sciendo Radiology and Oncology | Ljubljana | Slovenia | www.radioloncol.com RADIOLOGY AND ONCOLOGY

research article

The dose accumulation and the impact of deformable image registration on dose reporting parameters in a moving patient undergoing proton radiotherapy

Gasper Razdevsek[1], Urban Simoncic[1,2], Luka Snoj[2,1], Andrej Studen[1,2]

[1] Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana, Slovenia
[2] Jožef Stefan Institute, Ljubljana, Slovenia

Radiol Oncol 2022; 56(2): 248-258.

Received 11 April 2021
Accepted 18 February 2022

Correspondence to: Assist. Prof. Andrej Studen, Ph.D., Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia. E-mail: andrej.studen@fmf.uni-lj.si

Disclosure: No potential conflicts of interest were disclosed.

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction. Potential changes in patient anatomy during proton radiotherapy may lead to a deviation of the delivered dose. A dose estimate can be computed through a deformable image registration (DIR) driven dose accumulation. The present study evaluates the accumulated dose uncertainties in a patient subject to an inadvertent breathing associated motion.

Materials and methods. A virtual lung tumour was inserted into a pair of single participant landmark annotated computed tomography images depicting opposite breathing phases, with the deep inspiration breath-hold the planning reference and the exhale the off-reference geometry. A novel Monte Carlo N-Particle, Version 6 (MCNP6) dose engine was developed, validated and used in treatment plan optimization. Three DIR methods were compared and used to transfer the exhale simulated dose to the reference geometry. Dose conformity and homogeneity measures from International Committee on Radioactivity Units and Measurements (ICRU) reports 78 and 83 were evaluated on simulated dose distributions registered with different DIR algorithms.

Results. The MCNP6 dose engine handled patient-like geometries in reasonable dose calculation times. All registration methods were able to align image associated landmarks to distances, comparable to voxel sizes. A moderate deterioration of ICRU measures was encountered in comparing doses in on and off-reference anatomy. There were statistically significant DIR driven differences in ICRU measures, particularly a 10% difference in the relative $D_{98\%}$ for planning tumour volume and in the 3 mm/3% gamma passing rate.

Conclusions. The dose accumulation over two anatomies resulted in a DIR driven uncertainty, important in reporting the associated ICRU measures for quality assurance.

Key words: proton therapy; adaptive therapy; MCNP6; dose distribution measurement; Monte-Carlo; dose homogeneity; image registration

Introduction

Proton therapy is a well-established technique for treating cancer, exploiting a favourable depth-dose distribution of protons allowing greater sparing of normal tissues and improved local tumour control.[1,2] Evidence of a clinical advantage over photon therapy for certain disease sites and treatment scenarios exists, but often, the studies are inconclusive or show similar results for proton and other radiotherapy treatments.[3-5] To translate apparent benefits to clinical outcome, tight control over dose delivery and underlying uncertainties is required.[6,7]

Motion is a considerable problem in radiation treatment, resulting in deviations in the delivered dose from the prescribed dose plan.[8,9] This problem escalates in proton therapy due to the sensitivity of the proton path on tissue composition.[10-12] To guarantee target dose coverage and compliance of dose constraints for close-by critical structures, motion should be assessed and monitored, ideally prior to and during treatment.[13]

As the most comprehensive solution, magnetic resonance imaging (MRI) during dose delivery was suggested, but yet, there is no proton center operating an in-beam MRI due to associated technical and operational difficulties.[14-16] Standard modalities such as CT, are not considered viable as continuous in-beam monitoring techniques lead to dis-proportional radiation burden.[17] As a surrogate, 4D CT imaging is performed prior to the therapy. The 4D CT protocol dictates image collection synchronous to an independent breathing signal, division of the breathing cycle to several breathing phases (e.g., ten phases) and averaging of images within the phase to improve statistical properties and compensate for the blurring artifacts.[18,19]

Dose verification is a critical step in evaluating treatment success as even a small error in treatment planning, delivery, or dosimetry can lead to negative consequences and is often overlooked.[20] International Committee on Radioactivity Units and Measurements (ICRU) standards provide means for reliable dose reporting, which can be extended to non-static geometry should an appropriate algorithm for dose accumulation exist.[21,22]

The adaptive radiation therapy (ART) aims at compensating motion related dose deviation through adaptation of treatment plan to instant anatomy. Frequency in treatment adaptations vary: inter- or intra-fractional interventions are considered, particularly in recent studies.[23,24] The real-time Adaptive Particle Therapy of Cancer (RAPTOR) initiative is a collaborative effort to standardize and harmonize adaptive proton therapy, initiated by leading global proton and particle therapy centres, universities, vendors and proton therapy-related industries. As the initial goal, the impact of daily adaptations on treatment course was set, but the developed methodology already incorporates an extension to intra-fractional treatment adjustments.[25]

The critical components of adaptive protocol are contour propagation, dose accumulation, deformable image registration (DIR), quality assurance, response assessment, prescription for adaptive planning and decision for plan adaption.[26] The current study was focused on the impact of image registration and dose accumulation on dose reporting parameters within the frame of pencil-beam treatment setting.

Materials and methods

Patient geometry

The pair of opposite breathing phases were selected for a selected patient with identification code 4DCT1 from a publicly available database of 4D CT radiotherapy treatment planning images of oesophageal cancer patients free of pulmonary disease.[27] Following the convention of the dataset, a deep inspiration breath-hold (DIBH) phase is labelled as T00, and normal expiration phase as T50. The dataset contains an associated series of 300 anatomical landmarks identified in both breathing phases by an expert in thoracic radiology. The images and anatomic feature locations were transferred from a publicly available web portal.[28] The images have a voxel size of 0.97 mm in planar and 2.5 mm in the axial direction. As a disease model, a segmented gross tumour volume (GTV) of a publicly available cancer patient R005 from a Non-Small Cell Lung Cancer (NSCLC) Radiomics study was superimposed over a selected landmark in the T00 anatomy.[29,30] The landmark was selected to be in proximity to the equivalent position of the tumour in the R005 patient. Using Guy's deformable image registration, the tumour was projected to T50 anatomy. The planning volumes were delineated according to a recent photon- to proton radiotherapy comparison study.[31] A unity of GTV in both breathing phases was considered the internal tumour volume, while clinical tumour volume (CTV) was obtained by adding a uniform 8 mm margin. According to the study, a further 5 mm symmetrical margin was added to CTV to form the planning tumour volume (PTV). Treatment goals were for $D_{98\%}$ in PTV to be 100% of the prescribed dose with $D_{2\%}$ not to exceed 107% of the prescribed dose. No tissue was considered to be at risk for the purpose of planning. For dose reporting, the left lung with tumour volume removed was delineated as organ at risk (OAR).

