Four-dimensional Monte Carlo simulations demonstrating how the extent of intensity-modulation impacts motion effects in proton therapy lung treatments

Stephen Dowdell^{a)}

Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114

Clemens Grassberger

Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114 and Centre for Proton Therapy, Paul Scherrer Institut, 5232 Villigen-PSI, Switzerland

Harald Paganetti

Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114

(Received 13 June 2013; revised 27 September 2013; accepted for publication 20 October 2013; published 14 November 2013)

Purpose: To compare motion effects in intensity modulated proton therapy (IMPT) lung treatments with different levels of intensity modulation.

Methods: Spot scanning IMPT treatment plans were generated for ten lung cancer patients for 2.5Gy(RBE) and 12Gy(RBE) fractions and two distinct energy-dependent spot sizes ($\sigma \sim 8$ –17 mm and ~ 2 –4 mm). IMPT plans were generated with the target homogeneity of each individual field restricted to $<20\%$ (IMPT_{20%}). These plans were compared to full IMPT (IMPT_{full}), which had no restriction on the single field homogeneity. 4D Monte Carlo simulations were performed upon the patient 4DCT geometry, including deformable image registration and incorporating the detailed timing structure of the proton delivery system. Motion effects were quantified via comparison of the results of the 4D simulations (4D-IMPT_{20%}, 4D-IMPT_{full}) with those of a 3D Monte Carlo simulation (3D-IMPT_{20%}, 3D-IMPT_{full}) upon the planning CT using the equivalent uniform dose (EUD), V_{95} and D_1 – D_{99} . The effects in normal lung were quantified using mean lung dose (MLD) and $V_{90\%}$.

Results: For 2.5Gy(RBE), the mean EUD for the large spot size is $99.9\% \pm 2.8\%$ for 4D-IMPT_{20%} compared to $100.1\% \pm 2.9\%$ for 4D-IMPT_{full}. The corresponding values are $88.6\% \pm 8.7\%$ (4D-IMPT_{20%}) and $91.0\% \pm 9.3\%$ (4D-IMPT_{full}) for the smaller spot size. The EUD value is higher in 69.7% of the considered deliveries for 4D-IMPT_{full}. The V_{95} is also higher in 74.7% of the plans for 4D-IMPT_{full}, implying that IMPT_{full} plans experience less underdose compared to IMPT_{20%}. However, the target dose homogeneity is improved in the majority (67.8%) of plans for 4D-IMPT_{20%}. The higher EUD and V_{95} suggests that the degraded homogeneity in IMPT_{full} is actually due to the introduction of hot spots in the target volume, perhaps resulting from the sharper in-target dose gradients. The greatest variations between the IMPT_{20%} and IMPT_{full} deliveries are observed for patients with the largest motion amplitudes. These patients would likely be treated using gating or another motion mitigation technique, which was not the focus of this study.

Conclusions: For the treatment parameters considered in this study, the differences between IMPT_{full} and IMPT_{20%} are only likely to be clinically significant for patients with large (>20 mm) motion amplitudes. © 2013 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4829500>]

1. INTRODUCTION

Delivering proton therapy using scanning was first proposed in the early 1980s;¹ however, this technology is only now starting to be used clinically at multiple proton facilities. Scanning allows the potential of intensity-modulated proton therapy (IMPT),² which better exploits the physical advantages of protons compared to other techniques. The time-dependent nature of scanning fields makes them inherently more sensitive to intrafractional motion effects compared to passively scattered proton therapy. Any change induced in the target dose distribution due to relative motion between the target and the proton beam is defined as an interplay effect.³

IMPT involves the delivery of multiple inhomogeneous beams that result in a homogeneous target dose distribution.² Individual fields in an IMPT delivery can potentially possess steep dose gradients within the target volume, which may further increase the susceptibility of these treatments to motion effects. As such, full IMPT has typically not been adopted for routine use for treatment sites such as lung, which experience non-negligible tumor motion.

Several previous studies have investigated motion and interplay effects in proton^{4–7} and carbon ion therapy.^{8–10} These investigations of the interplay effect in proton and ion therapy have typically used single-field uniform dose (SFUD) fields that each deliver homogeneous target doses. Graeff

*et al.*¹⁰ have previously quantified motion effects in intensity modulated particle therapy using rescanning as a motion mitigation technique. However, this study represents the first comparisons of the motion effects in SFUD and IMPT lung treatments for a patient cohort.

2. METHODS

2.A. Variation in homogeneity

The homogeneity of a radiation field delivered upon a target volume is typically defined by the metric in Eq. (1).

$$\text{homogeneity} = \frac{\text{maximum dose} - \text{minimum dose}}{\text{mean dose}} \quad (1)$$

SFUD can be considered as a subset of IMPT, where the target dose homogeneity for each of the treatment fields is restricted to be within a given value, e.g. $\pm 2\%$.

There is currently no global definition or general consensus in the proton community as to what homogeneity threshold constitutes an SFUD field. Thus, we introduce the following nomenclature: IMPT_x , where x denotes the maximum homogeneity (in %), defined by Eq. (1), allowed in a single field during IMPT planning. We term this metric the single field homogeneity (SFH). An IMPT treatment with no SFH restriction will be denoted as $\text{IMPT}_{\text{full}}$.

