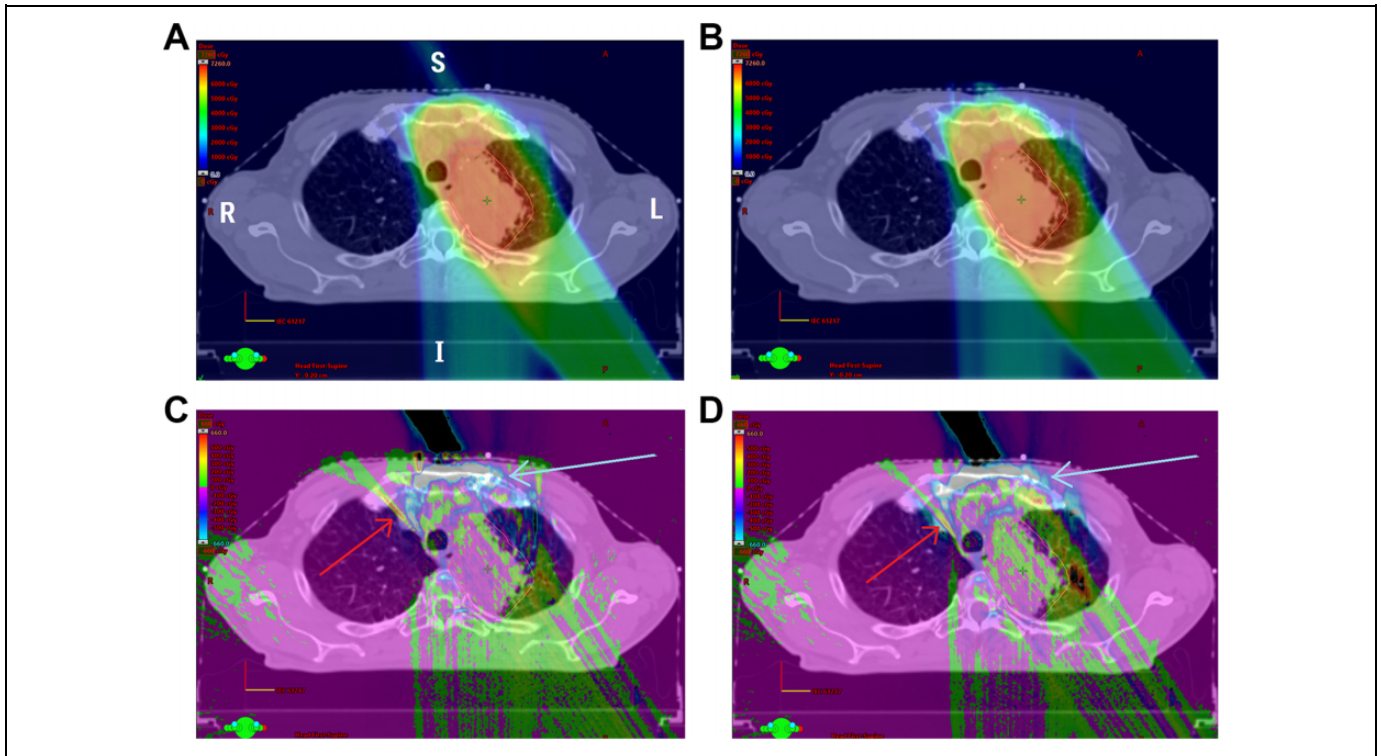
**Figure 6.** Comparison of inpatient DE1 and MC2 relative dose profiles for patient 10. A, Axial (left) and sagittal (right) planes of the DE1 – MC2 dose distribution. Arrows on the axial view indicate the directions of the 2 beams, and the lines show the positions of the (B) AP, (C) LR, and (D) SI dose profiles. For each of these profiles, the top graph directly compares DE1 and MC2 dose profiles and the bottom graph is the dose difference DE1 – MC2. All distances are measured from isocenter. AP indicates anterior to posterior; DE1, analytical dose engine; LR, left to right; MC2, fast Monte Carlo code; SI, superior to inferior.