Dose engine

The choice of particle tracking software was motivated by its use as a reference in neutron driven dose distribution in neutron reactor models and successful validations in proton therapy. At the

same time it was successfully used in radiation protection and dosimetry, radiation shielding, radiography, medical physics, nuclear criticality safety, detector design and analysis, nuclear oil well logging, accelerator target design and fission and fusion reactor design.[32,33] Backed by the nuclear energy application, Monte Carlo N-Particle, Version 6 (MCNP6) is subject to the most rigorous validation procedures.[34] The methodology of MCNP6 extended to proton and light nuclide cross-sections and transport properties in treatment scenarios would reduce existing uncertainties in dose calculations.

MCNP6 computational model geometry and material composition were created by using the mesh geometry feature in MCNP6. In this mode, each voxel is assigned its own material composition derived from isotopic material composition and associated material density. The voxelized geometry was derived directly from the selected CT images. A CT to material conversion was performed according to a lookup table (LUT), based on the interpolation in indicated ranges and material composition was derived from literature.[35] The errors on the simulation parameters, such as proton range, were found to be negligible for scanners where stoichiometric calibration was performed.[36,37] The default physics model of MCNP6, version 6.2 was used with a cut-off value of 10 keV for tracking of protons, photons and electrons, which, converted to range, allows particle tracking down to 1 μm. Neutron production and transport were modelled with the default physics list. A detailed study of the dose engine was performed with a water phantom where dose distribution and particle range equivalent to theory predicted values were observed. All calculations were performed at the Reactor research centre of the Jožef Stefan Institute on computers used and validated for simulating cores in nuclear power plants.

Treatment plan optimization

An in-house treatment plan optimization software was developed in the MATLAB software environment.[38] The treatment was assumed to be delivered by a set of pencil beams. The full treatment plan was formed by separating a sagittal irradiation field with a total size of 70 by 70 mm^2 to a grid of 14 by 14 independent beams with a spot diameter of 5 mm at tissue entry, as illustrated in Figure 1. Each beam was further divided to 25 energy bins spanning the range between 50 and 100 MeV, with the upper limit sufficient to deposit energy in the posterior of the PTV. Such an arrangement resulted

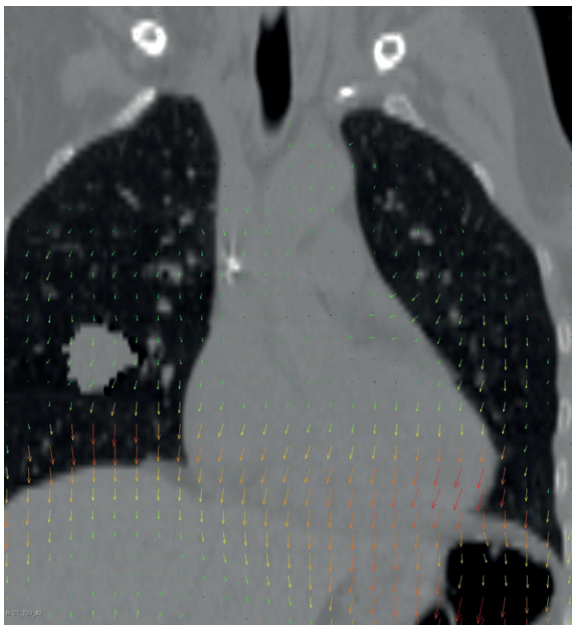

FIGURE 1. A coronal view of the deformable image transformation deformation field of moving exhale phase with respect to the reference deep inspiration breath-hold (DIBH) phase. Size and direction of the local motion is indicated by the arrows with the colour/whiteness of the arrow indicating the size of the translation.

in a set of 4900 beamlets with a fixed direction and energy.

The response of each beamlet was simulated individually using 10^5 protons per beamlet to generate the beamlet matrix A with rows corresponding to voxels and columns to each beamlet arrangement. The number of protons per beamlet was chosen to mimic recent studies in the field.[31] The column vector of 4900 beamlet weights **w** was then determined by requiring

$$\widehat{w} = \min\{\|w\|_1; A\boldsymbol{w} > \boldsymbol{b}, A\boldsymbol{w} < 1.07\boldsymbol{b}\}, \qquad [1]$$

where **b** corresponds to the dose plan in the space of the reference image, set to one within PTV and zero elsewhere, and ||w|| is the 1-norm (sum of the absolute values) of the weight vector. The second inequality in Eq. [1] was only applied to PTV. The optimization was performed as a linear programming optimization using MATLAB Version R2017B.[38]

The error due to limited number of protons per beamlet and in the full simulation was evaluated by generating two independent beam matrix realizations A_1 and A_2, associated optimized plans $\mathbf{w}_1$ and $\mathbf{w}_2$ and associated dose realizations $\mathbf{b}_1$ and $\mathbf{b}_2$. A dose ratio image was calculated as $R=D^{(2)}/D^{(1)}$,

where $D^{(1)}$ corresponds to $\mathbf{b}_1$ and $D^{(2)}$ to $\mathbf{b}_2$. As a measure of consistency, the average and standard deviation of R over voxels were evaluated.

The average relative dose difference ΔD was evaluated using N voxel locations within PTV labeled with index *i*, and two independent realizations, $D^{(1)}$ and $D^{(2)}$ as:

$$\Delta D = \frac{1}{N}\frac{\sum_{i=1}^{N}\left|D_i^{(1)}-D_i^{(2)}\right|}{D_i^{(1)}} \quad [2]$$

ΔD represents average statistical uncertainty in dose distribution on a voxel level due to a finite number of simulated particles.

Image registration

To accumulate dose distributions from different breathing frames to the reference frame, deformable image registration was performed using Elastix framework.[39] A fully continuous deformation mode was assumed with image warping function $\mathbf{u}(\mathbf{x}) = \mathbf{x}'-\mathbf{x}$ parameterized by B-splines, where $\mathbf{x}$ corresponds to a voxel location in the reference, DIBH phase and $\mathbf{x}'$ corresponds to its translation in the moving, exhale phase. Registration methods developed by Staring, Guy and Mattes were evaluated.

The Staring registration method was developed to measure tissue destruction in patients with pulmonary emphysema in clinical trials.[40] The method was used for intra-patient registration of images recorded at consecutive pulmonary function evaluation visits 29 months (median) apart, with inhale and exhale images recorded at both visits. Of the methods in the paper, the normalized correlation (NC) that seeks patch-wise matches between the fixed and the moving image, was chosen as it showed the best performance over 21 patients and 1849 annotated points.

The method of Guy *et al.* was targeted at recognizing large geometric changes in thorax primarily occurring due to treatment associated atelectasis in non-small cell lung cancer patients undergoing stereotactical radiotherapy treatment.[41] The challenge was to match the original patient image to the image with the most severe deformation. While a multitude of locally applied image metrics were used in the original paper, the method used in the present work was a combination of sum of squared tissue volume difference (SSTVD) and bending energy penalty, BE, an image preserving regularization term that penalizes non-physical transformations. The method selection was motivated by the most consistent results of the registration scheme in the absence of lobe label images.

The third registration alternative was derived for the purpose of the present study and is based on the Mattes advanced mutual information (MI) metrics.[42] The Mattes MI metric is a modality independent approach that can be used for monomodal or multi-modal image combinations. As in other approaches, adaptive stochastic gradient descent was used in MI maximization and multi-scale registration with six resolution levels was applied.

