ACTA ONCOLOGICA, 2017
VOL. 56, NO. 4, 531–540
http://dx.doi.org/10.1080/0284186X.2017.1293287

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

A study of the beam-specific interplay effect in proton pencil beam scanning delivery in lung cancer

Minglei Kang[a]*, Sheng Huang[a]*, Timothy D. Solberg[a,b], Rulon Mayer[c], Andy Thomas[c], Boon-Keng Kevin Teo[a], James E. McDonough[a], Charles B. Simone II[a,d] and Liyong Lin[a]

[a]Department of Radiation Oncology, University of Pennsylvania, Philadelphia, PA, USA; [b]Department of Radiation Oncology, University of California, San Francisco, CA, USA; [c]Department of Radiation Oncology, Walter Reed Military Hospital, Bethesda, MD, USA; [d]Department of Radiation Oncology, University of Maryland Medical Center, Baltimore, MD, USA

ABSTRACT

Background: For lung tumors with large motion amplitudes, the use of proton pencil beam scanning (PBS) can produce large dose errors. In this study, we assess under what circumstances PBS can be used to treat lung cancer patients who exhibit large tumor motion, based on the quantification of tumor motion and the dose interplay.

Material and methods: PBS plans were optimized on average 4DCT datasets using a beam-specific PTV method for 10 consecutive patients with locally advanced non-small-cell-lung-cancer (NSCLC) treated with proton therapy to 6660/180 cGy. End inhalation (CT0) and end exhalation (CT50) were selected as the two extreme scenarios to acquire the relative stopping power ratio difference (Δrsp) for a respiration cycle. The water equivalent difference (ΔWET) per radiological path was calculated from the surface of patient to the iCTV by integrating the Δrsp of each voxel. The magnitude of motion of voxels within the target follows a quasi-Gaussian distribution. A motion index (MI (>5mm WET)), defined as the percentage of target voxels with an absolute integral ΔWET larger than 5 mm, was adopted as a metric to characterize interplay. To simulate the treatment process, 4D dose was calculated by accumulating the spot dose on the corresponding respiration phase to the reference phase CT50 by deformable image registration based on spot timing and patient breathing phase.

Results: The study indicated that the magnitude of target underdose in a single fraction plan is proportional to the MI ($p < .001$), with larger motion equating to greater dose degradation and standard deviations. The target homogeneity, minimum, maximum and mean dose in the 4D dose accumulations of 37 fractions varied as a function of MI.

Conclusions: This study demonstrated that MI can predict the level of dose degradation, which potentially serves as a clinical decision tool to assess whether lung cancer patients are potentially suitable to receive PBS treatment.

ARTICLE HISTORY
Received 11 August 2016
Accepted 4 February 2017

Introduction

The most advanced proton therapy delivery technique currently available is pencil beam scanning (PBS) [1,2], in which spot intensities and energies can be optimized using cost functions within an inverse treatment planning process, accounting for both target coverage and constraints on normal tissues. Compared to double scattering (DS) techniques, the scanning beam can deliver conformal dose to both proximal and distal regions of the target without using patient-specific compensators [3]. Several previous investigations [4–9] have demonstrated that PBS treatment planning for thoracic tumors might reduce the dose to critical normal tissues relative to proton DS or intensity-modulated photon radiation therapy (IMRT) techniques. In spite of the potential dosimetric advantages of proton PBS, however, DS and IMRT remain the modalities of choice for treating complex thoracic tumors. This is due to two major factors: (1) the inability of commercial TPS to predict potential range variations due to respiratory motion, setup errors and inaccurate stopping power modeling using CT [8–11]; and (2) the interplay effect between scanning spots and moving target, which could cause the target dose to be degraded. The first difficulty also exists for the DS therapy of thoracic tumor, while the DS techniques are less susceptible to interplay effect than the PBS techniques [2,12]. To fully exploit the dosimetric and physical benefits of PBS in the treatment of lung cancer patients, these two critical issues must be addressed. With regard to range uncertainties, a 'smearing' approach was first proposed to add 0.5–1.0 cm proximal and distal margins to the target in the DS treatment of thoracic tumors by Moyers et al. [13]. For proton PBS treatment planning, Park et al. introduced a beam-specific PTV (BSPTV) [14] initially proposed by Rietzel and Bert [16] to explicitly include variation

CONTACT Minglei Kang Minglei.Kang@uphs.upenn.edu or mleikang@gmail.com 3400 Civic Center Blvd, TRC2, Philadelphia, PA 19104, USA
*These authors have contributed equally to this work.

Supplemental data for this article can be accessed here.

in the water equivalent path length along each beam direction as target margins. Knopf et al. [17] also proposed a range-adapted internal target volume (raITV) concept to compensate for possible range variations caused by motion. Lin et al. [8] further improved the BSPTV method using 4DCT to quantify the motion to provide appropriate motion margins and to minimize the target volume through beam gantry angle optimization.

For PBS treatment, it is often helpful to obtain a more accurate plan evaluation by incorporating the details of spot delivery and patient breathing pattern into the dose calculation, such as relating the time sequence of each spot to the corresponding 4DCT phase. A number of varying approaches to evaluate the interplay effect for treating moving targets have been proposed [2,15–23]. The 4D dose calculation method has been widely accepted to study the interplay effect from both the inter- and intra-fraction motion [7,21]. Chang et al. [20], Li et al. [21] and Kardar et al. [22] reported on the use of 4D dynamic dose by accumulating the spots dose on the corresponding respiration phase to the reference 4DCT phase, combined with the composite doses calculated by deformable image registration methods, to evaluate the actual delivered dose for PBS in the absence of respiratory motion mitigation. Zeng et al. [23] investigated the impact of the interplay effect for mediastinal lymphomas on PBS plan robustness using 4D dose calculation and determined it to be minimal when volumetric repainting and/or larger spots are employed in proton PBS. Chang et al. [20] and Kardar et al. [22] proposed standard work flows to guide the use of PBS to treat lung tumors at MD Anderson Cancer Center (MDACC). Chang et al. [20] used 5 mm geometric motion criteria to triage the use of PBS in lung patients. In contrast, Kardar et al. [22] used CTV dose coverage as a criteria; if the difference in 4D dose coverage between one fraction and the total fractions was smaller than 3%, intensity-modulated proton therapy (IMPT) was recommended, and otherwise, an isolayer rescanning method was adopted to mitigate the interplay effect. Dowdell et al. [2] and Grassberger et al. [19] also studied the interplay effect for lung tumor treatment using Monte Carlo simulation with a deformable registration method, demonstrating that motion interplay is a function of patient-specific parameters as well as PBS spot size.