MC2 than the TPS3 value. However, DE1 failed and TPS3 was closer to MC2 for the patient 11 right lung  $V_{40\%}$ . In this case, the MC2 index was 13.0%, the DE1 index was 14.7%, and the TPS3 index was 14.1%.

Figure 5 is a color wash comparison of plan dose on 1 axial slice of patient 10. This treatment plan contained 2 fields and used the ERS45 machine. Figure 5A is total plan dose as calculated by MC2 and Figure 5B is total plan dose calculated by DE1. Anatomical directions on the axial slice are shown in Figure 5A. The color range is from 0% to 110% of the prescription dose (40.05 Gy), and the CTV is outlined in magenta in all plots. Figure 5C shows the voxel-by-voxel dose difference (DE1 – MC2) in a color wash representing differences between –10% and +10% of prescription dose. Within the CTV, the difference between DE1 and MC2 dose was less than 3%. However, larger differences were seen at the distal edges of the CTV (relative to the 2 field directions). For example, at the distal end of the field coming from the lower left corner, DE1 dose was first up to 10% lower than MC2 (blue arrow) and then up to 10% larger than MC2 (red arrow). There were no similar discrepancies at proximal or lateral margins of the CTV. A similar but less drastic effect was seen in the dose difference

(TPS3 – MC2) in Figure 5D; nevertheless, TPS3 dose still underestimated dose to the distal margin of the CTV (the region indicated by a blue arrow). At some points, TPS3 was also up to 10% less than MC2.

Figure 6 shows orthogonal dose profiles of patient 10 which intersect at a point within the CTV. Figure 6A shows the difference between DE1 and MC2 on the axial dose plane from Figure 5C (left) and a sagittal dose plane (right). The anterior-to-posterior (AP) dose profile in Figure 6B is the vertical red line on the axial plane, and the left-to-right (LR) dose profile in Figure 6C is the horizontal red line. The vertical red line on the sagittal plane is the superior-to-inferior (SI) profile in Figure 6D, while the horizontal line is the AP dose profile. For each profile, the top panel contains the DE1 dose profile in blue and the MC2 profile in red. Positive directions are anterior, right, and superior, respectively. The bottom panel shows the dose difference (DE1 – MC2) as a black line. All positions are measured from isocenter, and doses are relative to prescription dose. Dose differences within and near the CTV were generally less than 3%. However, at the distal CTV margins located near +2 cm in the AP profile and near –5 cm and +4 cm in the LR profile, there were differences of 5% between DE1 and MC2 profiles.



**Figure 7.** Comparison of dose distributions for patient 9 using the ERS45 machine calculated by (A) MC2 and (B) DE1. C, The difference between the MC2 and DE1 dose distributions (DE1 – MC2). The color wash range is  $\pm 10\%$  of prescription dose. D, The difference between the MC2 and TPS3 dose distributions (TPS3 – MC2). Arrows point to dose differences that are referenced in the text. DE1 indicates analytical dose engine; MC2, fast Monte Carlo code; TPS, treatment planning system.

Figure 7 is a color wash comparison of plan dose on 1 axial slice of patient 9. Figure 7A shows MC2 dose, Figure 7B shows DE1 dose, and Figure 7C is the dose difference (DE1 – MC2). The difference in dose (TPS3 – MC2) in Figure 7D is provided for comparison. Anatomical directions in the slice are labeled in Figure 7A. Color wash ranges are from 0% to 110% of prescription dose in Figure 7A and B and from  $-10\%$  to  $+10\%$  of prescription dose in Figure 7C and D. In each plot, the CTV within the lung is outlined in magenta. Although it is surrounded by low-density tissue, MC2 and DE1 agree within the CTV. However, both DE1 and TPS3 underestimate dose downstream of the lung volume (indicated by blue arrows). Both analytical calculations also overestimated dose at the anterior edge of the right lung (indicated by red arrows), which was downstream of a density heterogeneity parallel to the beam originating at the lower right corner of the image.