The methods were compared in terms of residual accuracy by evaluating the mean and standard deviation of residuals distribution $\mathbf{r}_X = \mathbf{x}'-(\mathbf{x}+\mathbf{u}(\mathbf{x}))$ on a set of fiducial markers for method X where X stands for S for Staring, G for Guy and M for Mattes registration method. To check for potential bias, a Pearson *r* correlation test of $\mathbf{r}_X$ to their inferior-superior coordinate in the lung was performed.

The Slicer 3D toolkit with the Elastix and RT extension was used to derive the registration parameters and perform the transformation of landmarks and calculated doses between breathing phases. The tool was validated to contribute negligible contributions to dose uncertainty.[43]

Dose evaluation

In continuation, labelling adheres to the convention that an argument of the quantity denotes the breathing phase the irradiation is delivered at, while the quantity's index refers to the coordinate system of the quantity representation. The reference DIBH dose distribution $D_{T00}(T00)$ was predicted for a static case scenario, where the patient was not assumed to move from the inspiration position T00, used in plan optimization. The off-geometry dose $D_{T50}(T50)$ was simulated to be the dose delivered to the optimization blinded exhale anatomy T50, recorded in T50 coordinate system. For dose accumulation, transformed dose $D_{T00/X}(T50)$ was derived from $D_{T50}(T50)$ using DIR method X.

The resulting dose-volume histograms (DVH) were determined for both the PTV and the organ at risk volume (OAR), the left lung which surrounds the tumor. Standard DVH measures ($D_{2\%}$, $D_{50\%}$, $D_{98\%}$, V_{20Gy}, V_{35Gy}) and associated homogeneity index (HI),

$$HI = \frac{D_{2\%} - D_{98\%}}{D_{50\%}}$$

were evaluated on produced DVH.[44] For V_{20Gy} and V_{35Gy}, a 70 Gy prescribed dose was assumed.

Dose comparison was evaluated using gamma analysis.[45] To match the voxelization of the dose plan, passing criteria was evaluated for 3mm/3% distance to agreement/dose difference threshold

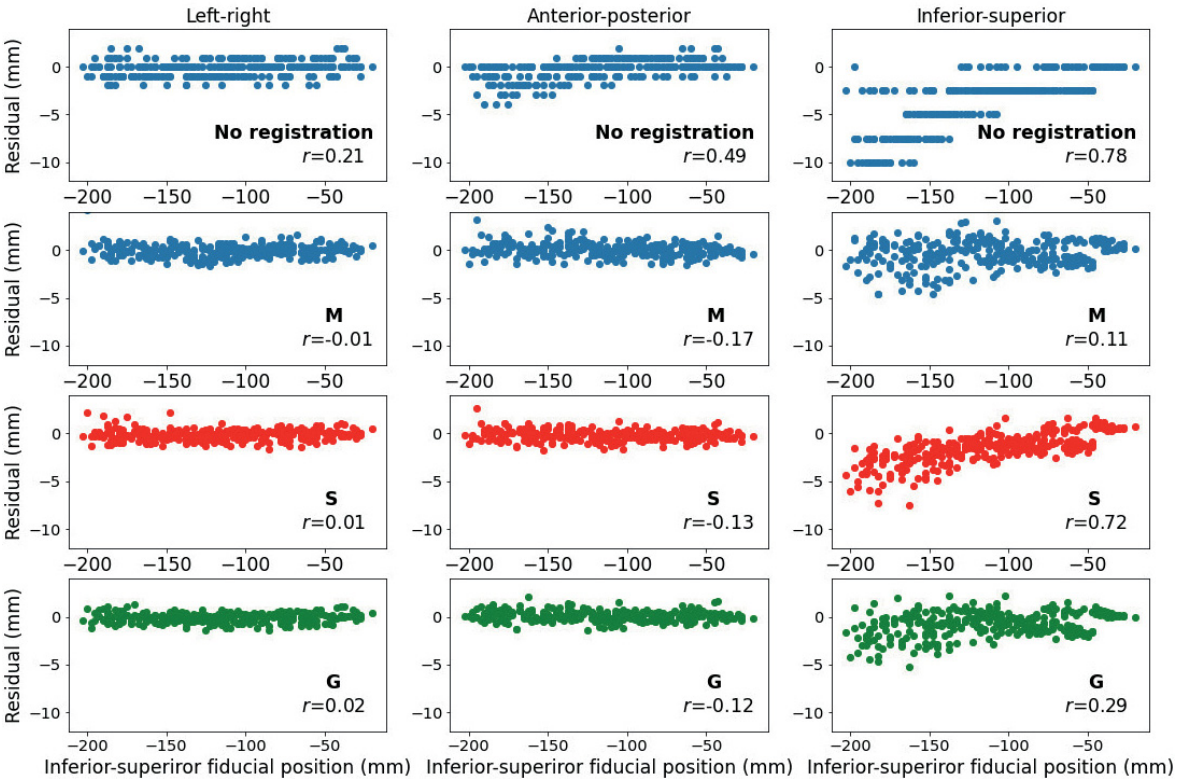

FIGURE 2. Scatter plots of the annotated landmarks with respect to the residual shift between expiration and inspiration positions on y-axis and location of landmark along the inferior-superior direction on the x-axis. A total of twelve panes indicate residual extent in the left-right direction (first column), anterior-posterior direction (second column) and inferior-superior direction (third column). The top row shows differences in landmark locations without registration, following rows correspond to (S) Staring deformable image registration (DIR), (G) Guy DIR and (M) Mattes DIR algorithms. Indicated numbers correspond to Pearson's *r* coefficient calculated for the shown distribution.

and 3 mm/1% threshold. To evaluate statistical uncertainty associated with dose engine, two independent realizations of identical plans were compared both for simulations of reference and off-reference geometries. To study agreement between image registration methods, pair-wise analysis was performed, always using the same T00 dose distribution as a reference but comparing it to off-reference T50 dose distribution projected to T00 using different registration methods. A comparison was performed for two realizations to estimate the uncertainty of gamma analysis results. In all cases, a passing rate where the gamma function is lower than unity is reported, and images of the gamma function were drawn for statistical and cross-method comparisons.