Previous studies have shown that the interplay effect between the moving targets and scanning spots is machine and patient specific and depends on beam characteristics, target location, tumor volume, and motion magnitude [1,2,19,21,24]. Based on specific beam line characteristics, the magnitude of the motion becomes very important criteria in determining the appropriateness of PBS. Therefore, it is necessary to first quantify the motion characteristics to determine those patients who could benefit dosimetrically from PBS over DS or IMRT techniques without motion mitigation and that need motion mitigation in order to receive PBS treatment. Mori et al. [25,26] first proposed analysis of proton range using 4DCT. The method was subsequently adopted in the interplay study of Kardar et al. [22] to correlate the internal target motion using water equivalent thickness (WET) changes between the different 4DCT phases. Chang et al. [20] also used WET analysis in their interplay study to determine the best gantry angles to minimize the motion effect. In this study, we further develop and enhance this approach by providing a statistical motion index (MI) to quantify the motion magnitude by comparing the proton range changes between the exhale and inhale 4DCT phases. Finally, we use the MI to assess the interplay effect for lung cancer PBS treatment with tumor motion larger than 5 mm.

Methods

Patient characteristics

Ten consecutive patients with stage III locally advanced non-small cell lung cancer (NSCLC) who were previously treated with DS were selected to study the PBS interplay effect. The target motion in the breathing cycle was analyzed along three directions: anterior-posterior (AP), left-right (LR), and superior-inferior (SI), on 4DCT images. Table 1 lists patient characteristics including the target volume, location, motion magnitude and respiration period. The target motion of the patients ranged from 4 to 11 mm with of a mean value 7.4 mm and a standard deviation (STD) of 2.0 mm. All but one patient had >5 mm target motion in the SI direction. We hypothesized that the interplay effect would have a large impact on the delivered dose in the presence of large amplitude motion. The respiration periods are the averaged breathing periods during the 4DCT acquisition with slight variations within 10%. For different treatment fractions, a patient's breathing magnitude and periods can vary. Knopf et al. [1], Zhang et al. [24], Dowdell et al. [2] and Grassberger et al. [19] used a default breathing period 5 s for all the patient studies. By adjusting the breathing period from 4.0–6.0 s, Dowdell et al. [2] found the varying breathing period of the patient will further smear out the dose distribution, additionally reducing the interplay effects. Li et al. [21] also studied the impact of variations within the breathing period on the interplay effect. They used both the sinusoidal function with fixed period of 4.2 s, and three realistic breathing patterns extracted from patient RPM data with various irregularities, characterized by the standard deviation in the length of breathing cycles (0.13, 0.46 and 0.93 s). Li et al. [21] observed very small difference for the delivery of a single fraction between these three breathing patterns, and no

Table 1. Tumor volume, location, motion characteristics and breathing period of the 10 selected patients.

				Tumor motion (mm)		
Patient No.	Volume (cc)	Location	Breathing period (s)	SI	AP	LR
1	102.7	LM	4.7	4	2	2
2	482.9	RU + RM	3.0	8	2	2
3	147	RM	6.9	8	4	2
4	135.5	RM	3.2	5	3	2
5	118.5	RM	4.2	7	5	2
6	396.4	RL	4.2	11	3	3
7	198.9	RU	3.5	7	3	2
8	220.3	RL	2.2	6	4	2
9	424.1	LM + LU	3.3	7	4	3
10	175.2	RM	3.6	10	4	3

LM: left middle; LU: left upper; RL: right lower; RM: right middle; RU: right upper.

observable coverage difference for 5, 10 and 35 fractions under different breathing conditions. This present study tries to establish the relationship between motion magnitude and interplay effect; irregularities of breathing patterns study are not included.

4DCT images and treatment planning

4DCT provides time-dependent 3D information of patient respiration, allowing quantitative description of motion tumor and thoracic/abdominal structures [25–27]. 4DCT images were acquired under free breathing conditions and patient breathing patterns recorded using the Varian Real-time Position Management (RPM) system (Varian Medical System, Palo Alto, CA). Each breathing period was divided into eight phases (0%, 12.5%, 25%, 37.5%, 50%, 62.5%, 75% and 87.5%). The CT images were time stamped and the corresponding CT phases were binned according to the RPM signal. The 0% phase images, denoted as CT0, corresponded to end inspiration while 50% phase images, denoted as CT50, corresponded to end exhalation. Voxel-to-voxel averaged CT images were generated following acquisition, and treatment plans were optimized on the averaged CT images.

The clinical target volume (CTV) was defined on all eight CT phases by board-certified thoracic radiation oncologists and summed to create an iCTV. Lung, heart and spinal cord were contoured on the average CT and on all eight phases of the 4DCT images. In conventional radiation therapy, an additional margin of 5–8 mm was used to expand the iCTV to a planning target volume (PTV) to account for uncertainty in patient setup. With PBS, variations in the proton range due to the inaccurate modeling of Hounsfield unit (HU) to stopping power ratio (SPR) combined with patient setup errors and respiratory motion can result in significant target underdose. In this study, we used a beam-specific PTV (BSPTV), which incorporates respiratory motion, patient setup errors and SPR uncertainties, to provide appropriate treatment margins and thus ensure adequate target coverage. The BSPTV were calculated using an in-house code and were subsequently imported into a commercial treatment planning system (Eclipse v13, Varian Medical System, Palo Alto, CA). The volume of the BSPTV changes as a function of the gantry angle; the algorithm is described in more detail in reference [8]. In Figure 1(a) and (b), the iCTV is shown in red, with the corresponding BSPTVs in blue and cyan contours at gantry angles of 180° and 210°, respectively. Two single-beam PBS plans were optimized on the averaged CT separately, and the BSPTV was targeted to deliver 1.8 GyE ×37 fractions covering 100% of the iCTV using a single field optimization (SFO) method. As illustrated in Figure 1(c), the final plan was obtained by summing the two beams together with a 50% weighting factor.

The scanning beam parameters

The cyclotron (IBA, Belgium) accelerates protons to 230 MeV, which are subsequently degraded to various beam energies to meet the clinical requirement to fully cover the distal-to-proximal extent of the target. Beam tuning and energy switching requires ∼2s. Spots are deflected using two orthogonal dipole magnets to cover the target in the direction perpendicular to beam. As the lowest energy from the beam line is 100 MeV, additional anterior and lateral bolus with a water equivalent thickness (WET) of 7.4 cm are used to further decrease the beam energy to cover the shallow portion of the targets. In this study, the plans were optimized using the beam configuration of the universal nozzle (UN), with spot σ in air varying from 2.5 mm to 5 mm for energies of 225 MeV to 100 MeV, respectively [28,29]. On our IBA system, each PBS spot requires ∼4–11 ms to be delivered, depending on each spot's monitor units (MU). For breathing periods of 2–6 s, each phase will have a length of ∼250–750 ms if divided into eight time-averaged CT phases.