Figure 8 shows orthogonal dose profiles of patient 9. Figure 8A shows the dose difference (DE1 – MC2) in both the axial plane from Figure 7C and a sagittal plane. Red lines are used to indicate the positions of the AP (Figure 8B), LR (Figure 8C), and SI (Figure 8D) profiles. Doses are relative to the prescription dose. Within the CTV, the dose difference is less than 3%. Dose discrepancy in the low-density tissue can be observed in the LR profile, in which the boundary between the solid tumor and the lung tissue occurs at approximately  $+3$  cm and  $-3$  cm. At positions beyond  $+3$  cm (within the left lung), there is up to

a 10% difference between DE1 and MC2 dose. The underestimation of dose downstream of the lung volume noted in Figure 7 is evident in the AP dose profile at  $+7$  cm. The anterior boundary of the CTV is located at  $+4$  cm on the AP profile, after which DE1 dose is consistently lower than MC2 dose by up to 15%.

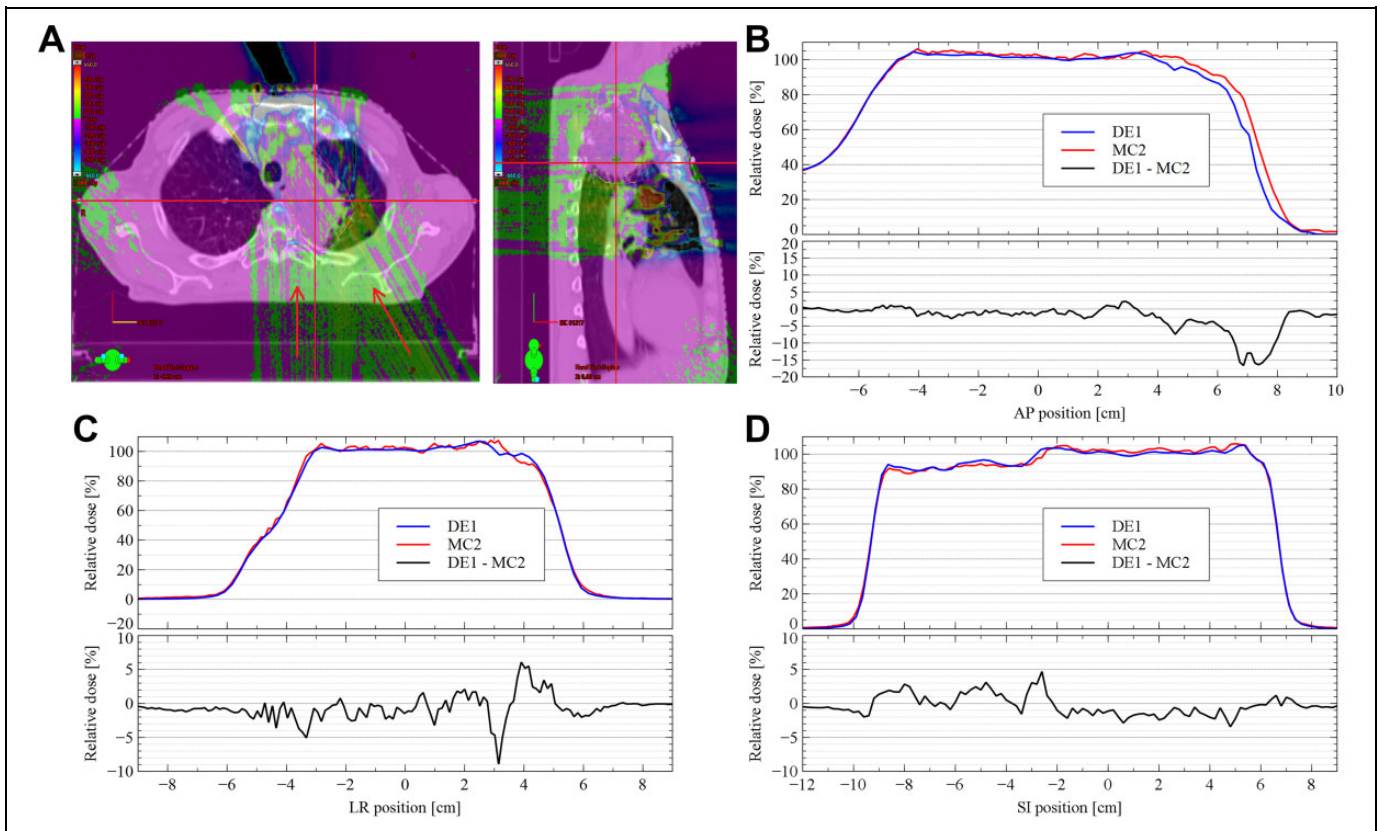
### Computational Efficiency of the DE1

Table 5 shows the time that was required to perform the DE1 plan dose calculation for each patient in the cohort. Dose calculation time for a treatment plan was nearly proportional to the product of the number of spots in the treatment plan and the volume of the patient's external structure BODY. For all patients except patient 12, plan dose was calculated in less than 4 minutes. Plan dose was calculated in less than 2 minutes for two-thirds of the plans in the cohort and in about 9 minutes for patient 12, who received a craniospinal irradiation. Since the calculation was parallelized, computational speed scaled with the number of additional processors.