2.B. Treatment planning

Two-field IMPT treatment plans were generated for ten lung cancer patients using ASTROID,¹¹ an in-house pencil beam scanning treatment planning platform. Plans were generated with a lateral spot spacing of 0.7σ for two distinct energy dependent spots sizes ($\sigma \sim 8\text{--}17$ mm and $\sigma \sim 2\text{--}4$ mm), which reflect the current extremes in clinical use across proton facilities. The planning method employed was based on the previous study of Kang *et al.*¹², which has also been used in previous studies.^{4,5}

The clinical tumor volume (CTV) was defined as the gross tumor volume (GTV) with a uniform expansion of 8 mm. The IGTV was defined to cover the GTV in all ten phases of the patient phase-binned 4DCT, while the ICTV covered the CTV in all phases. The density of the IGTV was overridden to 50 HU during treatment planning. This procedure has been shown to give the best compromise between target coverage and normal lung exposure.¹² A conventional planning target volume (PTV) was also defined to account for patient setup uncertainties, as a 5 mm uniform expansion of the ICTV.

ASTROID allows a SFH threshold to be implemented during planning. Treatment plans were generated for all patients on the average intensity projection CT for both $\text{IMPT}_{20\%}$ and $\text{IMPT}_{\text{full}}$ deliveries. A stricter SFH threshold than 20% was not employed as it could not be satisfied by all patients in the cohort for the larger spot size. This is partly due to the planning method, which required some spots on the PTV periphery to be placed directly in lung tissue and the large spot size.

All plans were optimized using multicriteria optimization¹³ such that the ICTV and PTV received

$\geq 95\%$ and $\geq 99\%$ of the prescription dose respectively. The dose to the normal lung, defined here as the combined lungs with the IGTV volume subtracted, and other critical structures was minimized as much as possible such that the target objectives remained satisfied.

Range shifters were required, primarily for the larger spot size at low energies, to meet the target planning constraints. The range shifter thicknesses were minimized individually for each beam such that the target planning constraints remained satisfied.

Treatment plans were generated according to two fractionation schemes: (1) $35 \times 2.5\text{Gy(RBE)}$ and (2) $4 \times 12\text{Gy(RBE)}$. Employing different fractionation schemes will alter the motion effects due to the change in the time structure of the radiation field, as outlined in a previous study.⁴

2.C. Monte Carlo simulations

The list of spots in each field was output from ASTROID and served as the basis for generating input files for the Monte Carlo simulations. The spot list contained proton energy, lateral position, and spot weight. All simulation input files were generated using MCAUTO-4D,⁴ an automated suite of scripts for handling the transfer of clinical data required for Monte Carlo simulations. All 4D Monte Carlo simulations were performed upon the patient 4DCT geometry using the TOPAS (Ref. 14) Monte Carlo code (version a9). The full treatment head was not simulated, as previous work¹⁵ has demonstrated the validity of this approach for in-field beam scanning applications.

MCAUTO-4D incorporates the full timing structure of the proton fields into the simulation framework. The time required to change the beam energy was assumed to be 1 s, with a 5 ms time delay for magnet settling invoked before the delivery of each spot. These values represent a compromise between the extremes currently used in clinical practice at different institutions. The beam current was set to 2 nA throughout each of the fields. Each spot within the treatment field was delivered to the appropriate phase of the patient 4DCT depending on its specific delivery time. If a single spot spanned multiple 4DCT phases, it was split such that the different portions of the spot were delivered to the correct 4DCT phase. The results of the individual simulations on the ten 4DCT phases for each considered configuration were combined using the deformable image registration package *plasmimatch*¹⁶ and end-of-exhale as the reference phase. Separate simulations were performed in which scanning commenced at eight locations evenly spaced in phase throughout the patient respiratory cycle, which was assumed to have a constant 5 s period.

To determine the degradation of the dose distribution due specifically to motion, the results of the 4D simulations were compared with those of a 3D Monte Carlo simulation performed on the planning CT. This methodology removed any bias based on difference in plan quality between $\text{IMPT}_{20\%}$ and $\text{IMPT}_{\text{full}}$. Theoretically, the heightened flexibility in $\text{IMPT}_{\text{full}}$ allows this technique to achieve a sharper dose-volume histogram (DVH) compared to $\text{IMPT}_{20\%}$. Comparing the

difference between the 3D (3D-IMPT_x) and 4D (4D-IMPT_x) dose distributions allows direct quantification of the motion effects while also removing any bias due to the initial plan quality. By looking at the values specifically from the 4D dose distributions, we can determine the expected dose distributions and relevant metrics for each of the techniques considered in the presence of motion.

Comparisons were performed upon the CTV using the equivalent uniform dose (EUD),¹⁷ the volume receiving 95% of the prescription dose (V_{95}) and the target dose homogeneity (D_1 - D_{99}). We did not make the additional step to determine the TCP or NTCP due to the uncertainty in the values used for such calculations. V_{95} was used as one of the planning constraints and examining this quantity in the 4D distributions shows degradation of the planned dose distribution and the amount of under-dosage. This definition of target dose homogeneity was employed, rather than D_5 - D_{95} used in previous studies^{5,18} due to its heightened sensitivity to motion effects. Previous work⁴ has shown that correlations and motion effects can be observed when examining D_1 - D_{99} , which are not observed using D_5 - D_{95} as the measure of target dose homogeneity.