Motion magnitude quantification with the beam-specific water equivalent thickness statistical analysis (WETSA) method

Evaluating the magnitude of motion identified from 4DCT images is a commonly accepted criterion to decide whether PBS is appropriate for the treatment of thoracic tumors. Chang et al. [20] observed that when the motion magnitude is larger than 5mm, IMPT was not routinely recommended at MDACC due to the dose degradation by interplay effect and range uncertainties. In that study, the motion magnitude in the SI, AP and RL directions was defined by the maximum

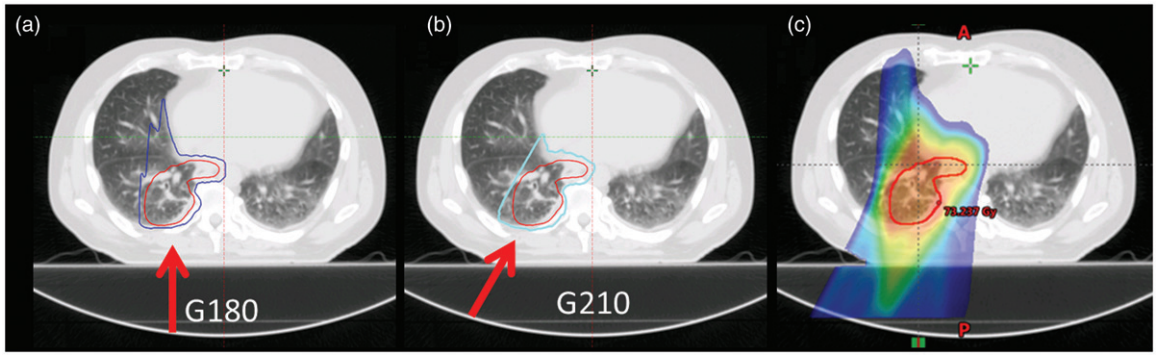

Figure 1. The BSPTV method for optimizing proton PBS plans: (a) and (b) the BSPTVs at gantry angles of 180° and 210° based on the same iCTV, (c) shows the final dose distribution by summing each of the two SFO plans with a 50% weighting factor.

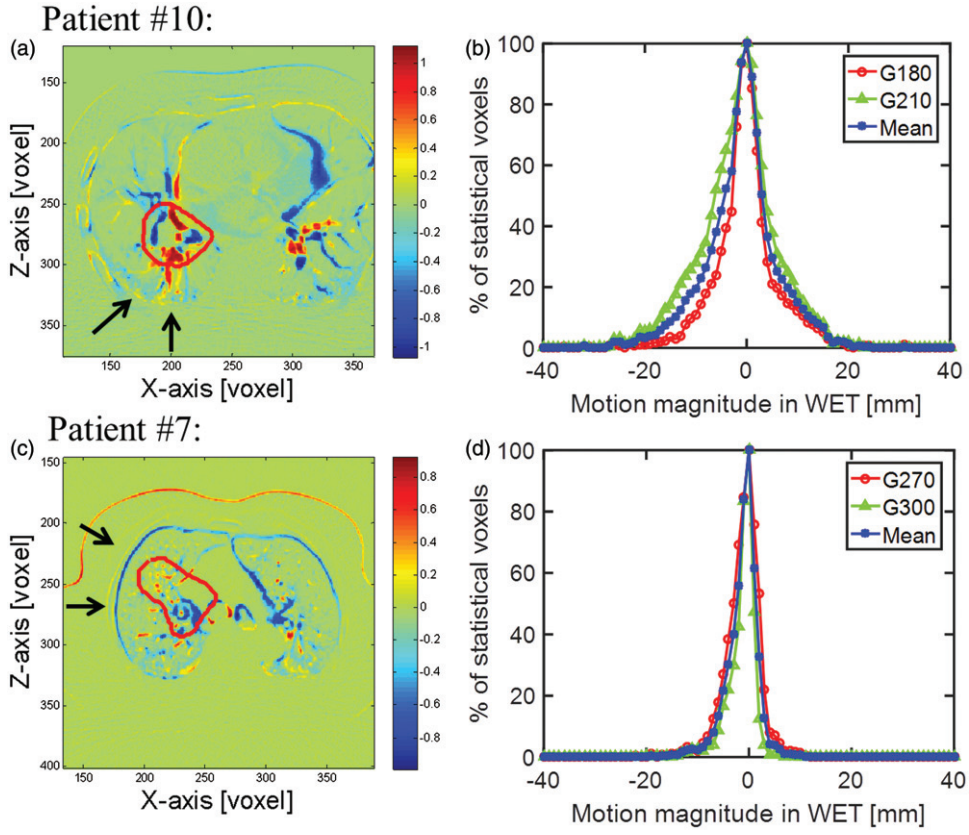

Figure 2. (a) and (c) are the Δrsp maps between CT0 and CT50 of two selected CT slices for patients 10 and 7, using CT0 (maximum inhalation) minus CT50 (maximum exhalation), the contour denotes iCTV, the arrows indicate the beam gantry angles chosen in the treatment planning; (b) and (d) are the statistical distribution of the target motion between CT0 and CT50 for all the sampled voxels of the iCTV at four different beam gantry angles 180°, 210°, 270° and 300° (the percentage of the voxels was normalized to the maximum value).

center-to-center displacement on the 4DCT images. This method was widely used in the tumor motion evaluation, but it also has its limitation to exclude some potential patients who could benefit from PBS treatment. Based on Mori [25,26] and Chang's [20], we further proposed an (MI) to quantify the tumor motion used in the 4D dose accumulation. As proton range is determined by the relative SPR or the radiological WET of individual beam path, the WET variation will change dramatically at different beam gantry angles due to tissue heterogeneity. Therefore, the beam-specific water equivalent thickness statistical analysis (WETSA) method may be a more reliable metric reflecting tumor motion. Additionally, using a statistical methodology that quantifies the motion in WET can more accurately represent tumor motion than the center-to-center motion criteria.