### Discussion

This work described and validated the dose calculation algorithm for our in-house developed independent second-dose check software in proton delivery systems with FWHM





**Figure 8.** Comparison of inpatient DE1 and MC2 relative dose profiles for patient 9. A, Axial (left) and sagittal (right) planes of the DE1 – MC2 dose distribution. Arrows on the axial view indicate the directions of the 2 beams, and the lines show the positions of the (B) AP, (C) LR, and (D) SI dose profiles. For each of these profiles, the top graph directly compares DE1 and MC2 dose profiles and the bottom graph is the dose difference DE1 – MC2. All distances are measured from isocenter. AP indicates anterior to posterior; DE1, analytical dose engine; LR, left to right; MC2, fast Monte Carlo code; SI, superior to inferior.

**Table 5.** Time Required to Compute Total Plan Dose Using DE1.<sup>a</sup>

Patient	Number of Spots ( $10^3$ )	BODY Volume ( $10^3 \text{ cm}^3$ )	DE1 Calculation Time (s)
1	5.1	46.0	105
2	4.3	24.8	94
3	3.2	19.8	67
4	4.9	55.6	134
5	4.6	14.8	35
6	21.4	18.8	75
7	6.4	6.8	115
8	18.0	34.8	95
9	25.3	44.9	179
10	11.7	13.2	97
11	56.3	36.8	212
12	48.3	83.0	532

Abbreviation: DE1, analytical dose engine.

<sup>a</sup>Calculation time depended on the product of total number of spots and body volume, which are provided in the second and third columns.

Note: BODY is an external structure defined to enclose the whole region of interests.

between 5 and 14 mm and a negligible spray component. We want to emphasize that the major purpose of this article is not to propose a new dose calculation algorithm for PBS, but rather to

describe and validate the dose calculation algorithm for our in-house developed second-dose check software in PBS. Unlike photon therapy, so far there are no commercial solutions for second-dose check software in the market.