3. RESULTS

Figure 1 shows differential DVHs for the two beams in the IMPT_{20%} and IMPT_{full} plans for patient 1 using the larger spot size. These histograms are calculated for the PTV using the dose calculation algorithm implemented in ASTROID and thus do not include any motion effects. They demonstrate that when no restriction on the SFH is imposed, the optimization process yields a more inhomogeneous target dose distribution

from the individual fields for this patient. Similar DVH plots were obtained for the smaller spot size and the other patients within the cohort. SFH values for each of the two fields used for all ten patients are shown in Table I for both spot sizes. The values in Table I are also obtained directly from ASTROID and thus do not include any motion effects. Furthermore, these values are independent of the fraction size. The homogeneity values for the larger spot size are typically closer to the imposed SFH restriction. The SFH for individual beams are primarily influenced by tumor location and the proximity of critical structures that must be spared during treatment, such as the heart, esophagus, and spinal cord.

Figure 2 shows the difference in the EUD values between the 3D and 4D simulations in addition to the values obtained from the 4D simulations for the different fraction and spot sizes. The data points are the average over the eight starting breathing phases, with the error bars equal to ± 1 standard deviation (SD). The average difference between the 3D and 4D EUD values is lower for IMPT_{full} for all ten patients for both fraction and spot sizes [Figs 2(a) and 2(c)]. The values in Figs. 2(b) and 2(d) show that for each of the ten patients, the average EUD values obtained from 4D-IMPT_{full} are either close to or higher than the corresponding 4D-IMPT_{20%} values.

The average EUD over all ten patients obtained with the larger spot size for the 2.5Gy(RBE) 4D-IMPT_{20%} treatments is $99.9\% \pm 2.8\%$ compared to $100.1\% \pm 2.9\%$ for 4D-IMPT_{full}. For the smaller spot size, the corresponding values are $88.6\% \pm 8.7\%$ (4D-IMPT_{20%}) and $91.0\% \pm 9.3\%$ (4D-IMPT_{full}). For 12Gy(RBE) fractions and the larger spot size, 4D-IMPT_{20%} gives an average EUD of $100.3\% \pm 2.0\%$ compared to $100.6\% \pm 2.2\%$ for 4D-IMPT_{full} over the ten patients. The smaller spot size gives an average EUD over

FIG. 1. Differential PTV dose-volume histograms for the beams 1 (left) and 2 (right) for patient 1 for both IMPT_{20%} and IMPT_{full} plans for an 87.5Gy(RBE) treatment.

TABLE I. Patient characteristics showing gross tumor volume (GTV) and motion amplitude along the superior-inferior (SI) motion axis. The single field homogeneity (SFH) values of the two IMPT fields in the treatment plans for each patient for the larger (LS) and smaller (SS) spot sizes are also shown. $\text{IMPT}_{20\%}$ = SFH restricted to $\leq 20\%$, $\text{IMPT}_{\text{full}}$ = no restriction on SFH.

Patient	GTV Volume (cc)	SI motion Amplitude (mm)	Single field homogeneity (%)							
			LS ($\text{IMPT}_{\text{full}}$)		LS ($\text{IMPT}_{20\%}$)		SS ($\text{IMPT}_{\text{full}}$)		SS ($\text{IMPT}_{20\%}$)	
1	21.1	30.9	88.8	59.7	28.5	28.8	93.0	67.4	19.8	18.0
2	64.9	5.0	75.7	69.8	29.4	28.6	95.8	97.2	17.7	15.0
3	26.0	10.7	88.8	59.7	30.7	30.3	92.6	76.6	18.7	14.5
4	82.3	15.2	108.1	92.2	32.1	32.1	105.6	102.7	17.4	17.1
5	4.0	15.1	47.1	40.6	13.9	13.9	93.6	86.5	12.7	12.9
6	21.7	8.0	66.9	83.1	28.2	28.2	91.8	114.8	31.6	31.1
7	2.6	5.6	45.6	60.6	26.5	26.5	71.1	66.1	17.1	17.9
8	15.4	16.8	68.8	57.2	26.7	26.7	67.5	65.4	10.8	11.6
9	24.5	10.0	69.0	69.4	29.2	29.2	98.1	85.3	18.1	14.2
10	33.9	3.2	98.4	64.2	23.3	30.2	115.2	80.0	15.6	11.5

all patients of $89.4\% \pm 10.8\%$ for $4\text{D-IMPT}_{20\%}$, while $4\text{D-IMPT}_{\text{full}}$ gives $91.8\% \pm 10.8\%$.