End inhalation (CT0) and end of exhalation (CT50) were selected as two extreme scenarios to acquire the WET difference (ΔWET) for a respiration cycle. Figure 2(a) and (c) show the voxel-to-voxel relative stopping power (Δrsp) maps for two selected CT slices for patients 10 and 7. The indicated color bar shows that the largest WET variations occur near the interface of the high- and low-density tissues. The integral ΔWET per unit radiological path was calculated from the surface of patient to the iCTV by integrating the Δrsp of each voxel between tissue and water using Equations (1) and (2), ρ_m and ρ_w denote the mass density of tissue and water, respectively, and $\overline{S_m}$ and $\overline{S_w}$ are the mean stopping powers of tissue and water, respectively. The motion magnitude of each voxel in the target was quantified by its integral ΔWET.

$$\Delta\mathrm{WET}_{x,y,z} = \int_{\mathrm{skin}}^{\mathrm{Depth}} \Delta\mathrm{rsp}(x,y,z)\mathrm{d}z \tag{1}$$

$$\Delta\mathrm{rsp}(x,y,z) = \left(\frac{\rho_m}{\rho_w}\frac{\overline{S_m}}{\overline{S_w}}\right)_{\mathrm{CT0}} - \left(\frac{\rho_m}{\rho_w}\frac{\overline{S_m}}{\overline{S_w}}\right)_{\mathrm{CT50}} \tag{2}$$

$$\mathrm{MI}\ (>5\mathrm{mm\ WET}) = \frac{\text{Number of voxels with motion} > 5\mathrm{mm\ WET}}{\text{Total reprsentative voxels of target}} \tag{3}$$

Figure 2(b) and (d) show the percentage of voxels in the target as a function of the motion magnitude in WET. Patient 10 had a much larger tumor motion than patient 7, while for patient 10, the beam from 180° contributed less motion uncertainty than that of the beam from 210°; while for patient 7, the beam from 300° had smaller motion uncertainty than the beam from 270°. To apply the WETSA method in a practical manner, we further defined an MI (>5mm WET)) as the percentage of the sample voxels of the target with absolute integral ΔWET larger than 5 mm, as shown in Equation (3). The histogram of the voxel WET change within the target follows a quasi-Gaussian distribution, we assume similar correlation can be found between MI at larger thresholds such as 4–8 mm and dose degradation. However, too small thresholds of 2–3 mm are unacceptable because only large proton range change in WET could cause large dose degradation. The higher the MI, the larger the

target motion. All the MIs of each of the two beam gantry angles for one PBS plan were calculated and averaged, listed in supplementary Table 1, as a reference for the interplay study. For example, MI (>5mm WET) of patient 7 at gantry angles 270° and 300° are 18% and 12%, while the MI (>5mm WET) of patient 10 at gantry angles 180° and 210° are 33% and 47%, which has a motion magnitude more than twice that of patient 7. As demonstrated, the motion varies greatly for different patients and is also beam angle dependent, thus a beam with a smaller MI (>5mm WET) can make a plan be more robust. The WETSA method applied onto different CT phases reflects the actual proton range variations within a breathing cycle. The motion of some voxels in Figure 2 (b) is as large as ~30 mm WET, indicating the largest range uncertainties between the inhale and exhale phases can exist in some parts of the tumor at a particular beam angle. The statistical values provide us with more accurate information of the motion to help make an informed decision for consideration of PBS treatment.

4D dose accumulation

The 4D dose calculation uses a method similar to that of the PSI [1,24], MDADCC [20–22], MGH [2,19] and GSI [30,31] groups. In addition to requiring the 3D static dose calculation and 4DCT images, we also used a dose deformable registration tool, while including the time sample of the spots in beam delivery, and the breathing patterns (see patient characteristics). Briefly, we simulated the dynamic treatment process by accumulating each spot dose on the corresponding respiration phase to the reference 4DCT phase (CT50).

For a typical field, one layer can have 600 or more spots; in the 4D dose simulation, therefore, the first step is to assign spots to each corresponding CT phase. For multiple treatment-room facilities, the beam needs to be requested for each treatment field. In a treatment plan with 2–3 fields, as is routinely used when treating lung cancer, the initial delivery time can begin at any point of the breathing cycle (assuming gating for respiration is not used). Thus, we chose the starting point of each beam delivery from a random distribution.

To acquire actual spot time samples, PBS plans were delivered using the UN in clinical mode. The spot time sequence of each treatment beam was extracted from the log file of the IBA scanning system and then registered onto the 8-phase 4DCT images, i.e., spot-segment, randomly choosing a delivery starting point within a breathing cycle using an in-house MatLab (MathWorks, MA) code. Therefore, the spots' time samples reflected the actual clinical treatment conditions. Supplementary Table 2 shows the total spots distribution in each energy layer for one beam of one patient, which has been divided into eight spot-segments. A '0' indicates that there is no spot at a particular location/phase. The dose for each spot-segment was calculated in Eclipse v13 on the corresponding CT phases, without changes to any other initial beam setting. Seven dose matrices (excluding CT50) were accumulated onto the exhale CT phase using a commercial deformable image registration package (MIM Maestro, MIM Software Inc., Cleveland, OH). 37 uncorrelated fractions were accumulated to simulate the dose interplay effect, and for each patient $37 \times 8 = 296$ segment plans were generated. The dose distributions between original and 4D dose simulation plans were compared for targets and organs-at-risk (OARs) corresponding to the average CT and CT50.

Statistical errors within the 4D dose simulation

In this study, 4D dose distributions for 1 fraction, 5 fractions and 37 fractions using statistical methods to assess the motion effect for each of our 10 patients were evaluated. 4D dose accumulation for a single fraction was determined from the expectation and standard deviation of all 37 uncorrelated fractions. The 37-fraction target coverage was averaged to obtain the expectation. When gating is not used, the initial beam delivery time or initial phase of a given fraction will occur at random. As the number of fractions increases, the distribution of initial phases will be close to a uniform distribution. The target coverage distribution of the fractions will therefore follow a uniform dose distribution. Choosing five plans from 37 fractions will generate 435,897 combinations, which is computationally prohibitive for an interplay study. Instead, we selected nine representative dose distributions (including the three that have the worst target coverage, three with the median target coverage and the three with the best coverage) to simulate 5 fractions from the 37 total 4D dose simulations to produce $C_9^5 = 126$ dose distribution combinations. 4D dose for the 5 fractions was determined from the expectation and standard deviation of the dose distribution of the 126 combinations.

To mitigate large tumor motion, we propose 'layer synchronization' as an alternative method to gating. Layer synchronization synchronizes each proton energy layer to an optimal breathing phase(s) instead of synchronizing beam-on to a breathing phase(s) in the gating method.

Results

Target robustness analysis of the 4D dose simulation

Target coverage is defined as the percentage of the CTV receiving the prescription dose. Table 2 lists the coverage of the static plans and 4D simulated plans with standard deviations. For a single fraction delivery, target coverage was significantly degraded, ranging from 98.3% ($\sigma = 1.2\%$) to 57.4%

Table 2. Target coverage for the original static plans and 4D simulation of 1, 5 and 37 FXs (standard deviation in parentheses).