As DE1 is based on a ray-casting pencil beam algorithm, DE1 can provide an accurate second-dose check that is complimentary<sup>9</sup> to TPS3's fluence-dose algorithm and that can be executed on widely available computer systems. The dose engine features a computationally efficient ray-casting algorithm, a triple Gaussian LPF, and the use of lateral WET to improve performance in inhomogeneous geometries. Details of the commissioning procedures and validation work used for benchmarking were provided. This commissioning process produced a dose model for which the in-water  $\gamma$ -index analysis DE1 versus MX was above 99% on average and had less than 1% standard deviation (see "Accuracy of DE1 Calculation in Water" section). Analytical dose engine in-water 2D-3D  $\gamma$ -index analysis passing rates were at least 91.9% and passing rates for the 3D-3D comparison to MC2 in water were at least 97.6%. Passing rates for all but one 3D-3D comparison to MC2 in patient geometry were greater than 90%; however, TPS3 had a lower passing rate than DE1 for this patient. For all patients, there was less than 2.1% difference between DE1 and MC2 in

both CTV mean dose and CTV  $D_{95\%}$ . Comparing DE1 to MC2 DVH indices for OARs, 94.4% of DVH indices were within 10% of MC2 values. Plan dose calculation times were usually less than 2 minutes, although DE1 was executed on a workstation with far less computing power than the servers used to run TPS3.

Modeling the field size effects has proven to be challenging and time-consuming for some commercial TPSs based on the fluence-dose model.<sup>6,7</sup> Since DE1 uses the in-water commissioning data generated either from MC simulation or from direct measurements, the 1.5% agreement in SOBP measurements with varied field sizes is reflected in the 99.5% average passing rate between DE1 and MC2 for the 3D-3D  $\gamma$ -index analysis (see Figures 1–3). In the 2D-3D analysis comparing DE1 and TPS3 to MX, the average DE1 passing rates were indeed superior to TPS3 (see Figure 3). The lowest TPS3 passing rates in water were found in breast treatment fields, which had a large field size. Analytical dose engine had an advantage in computing these large field sizes, while the TPS3 fluence model was limited by the number of parameters available for commissioning.<sup>7</sup> As shown in Figure 1, the majority of SOBP measurements used the  $10 \times 10$  cm field size. It would be possible to further improve the DE1 model by incorporating additional SOBP measurements with varying field sizes.

The DE1 passing rates for the 3D-3D  $\gamma$ -index analysis comparison to MC2 in patient geometries were comparable to but slightly lower than the TPS3 passing rates (see Figure 4). The outlier was patient 9 (lung); color wash and dose profile comparisons show that both DE1 and TPS3 underestimated dose downstream of the low-density lung volume (see Figures 7 and 8). Some differences were seen distal to the CTV even in relatively homogeneous geometries (see Figures 5 and 6). Nevertheless, CTV and OAR DVH indices were consistent with MC2-calculated indices. This is mainly because beam angles normally selected for treatment tend to avoid putting OARs distal to the CTV due to uncertainties in proton range and RBE. In cases where DE1 disagreed with MC2 such as the bladder and rectum  $D_{2cc}$  for patients with prostate disease, TPS3 matched the DE1 result, which indicated that some aspect of the geometry was likely problematic for both analytical codes. This is consistent with recent studies that have shown that analytic pencil beam algorithms do not perform as well as MC in heterogeneous geometries, including the Imaging and Radiation Oncology Core (IROC) lung phantom<sup>11</sup> as well as thoracic, prostate, and head and neck disease sites.<sup>10,12</sup> Our study thus agrees with Schaffner *et al*'s recommendation that MC simulations should be used as a final-dose check in highly complex geometries.<sup>9</sup>

Our dose engine used an altered ray-casting algorithm that is complimentary to the fluence-dose algorithm used in our commercial TPS with respect to the diverse set of heterogeneous patient geometries encountered in clinical practice.<sup>9</sup> This allowed our engine to provide an independent second-dose calculation check at our institution. For over 2 years, DE1 has been used as a second-dose check after plan optimization and to

calculate verification plan dose for every treatment plan as part of patient-specific quality assurance.<sup>39</sup>