The V_{95} values are shown in Fig. 3. $\text{IMPT}_{\text{full}}$ shows a smaller average difference between the 3D and 4D V_{95} over the eight starting breathing phases for all patients and

both fraction sizes except for patient 7 with the small spots and 12Gy(RBE) fraction. The magnitude of the degradation of V_{95} is higher for the smaller spot size [Fig. 3(c)]. For the larger spot size and both fraction sizes, all $4\text{D-IMPT}_{\text{full}}$ and $4\text{D-IMPT}_{20\%}$ V_{95} values coincide within the uncertainty

FIG. 2. Difference in the EUD from the 3D and 4D simulations for the larger (a) and smaller (c) spot sizes. The EUD values from the 4D simulations are shown in (b) and (d) for the larger and smaller spot sizes respectively. All values are percentages of the prescribed dose per fraction. The data points are the average values from the eight different starting breathing phases and the error bars represent ± 1 SD. $\text{IMPT}_{20\%}$ = single field homogeneity (SFH) restricted to $\leq 20\%$, $\text{IMPT}_{\text{full}}$ = unrestricted SFH.

FIG. 3. Difference in the V_{95} from the 3D and 4D simulations for the larger (a) and smaller (c) spot sizes. The V_{95} values from the 4D simulations are shown in (b) and (d) for the larger and smaller spot sizes respectively. All values are percentages of the prescribed dose per fraction. The data points are the average values from the eight different starting breathing phases and the error bars represent ± 1 SD. IMPT_{20%} = single field homogeneity (SFH) restricted to $\leq 20\%$, IMPT_{full} = unrestricted SFH.

limits when averaging over the eight starting breathing phases considered. For the smaller spot size, 4D-IMPT_{full} gives higher average V_{95} values for all ten patients [Fig. 3(d)].

To assess the target dose homogeneity, we use D_1 - D_{99} . Previous studies (e.g., Refs. 5, 18, and 19) have employed D_5 - D_{95} to quantify the dose homogeneity; however, we chose D_1 - D_{99} due to its heightened sensitivity to differences in the dose distributions. The maximum and minimum target doses were not used to assess the homogeneity due to their heightened sensitivity to the statistics of the simulations. Figure 4 shows the D_1 - D_{99} values for the IMPT_{full} and IMPT_{20%} treatments for the considered spot and fraction sizes. Compared to the larger spot size, treatments using the smaller spot size show larger differences between the 3D and 4D homogeneity values and decreased target dose homogeneity in the presence of motion. For most cases, the D_1 - D_{99} values are higher for 4D-IMPT_{full}, showing that unrestricted SFH leads to worse target dose homogeneity in the presence of motion. However, the average differences are small, with the majority of 4D-IMPT_{20%} and 4D-IMPT_{100%} values coinciding within ± 1 SD.

The values in Table II show the range of differences in the deterioration of the dose distributions due to motion between IMPT_{full} and IMPT_{20%}. The degradation due to motion effects is defined here as the difference in the chosen dose

metric (EUD, V_{95} or D_1 - D_{99}) in the 3D simulation upon the planning CT and the 4D simulation on the ten phases of the 4DCT. The values in Table II are obtained by subtracting the IMPT_{20%}(3D-4D) values from the IMPT_{full}(3D-4D) for the same starting breathing phase. Previous work⁴ has shown that the starting breathing phase plays a significant role in the magnitude of the interplay effect. Comparing values with the same starting breathing phase removes any bias based on the starting breathing phase by ensuring that the only difference between the two simulations is the actual treatment plans. A negative value in Table II implies that the IMPT_{full} delivery was less degraded by motion than the IMPT_{20%}.

Table II shows that, particularly for patients with largest motion amplitudes (especially patient 1), the differences between IMPT_{full} and IMPT_{20%} can be substantial. The differences in the degradation of the EUD, V_{95} , and D_1 - D_{99} can be as high as 28.5%, 78.6%, and 25.3% of the prescribed dose per fraction, respectively.

The values in Table II demonstrate the range of differences; however, it is of interest to know which of the two IMPT deliveries performs better for each patient, starting breathing phase, spot size, and fraction size. These values are shown in Table III.

The upper section of Table III shows the number of times the degradation of the dose distribution in IMPT_{full} plans

FIG. 4. Difference in the D_1 - D_{99} from the 3D and 4D simulations for the larger (a) and smaller (c) spot sizes. The D_1 - D_{99} values from the 4D simulations are shown in (b) and (d) for the larger and smaller spot sizes respectively. All values are percentages of the prescribed dose per fraction. The data points are the average values from the eight different starting breathing phases and the error bars represent ± 1 SD. $IMPT_{20\%}$ = single field homogeneity (SFH) restricted to $\leq 20\%$, $IMPT_{full}$ = no restriction on SFH.

is less than for $IMPT_{20\%}$ with the same spot size, fraction size, and starting breathing phase. The EUD experiences less degradation due to motion effects using $IMPT_{full}$. $IMPT_{full}$ shows less motion effects upon the D_1 - D_{99} for the larger spot size with worse degradation of the V_{95} . This is the opposite trend observed for the smaller spot size. These values are of high clinical importance as physicians typically approve treatment plans based on the 3D dose distribution, which will be degraded by motion during delivery.