Results

Figure 1 shows the coronal projection of the deformation field derived from the deformable image registration of the selected pair of inspiration-expiration images, using the Mattes method. The Figure shows that the dominant difference in geometry was in the inferior-superior direction z, due to the vertical motion of the diaphragm. The average shift was 3.6 mm with a deviation of 2.1 mm and a maximum range of 10 mm (4 slices) upwards and - 2.5 mm (1 slice) downwards. In the axial plane, the movement was less pronounced, giving an average shift of 0.2 mm with a deviation of 0.8 mm and 1.0 mm in left-right, x-direction and anterior-posterior, y-direction, respectively. The best alignment was achieved for the Mattes mutual information, where the residual shift in z-direction had an average of 0.41 mm with a deviation of 1.3 mm, while means of - 0.06 mm and 0.1 mm with a deviation of 0.6 mm and 0.7 mm were measured in x and y direction, respectively. For the Guy method, a slight, but statistically significantly worse alignment with a mean residual shift of 0.76 mm and a deviation of 1.3 mm in z-direction was achieved, whereas in-plane xy residual shift

was comparable to the Mattes method. The worse performance was recorded for the Staring method, where a mean residual z coordinate shift of 1.55 mm was recorded with a deviation of 1.6 mm. The un-registered exhale vertical shift correlated with the z location of fiducials with a correlation coefficient r = 0.78. This correlation was completely removed for Mattes and Guy registration but persisted in Staring method with r = 0.72 on z residuals versus z-axis fiducial location. A slight correlation, r = 0.5, between the shift in the anterior-posterior direction and the vertical position is also present in the fiducial locations but is effectively removed by all registration methods. The residuals and their correlations are shown in Figure 2.

The dose distributions for an optimized irradiation plan are shown in Figure 3, where the left pane illustrates the dose to the optimized, DIBH geometry while the right pane shows the dose for the same optimized dose plan used on the unaccounted exhale phase CT. Both images are shown in the laboratory system, therefore the direction of the beams remains unchanged while the patient shifts between the left and the right pane. For two independent runs of dose distribution simulation, the average dose consistency ratio R had a mean of 1.000, and a standard deviation of 0.022 and an average dose discrepancy of $\Delta D = 0.04\%$ on the voxel level, and the uncertainty of the dose integral over the PTV was significantly smaller than 0.02%.

The registration step was performed to transform dose distribution from T50 to the reference, T00 geometry, and the results are shown in Figure 4. The Figure shows both the dose delivered to the reference, T00 geometry as well as off-reference, T50 geometry. For the $D_{T00/G}(T50)$ dose in the bottom panes, Guy's DIR transformation was used. Lung mask is drawn as an anatomical reference, with red/dark arrows indicating areas of under-irradiation and blue/light arrows area of over-irradiation for the $D_{T00/G}(T50)$ dose distribution.

Figure 5 shows the comparison of cumulative DVH for different scenarios: $D_{T00}(T00)$ - irradiation at DIBH only, assuming cooperative patient with superb body control, $D_{T00/X}(T50)$, irradiation at off-geometry conditions assuming DIBH but encountering patient in exhale phase, and registrations of T50 dose back to the reference system using either Staring, Mattes or Guy DIR method. A partial undertreatment of tumour is seen. This is reflected in quantitative DVH parameters, collected in Table 1. Little or no difference is seen for $D_{2\%}$ and $D_{50\%}$ relative doses, while moderate changes were recorded for near-minimum, $D_{98\%}$ dose and homogeneity index HI. While moderate, statistically significant difference in both DVH and DVH associated parameters is seen for doses registered with different

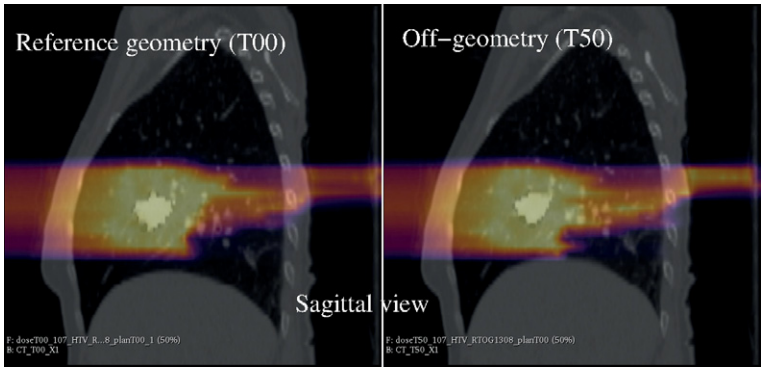

FIGURE 3. A sagittal view of the pair of the 4D CT image corresponding to the extreme breathing frames, left: deep inspiration breath-hold (DIBH) reference geometry, T00, right: exhale off-reference geometry, T50. A tumour delineated in Non-Small Cell Lung Cancer (NSCLC) Radiomics study, case R005, was artificially added into the lung. Dose delivered according to beamlet set that was optimized for irradiation in the reference frame is superimposed with colour scale/grey scale corresponding to the accumulated dose where bright colours/white corresponds to a higher accumulated dose. A slight relative upward motion of tumour in T50 geometry causes the right pane dose distribution to deviate from the planned distribution illustrated in the left pane.

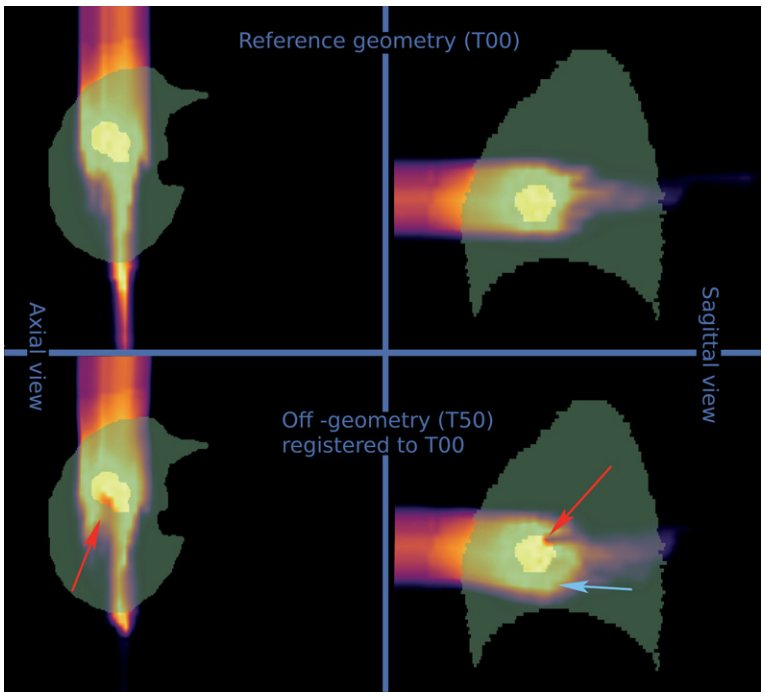

FIGURE 4. Colormap/grayscale image of dose distributions superimposed on lung mask for two views (axial and sagittal) and two irradiation conditions - irradiation geometry identical to planned geometry, T00, top panes and irradiation geometry non equal to the planned geometry, with dose registered back to the reference geometry T50, below. Red/dark arrows indicate areas of under-treatment for the off-geometry case. Blue/light arrows indicate areas of over-treatment in organs at risk.