The inclusion of water equivalent lateral distance resulted in adequate DE1 performance in patient geometry, although it was slightly less accurate than TPS3. This modification of the ray-casting algorithm was selected to determine whether or not more complex modifications were necessary. However, DE1 can be used as a platform to develop other methods that use different modifications of the ray-casting algorithm. Future work might also include other ways to improve accuracy in heterogeneous geometries such as a voxel-by-voxel scattering standard deviation calculation.<sup>28</sup> The effect of RSs on the spot width and the implications for a ray-casting algorithm should also be studied in greater detail. It is possible that an additional Gaussian component could better model RS scattering, just as a double Gaussian in-air fluence model has been used to model spray in the fluence-dose algorithm.<sup>6</sup> Improvements in the ray-casting algorithm, if able to substantially increase accuracy without loss of computational efficiency, would further improve DE1's effectiveness as a second-dose calculation check and as the basis for an internally developed TPS. We intend for this dose engine to form the basis of many future studies including an analysis of alternative ray-casting transformations.

## Conclusions

The results of our validation tests of DE1 were that (1) all treatment fields passed the validation criteria for the 2D-3D  $\gamma$ -index analysis comparison to MX in water, (2) all treatment fields passed the validation criteria for the 3D-3D  $\gamma$ -index analysis comparison to MC2 in water, (3) all plan doses passed the validation criteria for the 3D-3D  $\gamma$ -index analysis comparison to MC2 in patient geometry except 1 lung patient, (4) all CTV DVH indices passed the validation criteria of <3% dose difference, and (5) 68 of 72 DVH indices of OARs passed the validation criteria of <10% difference. As a result, DE1 was validated for calculating dose in water (eg, for verification plan dose calculation for patient-specific quality assurance). We also validated DE1 as a dose calculation engine in patient geometry for the following disease sites: prostate, head and neck, brain, breast, and craniospinal. Analytical dose engine was not validated for use in lung sites, and MC2 continues to be used as a final dose check for lung. In addition, we found that MC2 should be used to check DVH indices for OARs whenever they involve very small volumes (ie, point dose or maximum dose) or when the OAR is positioned distal to the target volume.

The major purpose of this article was to describe and validate the dose calculation algorithm for our in-house-developed second-dose check software—not to propose a new dose calculation algorithm for PBS. Unlike conventional radiotherapy, there are no commercial solutions for second-dose check software currently available on the market. Analytical dose engine was able to complete two-thirds of the dose calculations in 2 minutes with very modest hardware requirements. Since it uses

a ray-casting algorithm, DE1 provided an independent second-dose check for the fluence-dose algorithm used by our commercial TPS. However, we concur with previous studies that plan dose in highly complex geometries (such as the lung treatment site) should be independently verified by a well-benchmarked MC simulation.

### Authors' Note

Patient information was gathered retrospectively (posttreatment) for the purposes of this work. Patient consent forms are waived following a protocol approved by the institutional review board of Mayo Clinic (IRB #: 13-005709).

### Acknowledgments

The authors wish to thank Kurt Augustine, Michael Davis, and Dr Mirek Fatyga for developing the web interface for our dose engine and for other associated work. The authors are grateful to Dr Chris Beltran, Dr Hok Wan Chan Tseung, Dr Daniel Mundy, and the rest of the team at Mayo Clinic Rochester who developed the fast Monte Carlo code used in this study. The authors also wish to thank Dr Yixiu Kang for helpful conversation.


### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Cancer Institute (NCI) Career Developmental Award [K25CA168984]; the Arizona Biomedical Research Commission Investigator Award; The Lawrence W. and Marilyn W. Matteson Fund for Cancer Research; and The Kemper Marley Foundation.