The lower section of Table III shows the number of times the 4D- $IMPT_{full}$ metrics are improved compared to 4D- $IMPT_{20\%}$. Considering all patients, spot sizes, and fraction sizes, the EUD is higher in 69.7% of the $IMPT_{full}$ deliveries and the V_{95} in 74.7%. On the other hand, the target dose homogeneity (D_1 - D_{99}) is higher in 67.8% of the $IMPT_{20\%}$ deliveries.

The largest observed differences in MLD for any patient for either fraction size is 0.9% and 0.5% for the larger and smaller spot sizes, respectively. The definition of normal lung combined with the beam angles selected and the physical properties of protons leads to large lung volumes receiving zero or close to zero dose in both 3D and 4D simulations. Hence, when comparing individual patients, the differences in MLD are within 1.0% for the larger spot size for both

fraction sizes. Comparing the volume that receives 90% of the prescribed fraction dose ($V_{90\%}$) showed maximum differences of 3.1% and 4.1% for the larger spot size for the 2.5Gy(RBE) and 12Gy(RBE) fraction sizes. For the smaller spot size, the corresponding $V_{85\%}$ values for the smaller spot size are 2.0% [2.5Gy(RBE)] and 1.8% [12Gy(RBE)].

This is also reflected in the 4D plans for the larger spot size, which show $0.4\% \pm 0.7\%$ and $0.4\% \pm 0.9\%$ higher average lung $V_{90\%}$ for 4D- $IMPT_{full}$ compared to the 4D- $IMPT_{20\%}$ for the 2.5Gy(RBE) and 12Gy(RBE) fractions, respectively. For the smaller spot size, the 4D- $IMPT_{full}$ plans give an average $V_{90\%}$ $0.6\% \pm 0.5\%$ for both fraction sizes.

4. DISCUSSION AND CONCLUSION

We present here a new convention for the description of IMPT fields in terms of the allowed homogeneity per field. Currently, the criteria for a given proton field to be considered “SFUD” varies between proton facilities. The SFH, presented in Eq. (1), is based on commonly employed and available dose metrics, routinely used in current proton therapy practices. Adoption of this naming convention will increase consistency across the proton therapy community, which currently has disparities between individual institutions.

TABLE II. Maximum differences between $\text{IMPT}_{\text{full}}$ and $\text{IMPT}_{20\%}$ plan degradation (3D-4D) due to motion effects. The values are evaluated as $\text{IMPT}_{\text{full}}(3\text{D}-4\text{D}) - \text{IMPT}_{20\%}(3\text{D}-4\text{D})$ for the same spot size, fraction size, and starting breathing phase. All values are expressed as percentages of the dose per fraction and encompass the observed differences between $\text{IMPT}_{\text{full}}(3\text{D}-4\text{D})$ and $\text{IMPT}_{20\%}(3\text{D}-4\text{D})$. A negative value indicates larger differences in the 3D and 4D plans for $\text{IMPT}_{20\%}$.

Patient	EUD(3D)–EUD(4D)		V ₉₅ (3D)–V ₉₅ (4D)		D ₁ –D ₉₉ (3D)–D ₁ –D ₉₉ (4D)	
	2.5Gy(RBE)	12Gy(RBE)	2.5Gy(RBE)	12Gy(RBE)	2.5Gy(RBE)	12Gy(RBE)
Large spots						
1	–5.0–2.8	–20.5–14.8	–12.6–25.1	–78.6–34.3	–6.2–6.5	–8.0–14.3
2	0.3–1.3	0.1–1.5	–1.0–0.7	–0.4–1.8	–0.3–2.4	–1.5–1.6
3	–1.2–1.0	–2.8–2.0	–3.7–5.1	–2.7–3.4	–0.9–4.7	–3.5–4.1
4	–0.1–4.9	–3.5–8.7	–3.6–8.3	–36.1–6.0	–1.5–5.2	–0.9–9.3
5	0.2–1.9	0.5–1.8	–2.0–5.2	–0.7–2.5	–2.9–4.6	–5.3–3.0
6	0.6–2.2	0.5–3.6	–1.2–1.0	–2.5–6.2	–2.3–7.6	0.4–8.2
7	0.1–1.5	–0.2–1.0	–5.2– –0.7	–3.0– –0.5	0.2–2.7	–0.4–2.7
8	0.0–1.1	–0.3–1.8	–1.7–2.2	–2.2–2.9	–3.3–2.5	–4.1–2.8
9	–0.1–0.9	–1.0–1.2	–0.4–1.1	–5.5–0.9	–3.3–2.0	0.2–5.1
10	0.2–1.6	–1.0–4.1	–1.3–1.0	–4.9–2.1	–1.2–2.5	–3.5–2.5
Small spots						
1	5.5–28.5	–8.2–21.0	–10.1–16.0	–14.8–21.6	–7.6–8.3	–25.3–19.9
2	–0.0–1.8	1.2–1.5	–2.5–0.7	–1.8–0.4	–2.7–0.8	–2.6– –1.7
3	–1.3–23.7	–5.7–6.6	–3.4–8.5	–1.8–2.8	–9.1–2.6	–14.0–6.2
4	1.5–14.9	–3.6–7.8	10.2–24.0	3.0–30.8	–13.0– –2.0	–11.8–3.3
5	–2.0–4.7	1.8–4.2	–0.4–11.6	–3.9–16.3	–8.0– –0.4	–8.8–0.2
6	–0.2–15.5	1.6–12.6	4.1–20.4	–0.5–24.7	–8.9– –0.3	–11.8–0.2
7	–2.0–8.5	–1.9–1.6	–6.9–17.3	–18.1–11.8	–3.5–2.0	2.7–7.4
8	0.0–2.4	–0.6–1.4	4.8–8.3	–1.7–3.2	–2.3–3.7	0.6–10.0
9	–0.2–8.5	–12.8–10.3	1.7–7.0	–3.8–10.6	–8.8– –0.6	–13.3–15.7
10	8.3–9.4	–0.5–4.2	6.4–9.2	10.5–17.2	–0.7–0.0	–2.2– –3.4