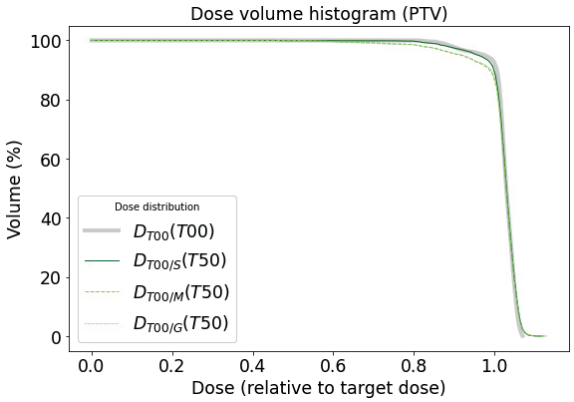

FIGURE 5. Dose-volume histogram (DVH) for planning tumour volume (PTV) under different irradiation conditions and volume representations. Comparison of D_{T00}(T00), thick solid line DVH for irradiation geometry identical to planned geometry, and $D_{T00/S}$(T50) $D_{T00/M}$(T50) and $D_{T00/G}$(T50), solid, dashed and dotted lines for DVH in off-geometry but evaluated in the reference geometry using Staring (S), Mattes (M) or Guy (G) deformable image registration (DIR) method, respectively. Curves for Mattes and Guy registration methods nearly overlap.

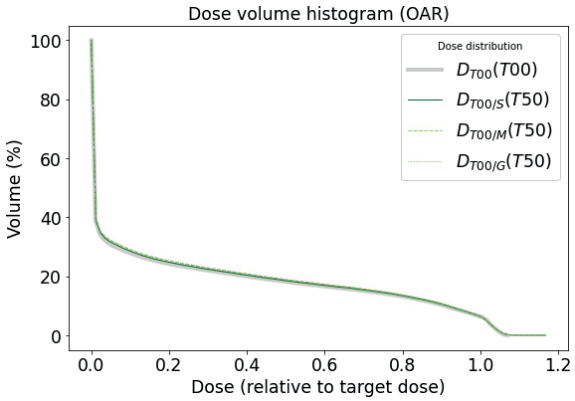

FIGURE 6. Dose-volume histogram (DVH) for planned organ at risk volume of organ at risk (OAR), left lung, under different irradiation conditions and volume representations. Comparison of D_{T00}(T00), thick solid line DVH for irradiation geometry identical to planned geometry, and $D_{T00/S}$(T50), $D_{T00/M}$(T50) and $D_{T00/G}$(T50), solid, dashed and dotted lines for DVH in off-geometry but evaluated in the reference geometry using Staring (S), Mattes (M) or Guy (G) deformable image registration (DIR) method, respectively.

registration methods, reaching 1% for CTV and 8% for PTV, changing HI by a factor of 1.6 for PTV (0.39 over 0.24, see Table 2).

The effects on the organ at risk are shown in Figure 6. The difference between DVH for OAR is at the limit of detection. Quantitative DVH parameters are collected in Table 2, where significant DIR differences of approximately 0.6% of the lung volume can be seen for the V_{20Gy} and V_{35Gy} volumes. The reported values are lower than for OAR encountered in irradiation of other disease sites such as rectum or bladder in prostate cancer, confirming exclusion of OAR in treatment goals.[22]

Passing rates for gamma analysis of selected combinations of registration methods and independent dose realizations are shown in Table 3. In the first and the last row, simulations in identical anatomies are compared, indicating between 90 and 100% passing rate for dose differences driven by statistical uncertainty. Passing rates are higher for clinically relevant lower tolerance criteria of 3 mm/20%, reaching perfect acceptance in PTV and 97–98% for lung. The more restrictive threshold of 3 mm/3% was chosen to expose dose differences due to breathing, amounting to passing rates as low as 67% over lung volume and 77% over PTV. The drop in passing rate for comparison of dose in the reference and off-reference system is not unexpected. There is, however a significant variation associated with selection of a registration method, on the order of few percent for 3 mm/20% comparison (passing values between 98 and 99% in PTV, 87 and 91% in lung) reaching as much as 10% for a more sensitive, 3 mm/3% gamma evaluation (77–85% in PTV, 67–79% in lung). The uncertainty associated with repeated dose simulation was on the order of 1%. Similar conclusion can be qualitatively drawn from Figure 7, where the top left and bottom right panel show gamma function evaluated over lung volume shown in sagittal projection indicating minor differences in dose for successive plan application. Larger differences in dose distributions are associated with comparing doses from reference and off-reference anatomy, where registration method was used to project the dose to the reference system, top right, or comparing doses registered to T00 using different registration methods, bottom left.

Discussion

The presented study is a comparison of three successful registration methods of CT images of lungs, including recently reported Mattes and Guy's registration methods, and their impact on quantitative measures such as DVH associated parameters identified in ICRU ($D_{2\%}$, $D_{50\%}$, $D_{98\%}$) and gamma analysis passing rates. Compared to similar studies in the field, the novelty lies in the simultaneous evaluation of landmark associated errors and uncertainties of the dose associated quantitative measures.[46]

There is a significant controversy associated with propagating the dose to a common, accumulation coordinate system using DIR methods.[47] This was reflected in the design of the study where the success of registration was measured independently prior to dose transfer and multiple methods using alternative registration approaches were evaluated to get a sense of uncertainty associated with the approach. While the selection of DIR methods was neither exhaustive nor complete, it should serve to illustrate the controversy and provide quantitative measures of uncertainty associated with using DIR as a dose accumulation method.

The study was designed to illustrate uncertainty associated with DIR methods in propagating the delivered dose to a common dose accumulation reference system. To achieve the goal, plan optimization software was blinded to the off-reference geometry. Such a setup is close to a clinical scenario where the instant patient anatomy is recorded, followed by a simulation of the previously optimized plan and checking for considerable deviations.[48,49] The re-optimization decision in such cases is based on gamma analysis or similar measures, which were shown in the presented study to depend on the registration method used.

The study has several limitations - the workflow was tested on a single case for illustrative purposes, containing only extreme breathing phases to indicate the span of the algorithm, limited relevance of treatment plans which were derived by in-house software rather than a commercial engine, smaller motion uncertainties not related to breathing were ignored and coarse voxelization of dose histograms driven by limited simulation statistics.

The Monte Carlo simulation was run roughly a day on standard computer hardware. The relatively large computing burden is due to the voxelized geometry and associated burden of particle parameter recalculation at voxel boundaries. In the presented work, the simulation time was shortened by using coarse voxelization in MCNP6 geometry definition. The other contributing factor is the requirement for the statistical component of uncertainty be small compared to the changes in dose distribution due to motion. The size of statistical error was checked by observing dose differences ΔD, variation in $D_{2\%}$, $D_{50\%}$, $D_{98\%}$, V_{20Gy}, V_{35Gy} and gamma passing rates. In all cases, DIR driven uncertainty was more pronounced than statistical uncertainty, justifying the computing strategy. Opportunities for improvement lie in software interface optimization, particularly in material definition and optimized simulation strategies.[50,51]