### ORCID iD

James E. Younkin  <https://orcid.org/0000-0002-8385-3716>

### References

- Gottschalk B, Cascio EW, Daartz J, Wagner MS. On the nuclear halo of a proton pencil beam stopping in water. *Phys Med Biol.* 2015;60(14):5627-5654.
- Clasie B, Depauw N, Franssen M, et al. Golden beam data for proton pencil beam scanning. *Phys Med Biol.* 2012;57(5):1147-1158.
- Lin L, Ainsley CG, McDonough JE. Experimental characterization of two-dimensional pencil beam scanning proton spot profiles. *Phys Med Biol.* 2013;58(17):6193-6204.
- Lin L, Ainsley CG, Mertens T, De Wilde O, Talla PT, McDonough JE. A novel technique for measuring the low-dose envelope of pencil-beam scanning spot profiles. *Phys Med Biol.* 2013;58(12):N171-180.
- Pedroni E, Scheib S, Böhringer T, et al. Experimental characterization and physical modelling of the dose distribution of scanned proton pencil beams. *Phys Med Biol.* 2005;50(3):541-561.
- Zhu XR, Poenisch F, Lii M, et al. Commissioning dose computation models for spot scanning proton beams in water for a commercially available treatment planning system. *Med Phys.* 2013;40(4):041723.
- Shen J, Liu W, Stoker J, et al. An efficient method to determine double Gaussian fluence parameters in the ECLIPSE™ proton pencil beam model. *Med Phys.* 2016;43(12):6544.
- Gottschalk B, Koehler A, Schneider R, Sisterson J, Wagner M. Multiple coulomb scattering of 160 MeV protons. *Nucl Instrum Methods Phys Res B.* 1993;74(4):467-490.
- Schaffner B, Pedroni E, Lomax A. Dose calculation models for proton treatment planning using a dynamic beam delivery system: an attempt to include density heterogeneity effects in the analytical dose calculation. *Phys Med Biol.* 1999;44(1):27-41.
- Schuemann J, Giantsoudi D, Grassberger C, Moteabbed M, Min CH, Paganetti H. Assessing the clinical impact of approximations in analytical dose calculations for proton therapy. *Int J Radiat Oncol Biol Phys.* 2015;92(5):1157-1164.
- Taylor PA, Kry SF, Followill DS. Pencil beam algorithms are unsuitable for proton dose calculations in lung. *Int J Radiat Oncol Biol Phys.* 2017;99(3):750-756.
- Yepes P, Adair A, Grosshans D, et al. Comparison of Monte Carlo and analytical dose computations for intensity modulated proton therapy. *Phys Med Biol.* 2018;63(4):045003.
- Beltran C, Jia Y, Slopsma R, Yeung D, Li Z. A simplified methodology to produce Monte Carlo dose distributions in proton therapy. *J Appl Clin Med Phys.* 2014;15(4):2-10.
- Wan Chan Tseung H, Ma J, Beltran C. A fast GPU-based Monte Carlo simulation of proton transport with detailed modeling of nonelastic interactions. *Med Phys.* 2015;42(6):2967-2978.
- Hong L, Goitein M, Bucciolini M, et al. A pencil beam algorithm for proton dose calculations. *Phys Med Biol.* 1996;41(8):1305-1330.
- Frühwirth R, Regler M. On the quantitative modelling of core and tails of multiple scattering by Gaussian mixtures. *Nucl Instrum Meth A.* 2001;456(3):369-389.
- Kusano Y, Kanai T, Kase Y, et al. Dose contributions from large-angle scattered particles in therapeutic carbon beams. *Med Phys.* 2007;34(1):193-198.
- Inaniwa T, Furukawa T, Nagano A, et al. Field-size effect of physical doses in carbon-ion scanning using range shifter plates. *Med Phys.* 2009;36(7):2889-2897.
- Inaniwa T, Kanematsu N, Hara Y, et al. Implementation of a triple Gaussian beam model with subdivision and redefinition against density heterogeneities in treatment planning for scanned carbon-ion radiotherapy. *Phys Med Biol.* 2014;59(18):5361-5386.
- Hirayama S, Takayanagi T, Fujii Y, et al. Evaluation of the influence of double and triple Gaussian proton kernel models on accuracy of dose calculations for spot scanning technique. *Med Phys.* 2016;43(3):1437-1450.
- Bellinzona VE, Ciocca M, Embriaco A, et al. On the parametrization of lateral dose profiles in proton radiation therapy. *Phys Med.* 2015;31(5):484-492.
- Li Y, Zhu RX, Sahoo N, Anand A, Zhang X. Beyond Gaussians: a study of single-spot modeling for scanning proton dose calculation. *Phys Med Biol.* 2012;57(4):983-997.
- Mackin D, Li Y, Taylor MB, et al. Improving spot-scanning proton therapy patient specific quality assurance with HPlusQA, a