The introduction of steeper dose gradients within the target volume using $\text{IMPT}_{\text{full}}$ leads to decreased target dose homogeneity in the presence of motion compared to $\text{IMPT}_{20\%}$, as shown by the homogeneity (D_1 – D_{99}) values. In their study, Widesott *et al.*²⁰ concluded that the smallest possible spot size should be used in IMPT to achieve similar lateral penumbræ compared to the most advanced photon techniques. This recommendation was based on the results from pencil beam calculations of individual head and neck, prostate and mesothelioma patients, and neglected any motion effects. As shown in this and previous studies,^{4,5} the susceptibility to motion effects is heightened when a smaller spot size is used, implying it may not be the best option for lung treatments.

Higher average V_{95} [Figs 3(a) and 3(c)] and D_1 – D_{99} [Figs 4(a) and 4(c)] values are observed for some patients for the 4D dose distributions compared to the 3D distributions. This results from small tumor motion potentially smearing hot and cold spots present in the 3D dose distributions.

The consistency in the MLD and lung $V_{90\%}$ values implies that the majority of the motion effects observed here are resultant from interplay and not due to dose blurring or other phenomena typically found in double scattering proton therapy or photon deliveries. Small rises in the lung doses are due to the tumor movement out of the beam, consequently resulting in increased lung exposure.

The larger spots show smaller variations in the considered dose metrics when moving from 4D- $\text{IMPT}_{20\%}$ to 4D- $\text{IMPT}_{\text{full}}$. This is likely due to the larger lateral penum-

brae, which means that the dose gradients introduced into the target volume through decreased SFH will not be as sharp as those encountered for the 4D- $\text{IMPT}_{\text{full}}$ plans which use the smaller spot size.

For the patients that show the most interplay, the variation in EUD between the different starting phases [i.e., the width of the error bars in Figs. 2(b) and 2(d)] is higher for 4D- $\text{IMPT}_{\text{full}}$ compared to 4D- $\text{IMPT}_{20\%}$ for both fraction and spot sizes, in some cases by $\sim 5\%$. This is primarily due to the presence of higher weighted spots in the $\text{IMPT}_{\text{full}}$ plans, most of which are found in the distal layers that are delivered first during the irradiation and hence have a strong dependence on the starting breathing phase. Previous work⁴ demonstrated the significance of the starting breathing phase. This is further highlighted by the error bar widths in Figs 2–4, which show the possible differences in the considered metrics based on the initial breathing phase. This provides motivation for including respiratory monitoring to commence treatment delivery at the most advantageous initial breathing phase or to include some level of robustness against initial breathing phase during treatment planning optimization.

The results of this study show that 4D- $\text{IMPT}_{\text{full}}$ and 4D- $\text{IMPT}_{20\%}$ lung treatments experience similar levels of motion effects for the patients and delivery conditions considered. It should be noted though, that $\text{IMPT}_{\text{full}}$ would be less robust to setup uncertainties, due to the steeper in-target dose gradients, which is not accounted for in this work. Techniques such as probabilistic treatment planning²¹ could be

TABLE III. Comparison of motion degradation (3D-4D) and 4D absolute values for IMPT_{full} and IMPT_{20%} for the ten patients for large (LS) and small (SS) spot sizes. The 3D-4D values represent the number of times the IMPT_{full} deliveries experience less motion degradation than IMPT_{20%} for the eight starting breathing phases for each patient. The 4D values are the number of times the 4D-IMPT_{full} value is higher than the corresponding 4D-IMPT_{20%} value for the eight starting breathing phases. The values shown are for 2.5Gy(RBE) fractions with the 12Gy(RBE) in parentheses. The “total” columns are the sum of the four different treatment scenarios for each patient. The “all” values are the summation of each treatment technique over the ten patients. A final value showing the 320 different deliveries is also included. The values in square brackets represent the possible maximum value for each variable.