		Coverage [%]			
Patients	MI (>5mm WET)	Static	1 FX	5 FX	37 FX
1	4.5	99.9	98.3 (1.2)	99.9 (0.0)	99.9 (0.0)
2	6.5	99.9	96.0 (0.9)	99.4 (0.3)	99.6 (0.5)
3	10.0	99.7	94.0 (3.1)	99.4 (0.1)	99.5 (0.1)
4	11.0	99.0	88.0 (5.3)	99.8 (0.3)	99.9 (0.3)
5	11.5	99.9	86.9 (4.7)	98.6 (1.1)	99.0 (0.5)
6	13.5	99.8	88.7 (4.3)	99.6 (0.3)	99.9 (0.1)
7	15.0	99.9	80.7 (6.2)	97.5 (1.8)	99.5 (0.9)
8	22.0	100.0	77.4 (5.2)	91.0 (2.1)	94.5 (0.7)
9	31.5	99.0	75.0 (8.4)	82.5 (5.8)	89.9 (0.9)
10	40.0	99.5	57.4 (19.4)	72.6 (12.3)	79.1 (6.2)

(σ =19.4) for the 10 selected patients. The magnitude of target underdose in single fraction plans is proportional to the MI (>5 mm WET) ($p < .001$), as shown by fitting the data with a linear equation in Figure 3, with larger motion equating to greater dose degradation and standard deviations. For three patients with the smallest motion, MI (>5mm WET) was less than 10%, with target dose coverage of 98.3% ($\sigma = 1.2\%$), 96.0% ($\sigma = 0.9\%$) and 94.0% ($\sigma = 3.1\%$) for a single fraction. Conversely, patient 10 had the largest MI (>5 mm WET) ~ 40% and the poorest target coverage at 57.4% ($\sigma = 19.4\%$). Similarly, target coverage could be maintained above 97.5% ($\sigma = 1.8\%$) in patients with intermediate motion magnitude, with MI (>5mm WET) ranging from 10–20%, in a five fractions course of delivery. With the delivery of 37 fractions, both small and intermediate motion patients maintained as target coverage equivalent to the original plan with no motion, as shown by the green curve in Figure 3. For the large motion patients 8, 9 and 10 with MI (>5mm WET) larger than 20%, adequate target coverage using proton PBS without any motion mitigation could not be achieved.

An alternate method of quantifying the dose degradation caused by interplay is to analyze the homogeneity. A homogeneity index (HI) was defined as $(D_5 - D_{95})/D_p \times 100$, where D_5 is the minimum dose to 5% of the target volume (indicating the 'maximum dose'), D_{95} is the minimum dose to 95% of the target volume (indicating the 'minimum dose'), and D_p is the prescribed dose. A smaller HI (the optimal value approaches 0) corresponds to a more homogeneous dose distribution [32,33]. As shown in Figure 4, for all the 10 patients (with the exception of patient 2), the smaller the MI, the better the dose homogeneity ($p < .001$).

Minimum, maximum and mean target dose were also used as indicators to quantify the interplay effect. From the statistical perspective, the interplay will cause cold and hot dose regions within the target. As shown in Figure 5, the minimum, maximum and mean dose of the 4D dose

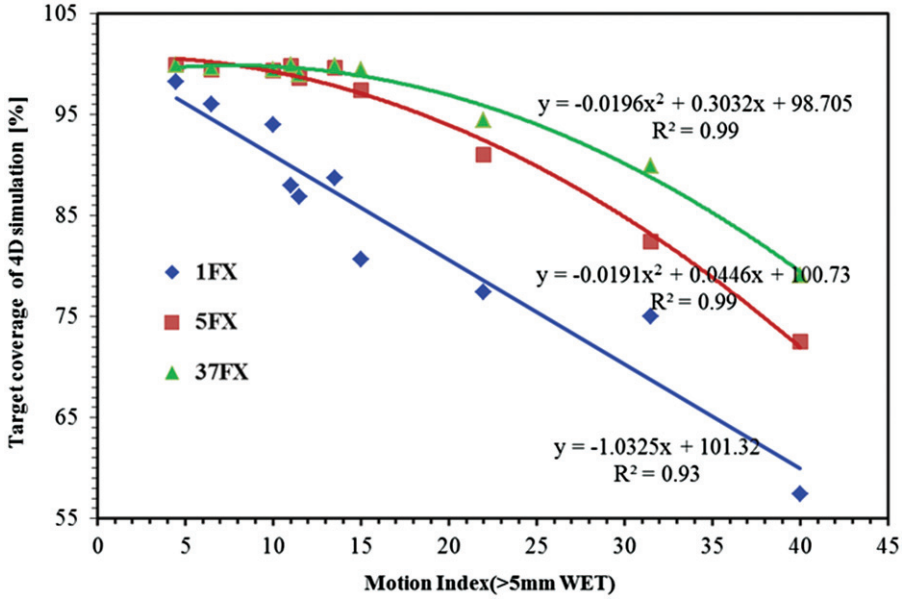

Figure 3. The CTV coverage in 4D dose simulations is changing as a function of MI for 10 patients. Patients 1–10 are plotted sequentially following the increasing MI and the three curves are the trends of the dose degradation in 1, 5 and 37 FXs, fitted by linear and polynomial functions with R square values of .93, .99 and .99. The coverage in a single fraction, 5 and 37 fractions is correlated with MI (>5mm WET) (p values <.001). Patients 1 and 10, who had the smallest and largest MI at 4.5% and 40%, also had the best and the poorest coverage (The abbreviation FX refers to the number of fractions.)

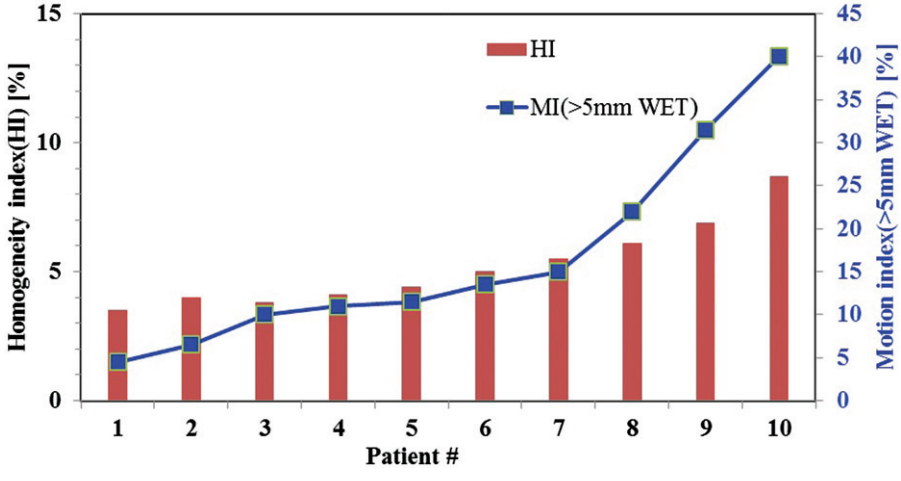

Figure 4. Variation of Homogeneity index (HI) as a function of the MI (>5mm WET) ($p < .001$).