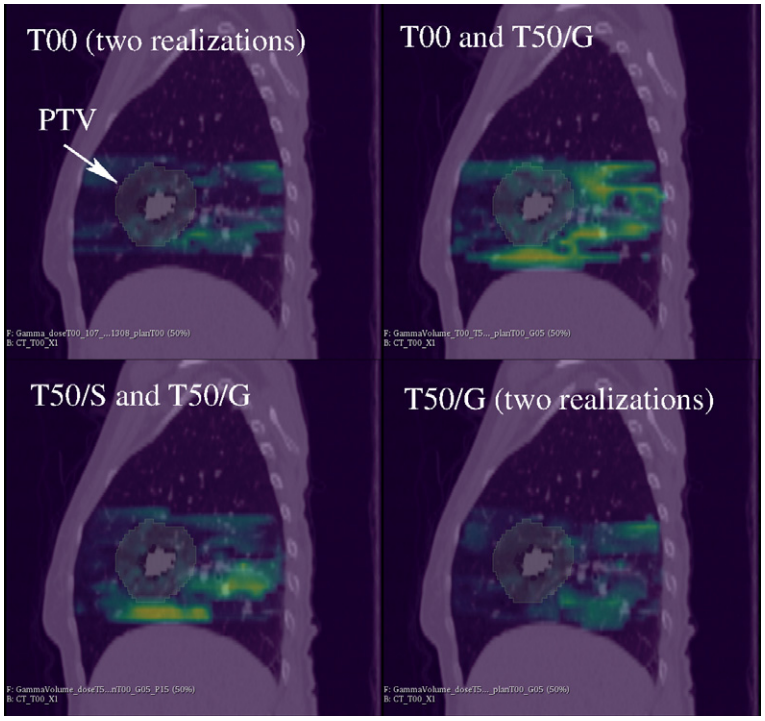

FIGURE 7. Gamma analysis for 3 mm/3% distance to agreement / dose difference tolerance criteria. All panes show sagittal projection of the CT image overlaid with planning tumour volume (PTV), indicated by arrow, and superimposed gamma function, where yellow/bright colour corresponds to a gamma value of approximately 2. Top left and bottom right pane show dose comparison of statistically independent realizations of the same dose plan, top left for reference geometry, bottom-right for off-reference geometry registered with Guy's deformable image registration (DIR). Top right is a comparison of T50 dose registered to T00 using Guy's method and T00 dose. Bottom left is a comparison of T50 dose registered to T00 using two registration methods, S for Staring and G for Guy.

TABLE 1. Dose-volume histogram parameters for planning tumour volume (PTV) and (CTV) and dose distributions delivered in reference, T00, geometry corresponding to reference deep inspiration breath-hold (DIBH) and off-reference T50, geometry, corresponding to exhale phase. Subscripts indicate coordinate system where the doses were evaluated with the letter following the slash indicating the deformable image registration (DIR) method used: S for Staring, M for Mattes and G for Guy. Relative doses are reported. The number of decimal places indicates statistical accuracy evaluated by repeated simulation

Dose	PTV				CTV			
distribution	$D_{2\%}$ (%)	$D_{50\%}$ (%)	$D_{98\%}$ (%)	HI	$D_{2\%}$ (%)	$D_{50\%}$(%)	$D_{98\%}$ (%)	HI
D_{T00}(T00)	106.4	102.9	89.5	0.2	106.7	103.4	100.8	0.05
$D_{T00/S}$(T50)	107.1	103.0	87.2	0.24	107.3	103.4	98.2	0.09
$D_{T00/M}$(T50)	107.1	102.9	80	0.39	107.4	103.4	97	0.10
$D_{T00/G}$(T50)	107.2	102.9	79	0.38	107.4	103.4	96	0.11

HI = homogeneity index

The concept of PTV is of limited relevance in particle therapy as the field is shifting towards robust optimization.[52] In the absence of a robust opti-

TABLE 2. Dose-volume histograms (DVH) parameters for OAR (left lung) and dose distributions delivered in deep inspiration breath-hold (DIBH), reference, T00 geometry and exhale, off-reference, T50 geometry. Subscripts indicate coordinate system where the dose was evaluated, with the letter following the slash indicating the deformable image registration (DIR) method used: S for Staring, M for Mattes and G for Guy. V_{20Gy} and V_{35Gy} are derived for a prescribed dose of 70 Gy. The uncertainty σ_D is the standard deviation of V_D for repeated simulation

Dose distribution	V_{20Gy} (%)	σ_{20Gy} (%)	V_{35Gy} (%)	σ_{35Gy} (%)
D_{T00}(T00)	26.2	0.2	18.3	0.1
DT00/S	26.8	0.1	18.7	0.1
$D_{T00/G}$(T50)	27.4	0.1	19.1	0.1
$D_{T00/M}$(T50)	27.3	0.1	19.0	0.1

TABLE 3. Gamma analysis pass rate for comparing the dose distributions in reference coordinate system achieved under different irradiation conditions and with different registration methods. For identical anatomies, the gamma passing rate between doses achieved for subsequent realizations was performed. For mismatching irradiation anatomies, the average pass rate over multiple realizations is reported. The number of decimal places indicates statistical uncertainty tested by repeated simulation

Dose pairs	Pass rate (%)			
	PTV		Lung	
	3 mm/20 %	3 mm/3 %	3 mm/20 %	3 mm/3 %
T00 with T00	100	97	97	90
T50/S with T00	99	85	91	79
T50/M with T00	98	77	87	67
T50/G with T00	98	83	89	77
T50/G with T50/G	100	96	98	92

PTV = planning tumour volume

mizer, PTV equivalent to volumetric modulated arc therapy (VMAT) approach was used in the present study.[31] The 5 mm margin was sufficient to limit dose degradation in optimization blinded T50 to $D_{98\%}$ dropping to 96–99%, range showing results for different DIR. The DIR uncertainty represents an additional source of uncertainty and should be included in robust optimization over multiple patient anatomies. A similar problem is encountered in gantry-less proton irradiation facilities, which consider patients in other than lying positions and motion must be included as part of the optimization.[53]

Dose accumulation is of particular interest for the adaptive treatment protocols. In principle, instantaneous changes in patient geometry could be compensated by beam delivery, however several technical challenges remain. Optimization of the plan should be performed for each included geometry phase, which might yield impractical plan preparation times. A rigorous synchronization between the delivery system and monitor of patient motion must be implemented. The treatment machine must be able to react to changes in prescribed operation in synchronization with geometry changes.[54] As a surrogate, plans adapted to daily patient geometry were chosen by the RAPTOR initiative. These would depend critically on the accurate dose accumulation presented in the present study.[25]

The standard deviation of the residual errors of half the slice width is equivalent to approximately two-thirds of the markers aligned to within the original voxel location in whichever of the DIR scheme. Nevertheless, the DIR driven discrepancy on $D_{98\%}$ in the PTV was on the order of 10 %, and this for a tumor inserted at an area, where dislocation between phases was moderate. This is comparable to reported differences in therapeutic volume coverage due to patient motion and to dose uncertainties reported in similar dose accumulation studies.[49,55] This uncertainty remains one of the most severe challenges of dose accumulation and its use to quantify treatment and correlate it to clinical outcomes in future studies of proton therapy benefits in cancer treatment.