(3D-IMPT _{full} -4D-IMPT _{full})-(3D-IMPT _{20%} -4D-IMPT _{20%})									
EUD			V ₉₅			D ₁ -D ₉₉			
Patient	LS [8]	SS [8]	Total [/32]	LS [8]	SS [8]	Total [/32]	LS [8]	SS [8]	Total [/32]
1	3 (6)	8 (6)	23	3 (5)	6 (6)	20	5 (5)	4 (2)	16
2	8 (8)	7 (8)	31	5 (3)	3 (1)	12	7 (2)	0 (0)	9
3	4 (4)	7 (4)	19	3 (3)	5 (5)	16	5 (4)	3 (2)	14
4	7 (6)	8 (6)	27	5 (3)	8 (8)	24	6 (6)	0 (2)	14
5	8 (8)	4 (8)	30	6 (6)	7 (5)	24	5 (3)	0 (2)	10
6	8 (8)	7 (8)	31	2 (3)	8 (7)	20	6 (8)	0 (1)	15
7	8 (4)	7 (4)	23	0 (0)	5 (4)	9	8 (7)	5 (8)	28
8	8 (7)	8 (4)	27	3 (4)	8 (5)	20	6 (2)	3 (8)	19
9	6 (5)	7 (5)	23	2 (2)	8 (6)	18	5 (8)	0 (5)	18
10	8 (7)	8 (7)	30	3 (3)	8 (8)	22	7 (3)	3 (1)	14
All [/80]	68 (63)	71 (60)	264 [/320]	31 (32)	64 (55)	185 [/320]	60 (48)	18 (31)	157 [/320]

4D-IMPT _{full} -4D-IMPT _{20%}									
EUD			V ₉₅			D ₁ -D ₉₉			
Patient	LS [8]	SS [8]	Total [/32]	LS [8]	SS [8]	Total [/32]	LS [8]	SS [8]	Total [/32]
1	2 (5)	5 (6)	18	3 (5)	6 (6)	20	2 (2)	2 (3)	9
2	7 (6)	8 (8)	31	7 (6)	8 (8)	29	0 (5)	0 (0)	5
3	6 (6)	4 (3)	19	4 (4)	8 (8)	24	0 (1)	3 (5)	9
4	7 (4)	7 (6)	24	6 (4)	8 (8)	26	2 (1)	3 (5)	11
5	8 (8)	8 (8)	32	7 (8)	8 (8)	31	0 (1)	0 (0)	1
6	6 (3)	8 (8)	27	6 (4)	8 (8)	26	2 (1)	1 (4)	8
7	1 (0)	8 (4)	13	0 (1)	7 (5)	13	7 (4)	3 (0)	14
8	1 (2)	8 (7)	18	3 (4)	8 (6)	21	6 (7)	4 (0)	17
9	1 (1)	8 (5)	15	6 (3)	8 (6)	23	3 (2)	6 (4)	15
10	8 (5)	6 (7)	26	5 (5)	8 (8)	26	1 (5)	3 (5)	14
All [/80]	47 (40)	70 (62)	223 [/320]	47 (44)	77 (73)	239 [/320]	23 (29)	25 (28)	103 [/320]

introduced into treatment planning systems to combat this issue as IMPT technology continues to become more widely available.

All our simulations and results assume that the patient has a constant breathing period of 5 s. Previous work⁴ has shown that the target dose homogeneity is correlated with breathing period. However, these still only assumed constant breathing periods over the treatment delivery. In a clinical treatment, which for current clinical proton therapy systems will cover multiple patient respiratory cycles, likely of varying period, any interplay and motion effects will be reduced compared to the results presented here due to increased averaging effects. The use of a single 4DCT implies that variations in the tumor trajectory over multiple breathing cycles or fractions are not accounted for here.

Baseline shifts were not considered in this study, neither were daily variations in tumor trajectory or volume changes over the treatment course. Repeat 4DCT and active respiratory and tumor trajectory monitoring could be used to ameliorate these issues, but were not possible with this patient

cohort. Motion mitigation techniques such as breath-hold, respiratory gating, rescanning, or tumor tracking were also not considered. Previous work⁵ suggests that for the majority of lung patients, such techniques may not be required depending on the delivery capabilities of the proton facility. Including motion mitigation techniques could also mask differences between IMPT_{full} and IMPT_{20%} deliveries. We chose to focus on characterizing the extent of the differences in the motion effects between the IMPT modalities, rather than specifically mitigating the motion effects, which has been investigated in several previous studies.^{9,10,18}

We do not make the step here to conclude that either of these IMPT delivery techniques is necessarily clinically viable or better suited than treatment methods currently in use. In particular, for the patients with large (>20 mm) motion, it is highly unlikely that they would be treated without respiratory gating. The aim of this study was to compare and determine the degradation due to motion of IMPT treatments with different maximum allowed SFH, not to necessarily establish their clinical viability.

Despite the different optimization strategies utilized, IMPT_{full} and IMPT_{20%} treatments show close agreement in absolute values of EUD, V_{95} , and D_1 - D_{99} for the majority of the cases investigated. The majority of the EUD (69.7%) and V_{95} (74.7%) results are improved in 4D-IMPT_{full}, with 4D-IMPT_{20%} more often showing improved target dose homogeneity (67.8%). For this set of delivery parameters and patient population, IMPT_{full} reduces motion effects; however, the differences compared to IMPT_{20%} are most likely not clinically significant.