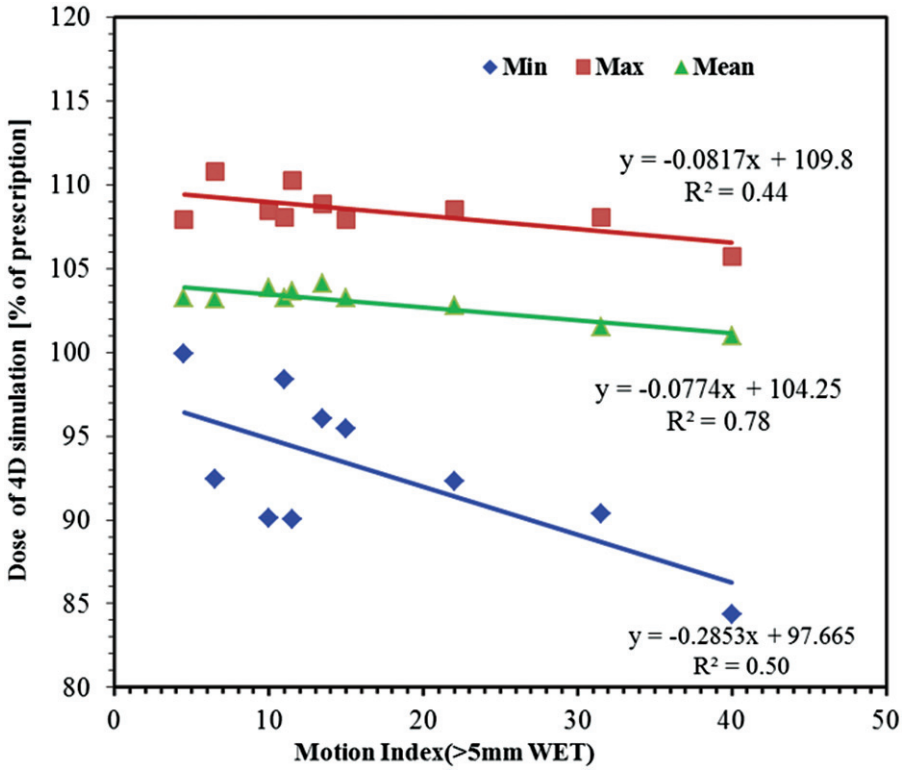

Figure 5. The minimum, maximum and mean CTV dose over 37 fractions using 4D accumulation for each of the ten patients as a function of MI (>5mm WET), $p < .05$, .05 and .001, respectively. The trends were fitted using linear functions, and the R squares are .50, .44 and .78, respectively. Patients 1–10 are plotted sequentially following the increasing MI.

accumulations of 37 fractions vary as a function of MI (>5mm WET). As the minimum and maximum dose quantities are values from single voxel, they are inclined to be more sensitive to the motion magnitude and have larger statistical errors ($p < .05$), compared with the mean dose ($p < .001$). The linear fitting equations in Figure 5 show the trend of the dose degradation of varying with MI (>5mm WET); the mean dose has a stronger correlation with MI (>5mm WET) than the minimum and maximum dose.

OARs dose comparison between the original plan and 4D dose simulation

OARs dose differences were calculated for lung V5, V20 and mean dose, heart V40, esophagus V55, and maximum cord dose and are shown in supplementary Figure 1. The maximum difference between the original plans and 37 fraction 4D dose accumulation were 1.2%, 1.6% and 0.8Gy, 1.3%, 2.2% and 1.5 Gy, respectively. For most of our 10 patients (with the exception of patients 2 and 9), the lung V5, V20 and mean dose were found to be lower than the static plan without any motion. For the heart, esophagus and cord, no significant variation was observed.

Discussion

Large respiratory motion can lead to significant dose degradation in PBS delivery to locally advanced NSCLC patients. In patients with an MI (>5mm WET) ≤10%, a single fraction 4D dose simulation could maintain target coverage of 94.0% ($\sigma = 3.1\%$), 96.0 ($\sigma = 0.9\%$) and 98.3% ($\sigma = 1.2\%$). In this study, for MI (>5mm WET) between [10–20%], acceptable coverage (97.5–99.8%) could be achieved in the three patients undergoing moderate motion in five fractions. For the patient with the largest MI (>5mm) of ~40%, however, acceptable target coverage could not be achieved even with the delivery of 37 fractions. The multiple fractions delivery simulation represents a kind of volumetric rescanning/painting [1,34,35], which cannot maintain an acceptable coverage for patients with these large motion characteristics. For these cases, we propose a layer synchronization method as an alternative option to mitigate the large tumor motion, which intentionally delivers PBS spots in regular, rather than random patterns.

In 4D dose simulation for PBS plans, the random starting point within one breathing cycle will determine the spot distribution onto the corresponding CT phases. As supplementary Table 1 shows, the number of MUs was determined accordingly to the numbers of spots on each CT phase. The spot/MU distribution over the eight CT phases in one 4D dose simulation was referred to as a spot/MU pattern. In Figure 6(a) and (b), MU distribution is plotted as a function of the eight CT phases, with each MU plot related to one unique dose distribution in the 4D dose simulation (see Figure 6(c)). 4D dose accumulation demonstrates that some MU patterns (see Figure 6(a)), have poorer dose coverage, lower than 15%, while other patterns (see Figure 6(b)) can achieve good dose coverage, greater than 90%. If the patient breathing signal can be monitored and simultaneously integrated into the PBS spot delivery system, it is conceivable