Conclusions

A quantitative method to evaluate the impact of patient motion on treatment reporting parameters was developed based on strictly validated MCNP6 nuclear radiation simulation tool. A set of DIR methods were tried and shown to align images with maximum differences driven by breathing to within voxel sizes. Nevertheless, moderate but statistically significant DIR driven differences were reported for some common dose evaluation parameters such as 10% difference in $D_{98\%}$ of PTV and 10% difference in gamma analysis passing rates. A caution must be exercised in using the dose accumulation and associated measures in adaptive therapy optimization algorithms and studies like the presented can help in the determination of the limits of agreement among different registration algorithms.

Acknowledgment

The authors acknowledge the financial support from the Slovenian Research Agency (research

core funding No. P1-0389 and P2-0073. We would like to thank the reviewers for excellent comments and providing compelling reference material for discussion.

References

1. Pedroni E, Bacher R, Blattmann H, Bohrinaer T, Coray A, Lomax A, et al. The 200-MeV proton therapy project at the Paul Scherrer Institute: conceptual design and practical realization. *Med Phys* 1995; **22:** 37-53. doi: 10.1118/1.597522

2. Degiovanni A, Amaldi U. History of hadron therapy accelerators. *Phys Medica* 2015; **31:** 322-32. doi: 10.1016/j.ejmp.2015.03.002.

3. Terasawa T, Dvorak T, Ip S, Raman G, Lau J, Trikalinos TA. Systematic review: charged-particle radiation therapy for cancer. *Ann Intern Med* 2009; **151:** 556-65. doi: 10.7326/0003-4819-151-8-200910200-00145

4. Verma V, Rwigema J-CM, Malyapa RS, Regine WF, Simone CB. Systematic assessment of clinical outcomes and toxicities of proton radiotherapy for reirradiation. *Radiother Oncol* 2017; **125:** 21-30. doi: 10.1016/j.radonc.2017.08.005

5. Liao Z, Lee JJ, Komaki R, Gomez DR, O'Reilly MS, Fossella FV, et al. Bayesian adaptive randomization trial of passive scattering proton therapy and intensity-modulated photon radiotherapy for locally advanced non–small-cell lung cancer. *J Clin Oncol* 2018; **36:** 1813-22. doi: 10.1200/JCO.2017.74.0720

6. Jones B. Towards achieving the full clinical potential of proton therapy by inclusion of LET and RBE models. *Cancers* 2015; **7:** 460-80. doi: 10.3390/cancers7010460

7. Hu M, Jiang L, Cui X, Zhang J, Yu J. Proton beam therapy for cancer in the era of precision medicine. *J Hematol Oncol* 2018; **11:** 136. doi: 10.1186/s13045-018-0683-4

8. Kissick MW, Boswell SA, Jeraj R, Mackie TR. Confirmation, refinement, and extension of a study in intrafraction motion interplay with sliding jaw motion. *Med Phys* 2005; **32:** 2346-50. doi: 10.1118/1.1935774

9. Yu CX, Jaffray DA, Wong JW. The effects of intra-fraction organ motion on the delivery of dynamic intensity modulation. *Phys Med Biol* 1998; **43:** 91-104. doi: 10.1088/0031-9155/43/1/006

10. Engelsman M, Schwarz M, Dong L. Physics controversies in proton therapy. *Semin Radiat Oncol* 2013; **23:** 88-96. doi: 10.1016/j.semradonc.2012.11.003

11. Sonke JJ, Zijp L, Remeijer P, Van Herk M. Respiratory correlated cone beam CT. *Med Phys* 2005; **32:** 1176-86. doi: 10.1118/1.1869074

12. Widesott L, Amichetti M, Schwarz M. Proton therapy in lung cancer: clinical outcomes and technical issues. A systematic review. *Radiother Oncol* 2008; **86:** 154-64. doi: 10.1016/j.radonc.2008.01.003

13. De Ruysscher D, Sterpin E, Haustermans K, Depuydt T. Tumour movement in proton therapy: solutions and remaining questions: a review. *Cancers* 2015; **7:** 1143-53. doi: 10.3390/cancers7030829

14. Moteabbed M, Schuemann J, Paganetti H. Dosimetric feasibility of real-time MRI-guided proton therapy. *Med Phys* 2014; **41:** 111713. doi: 10.1118/1.4897570

15. Pollard JM, Wen Z, Sadagopan R, Wang J, Ibbott GS. The future of image-guided radiotherapy will be MR guided. *Br J Radiol* 2017; **90:** 20160667. doi: 10.1259/bjr.20160667

16. Padilla-Cabal F, Georg D, Fuchs H. A pencil beam algorithm for magnetic resonance image-guided proton therapy. *Med Phys* 2018; **45:** 2195-204. doi: 10.1002/mp.12854

17. Ding GX, Alaei P, Curran B, Flynn R, Gossman M, Mackie TR, et al. Image guidance doses delivered during radiotherapy: quantification, management, and reduction: report of the AAPM Therapy Physics Committee Task Group 180. *Med Phys* 2018; **45:** e84-99. doi: 10.1002/mp.12824

18. Vedam SS, Keall PJ, Kini VR, Mostafavi H, Shukla HP, Mohan R. Acquiring a four-dimensional computed tomography dataset using an external respiratory signal. *Phys Med Biol* 2003; **48:** 45-62. doi: 10.1088/0031-9155/48/1/304

19. Pan T, Lee T-Y, Rietzel E, Chen GTY. 4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT. *Med Phys* 2004; **31:** 333-40. doi: 10.1118/1.1639993

20. Malicki J. The importance of accurate treatment planning, delivery, and dose verification. *Reports Pract Oncol Radiother* 2012; **17:** 63-5. doi: 10.1016/j.rpor.2012.02.001

21. Gregoire V, MacKie TR. Dose prescription, reporting and recording in intensity-modulated radiation therapy: a digest of the ICRU Report 83. *Imaging Med* 2011; **3:** 367-73. doi: 10.2217/IIM.11.22

22. Jones D, Suit H, Kanematsu N, Tatsuzaki H, Tsujii H. Recording, and reporting proton-beam therapy ICRU Report 78. [Internet]. *J ICRU* 2007; **7:** 1-210. [cited 2021 Mar 15]. Available at : https://www.icru.org/report/prescribing-recording-and-reporting-proton-beam-therapy-icru-report-78/

23. Yan D, Vicini F, Wong J, Martinez A. Adaptive radiation therapy. *Phys Med Biol* 1997; **42:** 123-32. doi: 10.1088/0031-9155/42/1/008

24. Dolde K, Naumann P, David C, Gnirs R, Kachelrieß M, Lomax AJ, et al. 4D dose calculation for pencil beam scanning proton therapy of pancreatic cancer using repeated 4DMRI datasets. *Phys Med Biol* 2018; **63:** 165005. doi: 10.1088/1361-6560/aad43f

25. European Commission. CORDIS EU research results. Real-time Adaptive Particle Therapy of Cancer. RAPTOR [Internet]. [cited 2021 Mar 16]. Available at: https://cordis.europa.eu/project/id/955956

26. Zhong H, Jin J-Y. Recent advances and challenges in adaptive radiotherapy for patients with locally advanced NSCLC. *Ann Radiat Ther Oncol* 2017; **1:** 1008. doi: 10.25107/2577-8757/arto-v1-id1008