This study represents the first in-depth comparison of IMPT treatments with varying SFH for a series of lung patients. 4D Monte Carlo simulations clearly show that relative motion between the tumor and proton beam induces changes in the dose distribution. Comparison with 3D simulations on the planning CT shows that the degradation of the dose distribution due to motion effects is of the same order for IMPT_{full} and IMPT_{20%} treatments for 2.5Gy(RBE) and 12Gy(RBE) fraction sizes for both spot sizes considered.

ACKNOWLEDGMENTS

This project was supported by NIH/NCI Grant No. R01 CA111590.

^{a)}Author to whom correspondence should be addressed. Electronic mail: sdowdell@mgh.harvard.edu

¹M. Goitein and G. T. Y. Chen, "Beam scanning for heavy charged particle radiotherapy," *Med. Phys.* **10**(6), 831–840 (1983).

²A. J. Lomax *et al.*, "Intensity modulated proton therapy: A clinical example," *Med. Phys.* **28**(3), 317–324 (2001).

³T. Bortfeld, K. Jokivarsi, M. Goitein, J. Kung, and S. B. Jiang, "Effects of intra-fraction motion on IMRT dose delivery: Statistical analysis and simulation," *Phys. Med. Biol.* **47**(13), 2203–2220 (2002).

⁴S. Dowdell, C. Grassberger, G. C. Sharp, and H. Paganetti, "Interplay effects in proton scanning for lung: A 4D Monte Carlo study assessing the impact of tumor and beam delivery parameters," *Phys. Med. Biol.* **58**, 4137–4156 (2013).

⁵C. Grassberger *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.* **86**(2), 380–386 (2013).

⁶Y. Zhang, D. Boye, C. Tanner, A. J. Lomax, and A. Knopf, "Respiratory liver motion estimation and its effect on scanned proton beam therapy," *Phys. Med. Biol.* **57**(7), 1779–1795 (2012).

⁷J. Seco, D. Robertson, A. Trofimov, and H. Paganetti, "Breathing interplay effects during proton beam scanning: Simulation and statistical analysis," *Phys. Med. Biol.* **54**(14), N283–N294 (2009).

⁸C. Bert, S. O. Grözinger, and E. Rietzel, "Quantification of interplay effects of scanned particle beams and moving targets," *Phys. Med. Biol.* **53**(9), 2253–2265 (2008).

⁹C. Bert, A. Gemmel, N. Saito, and E. Rietzel, "Gated Irradiation With Scanned Particle Beams," *Int. J. Radiat. Oncol., Biol., Phys.* **73**(4), 1270–1275 (2009).

¹⁰C. Graeff, M. Durante, and C. Bert, "Motion mitigation in intensity modulated particle therapy by internal target volumes covering range changes," *Med. Phys.* **39**(10), 6004–6013 (2012).

¹¹H. Kooy *et al.*, "A case study in proton pencil-beam scanning delivery," *Int. J. Radiat. Oncol., Biol., Phys.* **76**(2), 624–630 (2010).

¹²Y. Kang *et al.*, "4D Proton treatment planning strategy for mobile lung tumors," *Int. J. Radiat. Oncol., Biol., Phys.* **67**(3), 906–914 (2007).

¹³W. Chen, D. Craft, T. M. Madden, K. Zhang, H. M. Kooy, and G. T. Herman, "A fast optimization algorithm for multicriteria intensity modulated proton therapy planning," *Med. Phys.* **37**(9), 4938–4945 (2010).

¹⁴J. Perl, J. Shin, J. Schumann, B. Faddegon, and H. Paganetti, "TOPAS: An innovative proton Monte Carlo platform for research and clinical applications," *Med. Phys.* **39**(11), 6818–6837 (2012).

¹⁵S. Dowdell *et al.*, "Monte Carlo study of the potential reduction in out-of-field dose using a patient-specific aperture in pencil beam scanning proton therapy," *Phys. Med. Biol.* **57**(10), 2829–2842 (2012).

¹⁶G. C. Sharp, M. Peroni, R. Li, J. Shackleford, and N. Kandasamy, "Evaluation of plastimatch B-spline registration on the EMPIRE10 data set," *Medical Image Analysis for the Clinic: A Grand Challenge, MICCAI 2010*, Beijing, China (2010), pp. 99–108.

¹⁷A. Niemierko, "Reporting and analyzing dose distributions: A concept of equivalent uniform dose," *Med. Phys.* **24**(1), 103–110 (1997).

¹⁸A.-C. Knopf, T. S. Hong, and A. Lomax, "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.* **56**(22), 7257–7271 (2011).

¹⁹K. M. Kraus, E. Heath, and U. Oelfke, "Dosimetric consequences of tumour motion due to respiration for a scanned proton beam," *Phys. Med. Biol.* **56**(20), 6563–6581 (2011).

²⁰L. Widesott, A. J. Lomax, and M. Schwarz, "Is there a single spot size and grid for intensity modulated proton therapy? Simulation of head and neck, prostate and mesothelioma cases," *Med. Phys.* **39**(3), 1298–1308 (2012).

²¹J. Unkelbach, T. Bortfeld, B. C. Martin, and M. Soukup, "Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning," *Med. Phys.* **36**(1), 149–163 (2009).