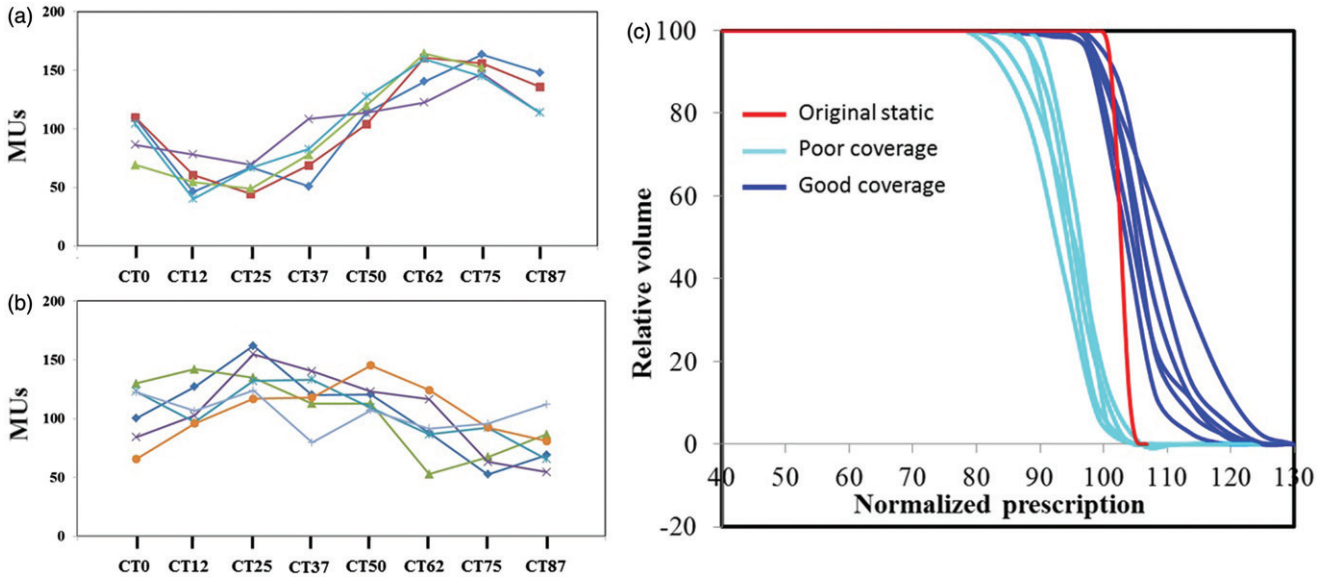

Figure 6. In the 4D dose simulation, the MUs assignment to the CT phases of each fraction has a random distribution. (a) and (b) show the MUs patterns of 11 fractions, (c) shows the target DVH. Each plot in (a) and (b) corresponded to the poor (<15%) and good (>90%) target coverage as shown in (c), the original static plan target DVH was plotted for reference.

that the beam can be triggered at a point of breathing cycle for each layer to deliver the spots accurately within the MU patterns. Although layer synchronization requires monitoring respiration and beam delivery simultaneously, it has potential advantages over gating methods. First, it can have much higher spot delivery efficiency than delivery within a gating window. Once an energy layer is trigged to start, all the spots within this layer will be delivered without stopping. In most cases, a single layer can be delivered within one breathing cycle (for example in supplementary Table 2), thus each layer can be split into 1–3 segments onto the adjacent CT phases, which can be delivered in less than three-phases of one breathing cycle. Compared with traditional gating method, layer synchronization has a larger duty cycle (beam-on percentage), thus the total beam delivery time will be determined by the number of layers, which can be equal to the treatment time without any gating. Secondly, 4D simulation indicates that numerous patterns that fall into the bands in Figure 6(b) will obtain similar good dose coverage. Therefore, PBS delivery can be selected from one of numerous patterns instead of one particular MUs pattern to enable multiple breathing phases to be suitable for each energy layer. Layer synchronization provides the flexibility of multiple-pattern combinations for spot delivery. Finally, because breathing periods (3–6 s for most patients) are comparable to the energy switching time (2 s for the IBA system), the energy switching period can be utilized to select an optimal breathing phase.

It is also important to be aware of potential limitations of layer synchronization, as using only one MU pattern can still yield poor target uniformity in patients with extremely large motion (MI (>5mm WET) larger than ~40%). For these patients, the dose distribution quality can be improved by using several MU patterns. Using the patient 10 as an example, supplementary Figure 2 (a) and (b) show one-slice of the 2D original static PBS plan dose distribution and 37 fractions 4D dose accumulation, respectively. The low dose region exists for the 37 fractions 4D dose accumulation; the DVH in supplementary Figure 2(f) only has 80% coverage. Using 1, 4 and 6 MU patterns can maintain coverage >97%, the max doses are 84, 82 and 79Gy, and the high-dose area was decreased dramatically from supplementary Figure 2(c), (d) to (e), the uniformity was also increased in the DVH of supplementary Figure 2(f). The layer synchronization spot delivery method is potentially an alternative solution for improving delivery efficiency and achieving target coverage for proton PBS treatment with fewer fractions/paintings. However, to achieve the true layer synchronization, the signal from a motion monitoring system such as RPM should be synchronized dedicated with the spot delivery system. Ideally, patients would also be coached to keep a more reproducible breathing pattern during each treatment. Additional study is needed to verify the feasibility of this alternative motion mitigation option.

While 4D dose accumulation is a useful verification tool to evaluate the severity of dose variation prior to proton PBS treatment, a single 4DCT set may not reflect the actual treatment condition, and 4D dose simulation accuracy also depends on the quality of the deformable image registration. Setup errors, breathing pattern changes, intra- and inter-fraction anatomical changes, baseline shifts, motion mitigation methods and characteristics of the beam line could cause dose errors. The dose mapping by deformable image registration can also introduce propagation errors [36]. Each of the above limitations should be carefully considered when the 4D dose simulation is performed to estimate the dose variations during clinical applications.

Conclusions

The WETSA method generates a quantitative metric that potentially serves as a clinical decision tool to assess whether patients are suitable to receive PBS treatment. Degradation

in target coverage was proportional with the MI (>5mm WET) ($p < .001$). For the three patients with the smallest motion, MI (>5mm WET) $\leq$ 10%, and target dose coverage could be maintained even for a single fraction (98.3% ($\sigma = 1.2\%$), 96% ($\sigma = 0.9\%$) and 94.0% ($\sigma = 3.1\%$)). In contrast, in patients with moderate motion, MI (>5mm WET) between 10% and 20%, coverage in a single fraction was inadequate while treatment over five fractions of delivery could achieve a dose coverage above 97%. For patients with the largest motion, however, (MI (>5mm WET) > 40%) PBS without any mitigation could not provide acceptable dose coverage. A layer synchronization method is therefore proposed to improve the PBS delivery efficiency and mitigate interplay effects to achieve target dose uniformity with fewer treatment fractions/paintings for patients with largest MI (>5mm WET).

Acknowledgements

Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the US Army. The authors would like to thank David Alkins for his help in upgrading the test TPS system.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Funding

This work was supported in part by the U.S. Army Medical Research and Materiel Command under Contract Agreement No. DAMD17-W81XWH-07-2-0121 and W81XWH-09-2-0174, and in part by a Varian Medical Systems grant.