27. Castillo R, Castillo E, Guerra R, Johnson VE, McPhail T, Garg AK, et al. A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. *Phys Med Biol* 2009; **54:** 1849-70. doi: 10.1088/0031-9155/54/7/001

28. Castillo R. The deformable image registration laboratory. [Internet]. [cited 2021 Mar 17]. Available at: http://www.dir-lab.com/

29. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014; **5:** 4006. doi: 10.1038/ncomms5006

30. Nationa Cancer Institute. Cancer Imaging program. The cancer imaging archive [Internet]. [cited 2021 Mar 18]. Available at: https://www.cancerimagingarchive.net/

31. Teoh S, Fiorini F, George B, Vallis KA, Van den Heuvel F. Proton vs photon: a model-based approach to patient selection for reduction of cardiac toxicity in locally advanced lung cancer. *Radiother Oncol* 2020; **152:** 151-62. doi: 10.1016/j.radonc.2019.06.032

32. Mashnik SG. Stepan G. Validation and verification of MCNP6 as a new simulation tool useful for medical applications. [Internet]. *44th Annu Midyear Meet Heal Phys Soc* 2011, *Charleston, SC (United States);* 6 Jan 2011; 24 p; Report No. LA-UR-11-00083. Avalable at: https://inis.iaea.org/search/search.aspx?orig_q=RN:43119331

33. Ardenfors O, Dasu A, Kopeć M, Gudowska I. Modelling of a proton spot scanning system using MCNP6. *J Phys Conf Ser* 2017; **860:** 012025. doi: 10.1088/1742-6596/860/1/012025.

34. Goorley T, James M, Booth T, Brown F, Bull J, Cox LJ, et al. Features of MCNP6. *Ann Nucl Energy* 2016; **87:** 772-83. doi: 10.1016/j.anucene.2015.02.020

35. Schneider W, Bortfeld T, Schlegl W. Correlation between CT numbers and tissue parameters needed for Monte Carlo simulation of clinical dose distributions. *Phys Med Biol* 2000; **45:** 459-78. doi: 10.1088/0031-9155/45/2/314

36. Schneider U, Pedroni E, Lomax A. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Phys Med Biol* 1996; **41:** 111-24. doi: 10.1088/0031-9155/41/1/009

37. Paganetti H. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Phys Med Biol* 2012; **57:** R99-117. doi: 10.1088/0031-9155/57/11/R99

38. The Mathworks, Inc. MATLAB. version 9.3.0.713579 (R2017b). 2017. Natick, Massachusetts; 2017.

39. Klein S, Staring M, Murphy K, Viergever MA, Pluim JPW. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging* 2010; **29:** 196-205. doi: 10.1109/TMI.2009.2035616

40. Staring M, Bakker ME, Stolk J, Shamonin DP, Reiber JH, Stoel BC. Towards local progression estimation of pulmonary emphysema using CT. *Med Phys* 2014; **41:** 021905. doi: 10.1118/1.4851535

41. Guy CL, Weiss E, Christensen GE, Jan N, Hugo GD. CALIPER: a deformable image registration algorithm for large geometric changes during radiotherapy for locally advanced non-small cell lung cancer. *Med Phys* 2018; **45:** 2498-508. doi: 10.1002/mp.12891

42. Mattes D, Haynor DR, Vesselle H, Lewellyn TK, Eubank W. Nonrigid multimodality image registration. *Proc SPIE Med Imaging* 2001; **4322:** 1609-20. doi: 10.1117/12.431046

43. Pinter C, Lasso A, Wang A, Jaffray D, Fichtinger G. SlicerRT. Radiation therapy research toolkit for 3D Slicer. *Med Phys* 2012; **39:** 6332-8. doi: 10.1118/1.4754659

44. Gregoire V, Mackie TR, De Neve W, Gospodarowicz M, van Herk M, Niemierko A. Prescribing, recording, and reporting intensity-modulated photon-beam therapy (IMRT) ICRU Report 83. *J ICRU* 2010; **10:** 1-35. doi: 10.1093/jicru/ndq001

45. Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys* 1998; **25:** 656-61. doi: 10.1118/1.598248

46. Amstutz F, Nenoff L, Albertini F, Ribeiro CO, Knopf AC, Unkelbach J, et al. An approach for estimating dosimetric uncertainties in deformable dose accumulation in pencil beam scanning proton therapy for lung cancer. *Phys Med Biol* 2021; **66:** 105007. doi: 10.1088/1361-6560/abf8f5

47. Schultheiss TE, Tome WA, Orton CG. Point/counterpoint: it is not appropriate to "deform" dose along with deformable image registration in adaptive radiotherapy. *Med Phys* 2012; **39:** 6531-3. doi: 10.1118/1.4722968

48. Schaly B, Kempe J, Venkatesan V, Mitchell S, Battista JJ. Using gamma index to flag changes in anatomy during image-guided radiation therapy of head and neck cancer. *J Appl Clin Med Phys* 2017; **18:** 79-87. doi: 10.1002/acm2.12180

49. Houweling AC, Crama K, Visser J, Fukata K, Rasch CRN, Ohno T, et al. Comparing the dosimetric impact of interfractional anatomical changes in photon, proton and carbon ion radiotherapy for pancreatic cancer patients. *Phys Med Biol* 2017; **62:** 3051-64. doi: 10.1088/1361-6560/aa6419

50. Rehfeld NS, Stute S, Apostolakis J, Soret M, Buvat I. Introducing improved voxel navigation and fictitious interaction tracking in GATE for enhanced efficiency. *Phys Med Biol* 2009; **54:** 2163-78. doi: 10.1088/0031-9155/54/7/021

51. Yuan J, Chen Q, Brindle J, Zheng Y, Lo S, Sohn J, et al. Investigation of nonuniform dose voxel geometry in Monte Carlo calculations. *Technol Cancer Res Treat* 2015; **14:** 419-27. doi: 10.1177/1533034614547459

52. Liu W, Zhang X, Li Y, Mohan R. Robust optimization of intensity modulated proton therapy. *Med Phys* 2012; **39:** 1079-91. doi: 10.1118/1.3679340

53. Yan S, Depauw N, Flanz J, Adams J, Gorissen BL, Shih H, et al. SU-F-T-207: does the greater flexibility of pencil beam scanning reduce the need for a proton gantry? *Med Phys* 2016; **43:** 3509-10. doi:10.1118/1.4956345

54. Graeff C, Lüchtenborg R, Eley JG, Durante M, Bert C. A 4D-optimization concept for scanned ion beam therapy. *Radiother Oncol* 2013; **109:** 419-24. doi: 10.1016/j.radonc.2013.09.018

55. Nenoff L, Ribeiro CO, Matter M, Hafner L, Josipovic M, Langendijk JA, et al. Deformable image registration uncertainty for inter-fractional dose accumulation of lung cancer proton therapy. *Radiother Oncol* 2020; **147:** 178-85. doi: 10.1016/j.radonc.2020.04.046