- second-check dose calculation engine. *Med Phys.* 2013;40(12):121708.
24. Würfl M, Englbrecht F, Parodi K, Hillbrand M. Dosimetric impact of the low-dose envelope of scanned proton beams at a ProBeam facility: comparison of measurements with TPS and MC calculations. *Phys Med Biol.* 2016;61(2):958-973.
25. Gillin MT, Sahoo N, Bues M, et al. Commissioning of the discrete spot scanning proton beam delivery system at the University of Texas MD Anderson Cancer Center, proton therapy center, Houston. *Med Phys.* 2010;37(1):154-163.
26. Shen J, Tryggestad E, Younkin JE, et al. Technical Note: Using experimentally determined proton spot scanning timing parameters to accurately model beam delivery time. *Med Phys.* 2017;44(10):5081-5088.
27. Both S, Shen J, Kirk M, et al. Development and clinical implementation of a universal bolus to maintain spot size during delivery of base of skull pencil beam scanning proton therapy. *Int J Radiat Oncol Biol Phys.* 2014;90(1):79-84.
28. Soukup M, Fippel M, Alber M. A pencil beam algorithm for intensity modulated proton therapy derived from Monte Carlo simulations. *Phys Med Biol.* 2005;50(21):5089.
29. Dagum L, Menon R. OpenMP: an industry-standard API for shared-memory programming. *IEEE Comput Sci Eng.* 1998; 5(1):46-55.
30. Bortfeld T. An analytical approximation of the Bragg curve for therapeutic proton beams. *Med Phys.* 1997;24(12):2024-2033.
31. Zhang X, Liu W, Li Y, et al. Parameterization of multiple Bragg curves for scanning proton beams using simultaneous fitting of multiple curves. *Phys Med Biol.* 2011;56(24):7725-7735.
32. Schaffner B. Proton dose calculation based on in-air fluence measurements. *Phys Med Biol.* 2008;53(6):1545-1562.
33. Highland VL. Some practical remarks on multiple scattering. *Nucl Instrum Methods.* 1975;129(2):497-499.
34. Ding X, Liu W, Shen J, et al. Use of a radial projection to reduce the statistical uncertainty of spot lateral profiles generated by Monte Carlo simulation. *J Appl Clin Med Phys.* 2017;18(6):88-96.
35. Allison J, Amako K, Apostolakis J, et al. Geant4 developments and applications. *IEEE Trans Nucl Sci.* 2006;53(1):270-278.
36. Cirrone GP, Cuttone G, Mazzaglia SE, et al. Hadron therapy: a geant4-based tool for proton/ion-therapy studies. *Prog Nucl Sci Technol.* 2011;2:207-212.
37. Allison J, Amako K, Apostolakis J, et al. Recent developments in geant4. *Nucl Instrum Meth A.* 2016;835:186-225.
38. Ulmer W, Schaffner B. Foundation of an analytical proton beamlet model for inclusion in a general proton dose calculation system. *Radiat Phys Chem.* 2011;80(3):378-389.
39. Hernandez Morales D, Shan J, Liu W, et al. Automation of routine elements for spot-scanning proton patient-specific quality assurance. *Med Phys.* 2019; 46(1):5-14.
40. Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys.* 1998; 25(5):656-661.
41. Miften M, Olch A, Mihailidis D, et al. Tolerance limits and methodologies for IMRT measurement-based verification QA: recommendations of AAPM task group No. 218. *Med Phys.* 2018;45(4): e53-e83.