ORCID

Minglei Kang http://orcid.org/0000-0002-5054-7587
Liyong Lin http://orcid.org/0000-0002-0820-8063

References

[1] Knopf AC, Hong TS, Lomax A. Scanned proton radiotherapy for mobile targets – the effectiveness of re-scanning in the context of different treatment planning approaches and for different motion characteristics. Phys Med Biol. 2011;56:7257–7271.
[2] Dowdell S, Grassberger C, Sharp GC, et al. Interplay effects in proton scanning for lung: a 4D Monte Carlo study assessing the impact of tumor and beam delivery parameters. Phys Med Biol. 2013;58:4137–4156.
[3] Kooy H, Clasie BM, Lu HM, et al. A case study in proton pencil-beam scanning delivery. Int J Radiat Oncol Biol Phys. 2010;76:624–630.
[4] Chang JY, Cox JD. Improving radiation conformality in the treatment of non-small cell lung cancer. Semin Radiat Oncol. 2010;20:171–177.
[5] Chang JY, Zhang RX, Wang X, et al. Significant reduction of normal tissue dose by proton radiotherapy compared with three-dimensional conformal or intensity-modulated radiation therapy in Stage I or Stage III non–small-cell lung cancer. Int J Radiat Oncol Biol Phys. 2006;65:1087–1096.
[6] Kang Y, Zhang X, Chang JY, et al. 4D proton treatment planning strategy for mobile lung tumors. Int J Radiat Oncol Biol Phys. 2007;67:906–914.
[7] Zhang RX, Li YP, Pan X, et al. Intensity-modulated proton therapy reduces the dose to normal tissue compared with intensity-modulated radiation therapy or passive scattering proton therapy and enables individualized radical radiotherapy for extensive stage IIIB non-small-cell lung cancer: a virtual clinical study. Int J Radiat Oncol Biol Phys. 2010;77:357–366.
[8] Lin L, Kang M, Huang S, et al. Beam specific planning target volumes incorporating 4DCT for pencil beam scanning proton therapy of thoracic tumors. J Appl Clin Med Phys. 2016;16: 281–292.
[9] Kesarwala AH, Ko CJ, Ning H, et al. Intensity-modulated proton therapy for elective nodal irradiation and involved-field radiation in the definitive treatment of locally advanced non–small-cell lung cancer: a dosimetric study. Clin Lung Cancer. 2015;16:237–244.
[10] Wink KC, Roelofs E, Solberg TD, et al. Particle therapy for non-small cell lung tumors: where do we stand? A systematic review of the literature. Front Oncol. 2014;4:292
[11] Paganetti H. Range uncertainties in proton therapy and the role of Monte Carlo simulations. Phys Med Biol. 2012;57:R99–R117.
[12] Bortfeld T, Jokivarsi K, Goitein M, et al. Effects of intra-fraction motion on IMRT dose delivery: statistical analysis and simulation. Phys Med Biol. 2002;47:2203–2220.
[13] Moyers MF, Miller DW, Bush DA, et al. Methodologies and tools for proton beam design for lung tumors. Int J Radiat Oncol Biol Phys. 2001;49:1429–1438.
[14] Park PC, Zhu XR, Lee AK, et al. A beam-specific planning target volume (PTV) design for proton therapy to account for set up and range uncertainties. Int J Radiat Oncol Biol Phys. 2012;82:329–336.
[15] Bert C, Rietzel E. 4D treatment planning for scanned ion beams. Radiat Oncol. 2007;2:24.
[16] Rietzel E, Bert C. Respiratory motion management in particle therapy. Med Phys. 2010;37:449–460.
[17] Knopf AC, Boye D, Lomax A, et al. Adequate margin definition for scanned particle therapy in the incidence of intrafractional motion. Phys Med Biol. 2013;58:6079–6094.
[18] Bert C, Grozinger SO, Rietzel E. Quantification of interplay effects of scanned particle beams and moving targets. Phys Med Biol. 2008;53:2253–2265.
[19] Grassberger C, Dowdell S, Lomax A, et al. Motion interplay as a function of patient parameters and spot size in spot scanning proton therapy for lung cancer. Int J Radiat Oncol Biol Phys. 2013;86:380–386.
[20] Chang JY, Li H, Zhu XR, et al. Clinical implementation of intensity modulated proton therapy for thoracic malignancies. Int J Radiat Oncol Biol Phys. 2014;90:809–818.
[21] Li YP, Kardar L, Li X, et al. On the interplay effects with proton scanning beams in stage III lung cancer. Med Phys. 2014;41:021721.
[22] Kardar L, Li YP, Li X, et al. Evaluation and mitigation of the interplay effects of intensity modulated proton therapy for lung cancer in a clinical setting. Pract Radiat Oncol. 2014;4:e259–e268.
[23] Zeng C, Plastaras JP, Tochner ZA, et al. Proton pencil beam scanning for mediastinal lymphoma: the impact of interplay between target motion and beam scanning. Phys Med Biol. 2015;60:3013.
[24] Zhang Y, Boye D, Tanner C, et al. Respiratory liver motion estimation and its effect on scanned proton beam therapy. Phys Med Biol. 2012;57:1779–1795.
[25] Mori S, Wolfgang J, Lu HM, et al. Quantitative assessment of range fluctuations in charged particle lung irradiation. Int J Radiat Oncol Biol Phys. 2008;70:253–261.
[26] Mori S, Dong L, Starkschall G, et al. A serial 4DCT study to quantify range variations in charged particle radiotherapy of thoracic cancers. J Radiat Res. 2014;55:309–319.
[27] Zhang XD, Zhao KL, Guerrero TM, et al. 4D CT-based treatment planning for intensity-modulated radiation therapy and proton therapy for distal esophagus cancer. Int J Radiat Oncol Biol Phys. 2008;72:278–287.

[28] Lin L, Ainsley CG, McDonough JE. Experimental characterization of two-dimensional pencil beam scanning proton spot profiles. Phys Med Biol. 2013;58:6193–6204.
[29] Lin L, Ainsley CG, Solberg TD, et al. Experimental characterization of two-dimensional spot profiles for two proton pencil beam scanning nozzles. Phys Med Biol. 2014;59:493–504.
[30] Richter D, Graeff C, Jäkel O, et al. Residual motion mitigation in scanned carbon ion beam therapy of liver tumors using enlarged pencil beam overlap. Radiother Oncol. 2014;113:290–295.
[31] Richter D, Saito N, Chaudhri N, et al. Four-dimensional patient dose reconstruction for scanned ion beam therapy of moving liver tumors. Int J Radiat Oncol Biol Phys. 2014;89:175–181.
[32] Gong Y, Wang J, Bai S, et al. Conventionally-fractionated image-guided intensity modulated radiotherapy (IG-IMRT): a safe and effective treatment for cancer spinal metastasis. Radiat Oncol. 2008;22:3–11.
[33] Kataria T, Sharma K, Subramani V, et al. Homogeneity Index: an objective tool for assessment of conformal radiation treatments. J Med Phys. 2012;37:207.
[34] Seco J, Robertson D, Trofimov A, et al. Breathing interplay effects during proton beam scanning: simulation and statistical analysis. Phys Med Biol. 2009;54:N283–N294.
[35] Zenklusen SM, Pedroni M, Meer D. A study on repainting strategies for treating moderately moving targets with proton pencil beam scanning at the new Gantry 2 at PSI. Phys Med Biol. 2010;55:5103–5121.
[36] Li H, Zhong H, Kim J, et al. Direct dose mapping versus energy/mass transfer mapping for 4D dose accumulation: fundamental differences and dosimetric consequences. Phys Med Biol. 2014;59:173–188.