4D robust optimization including uncertainties in time structures can reduce the interplay effect in proton pencil beam scanning radiation therapy

Erik Engwall,[a] Albin Fredriksson, and Lars Glimelius
RaySearch Laboratories, Sveavägen 44, Stockholm SE-111 34, Sweden

(Received 27 February 2018; revised 4 July 2018; accepted for publication 4 July 2018; published 3 August 2018)

Purpose: Interplay effects in proton radiotherapy can create large distortions in the dose distribution and severely degrade the plan quality. Standard methods to mitigate these effects include abdominal compression, gating, and rescanning. We propose a new method to include the time structures of the delivery and organ motion in the framework of four-dimensional (4D) robust optimization to generate plans that are robust against interplay effects.

Methods: The method considers multiple scenarios reflecting the uncertainties in the delivery and in the organ motion. In each scenario, the pencil beam scanning spots are distributed to different phases of the breathing cycle according to each individual spot time stamp, and a partial beam dose is calculated for each phase. The partial beam doses are accumulated on a reference phase through deformable image registrations. Minimax optimization is performed to take all scenarios into account simultaneously. For simplicity, the uncertainties in this proof of concept study are limited to variations in the breathing pattern. The method is evaluated for three different nonsmall cell lung cancer patients and compared to plans using conventional 4D robust optimization both with and without rescanning. We assess the ability of the method to mitigate distortions from the interplay effect over multiple evaluation scenarios using 4D dose calculations. This interplay evaluation is performed in an experimentally validated framework, which is independent of the optimization in the plan generation step.

Results: For the three studied patients, 4D optimization including time structures is efficient, especially for large tumor motions, where rescanning of conventional 4D robustly optimized plans is not sufficient to mitigate the interplay effect. The most efficient approach of the new method is achieved when it is combined with rescanning. For the patient with the largest motion, the mean V95% is 99.2% and mean V107% is 3.65% for the best rescanned 4D plan optimized with time structure. This can be compared to conventional 4D optimized plans with mean V95% of 92.7% and mean V107% of 13.1%.

Conclusions: The current study shows the potential of reducing interplay effects in proton pencil beam scanning radiotherapy by incorporating organ motion and delivery characteristics in a 4D robust optimization. [https://doi.org/10.1002/mp.13094]

Key words: 4D optimization, interplay effect, motion mitigation, proton therapy, rescanning, robust optimization

1. INTRODUCTION

In active proton treatments, interference between the scanning beam and intrafractional motion can severely degrade the plan quality. This is commonly referred to as the interplay effect.[1,2,3,4] Despite this challenge, many proton centers plan to use active scanning in sites with moving targets, such as the lung and liver regions.[5] The reason is twofold: (a) many treatments benefit from the sharp dose gradients of proton therapy, providing a possibility to precisely treat the target while sparing healthy tissue, and (b) new proton centers tend to install dedicated active beam scanning rooms only.

Means of mitigating the interplay effect include abdominal compression, breath hold, gating, rescanning and, in the longer perspective, tumor tracking.[4] Abdominal compression and breath hold will suppress the motion itself, while the other techniques aim at providing satisfactory dose distributions also during motion. Tumor tracking[6,7] could potentially lead to a very conformal dose distribution, but more research and development is needed both on the software and hardware side before this method could enter clinical practice. Rescanning can provide efficient means to mitigate the interplay effects for moderate motion amplitudes by statistically averaging out the motion effects through division of each energy layer into a number of rescans.[8,9,10] In gating, the beam is only delivered during a limited part of the breathing cycle using a connection between the delivery system and motion monitoring equipment.[11,12,13] Both gating and rescanning have the drawback of prolonging the treatment time. This is also true for breath hold techniques. However, breath hold has been suggested to provide an efficient way to deliver

high-quality plans, as long as the plans can be delivered within a few breath holds.[14] Many of the mentioned techniques could be used in combination to provide a higher level of interplay effect mitigation.

In this paper, an alternative way to handle interplay effects is presented: we propose to include the time structure of the delivery and of the respiratory cycle in the framework of robust optimization to create treatment plans that are robust against interplay effects. Bernatowicz et al.[15] have previously presented a computationally efficient method for four-dimensional (4D) optimization including the time structures of delivery and intrafractional motion. As long as the time structures are identical between planning and delivery, excellent plans can be produced. However, the shortcoming of this method is that it is very sensitive to small deviations in the time structures between planning and delivery. Such deviations will arise in any real treatment situation and could potentially be handled by the use of robust optimization. Robust optimization taking setup and range uncertainties into account has previously been successfully implemented and used,[16,17,18,19,20] and has also been observed to provide robustness against intrafractional motion.[21,22] Moreover, intrafractional changes in the patient geometry have been included in the optimization (also known as 4D optimization).[23,24,25,26]

In the standard 4D optimization approach, the entire beam dose is computed and accumulated over the 4DCT images. This results in robust plans if the entire beam dose is delivered to each phase, as would be the case for scattered proton beams. However, the time structures are not taken into account and the plans are not robust against the interplay effect *per se*. To address this problem Graeff et al.[27] proposed a 4D optimization approach where the target was divided into different sectors assigned to different motion phases. Each sector was targeted by a set of spots, and the sets of spots were optimized simultaneously to achieve a homogeneous dose distribution. During delivery, the spots were subsequently sorted by the treatment control system (TCS) to the correct phase by active motion monitoring. Another proposed method combines 4D optimization with beam tracking,[28] which also requires active implementation in the TCS. In contrast to these suggestions, our approach aims at creating robust plans without the additional implementation in the TCS by including the time structures already in the optimization.

Our method is applied for three different nonsmall cell lung cancer (NSCLC) patients for plans both with and without rescanning. The resulting plans are compared to 4D robustly optimized plans (worst case optimization over all phases) with and without rescanning. NSCLC is one of the cancer types that has been suggested to be well-suited for proton therapy.[29,30,31] However, lung treatments also present the most complicated region for active proton treatments, not only because of the interplay effect but also due to the large density inhomogeneities, which are difficult from the point of view of plan robustness even for stationary targets[32,33,34] and accuracy of analytical dose calculation algorithms,[35,36] which until recently was the only option in the most common commercial proton treatment planning systems. In our study, the dose calculation employs a Monte Carlo method to fully account for the density heterogeneities in the lung region. All plans are created simulating delivery of an IBA Proteus Plus system (IBA, Louvain-La-Neuve, Belgium) equipped with the Dedicated PBS nozzle.

2. MATERIALS AND METHODS

2.A. Robust optimization with respect to time structure

Given a treatment plan, the time structure of the delivery, and the time spent in each phase of the breathing cycle (the phase occupancy), the partial beam doses delivered to each phase are calculated. Deformable image registration is used to accumulate all partial beam doses onto a reference phase, which results in total beam doses. Uncertainties in the phase occupancy are handled by robust optimization.

For brevity, the following exposition concerns a plan using a single beam. The extension to multiple beams is straightforward. Multiple beams are used in the experiments.

2.A.1. Time structure

The spot delivery times are provided by a connection to the IBA ScanAlgo system, which simulates the delivery sequence. The simulation tool provides accurate delivery characteristics that are used in the optimization process. Given ordered spot energies, positions, and weights, the interface returns the time structure—a vector t of time stamps at which the spots will be delivered.

In our experiments, the time structure is assumed to be constant while the optimization is running. Every 10 iterations, the optimization is paused, and the time structure is updated to reflect the current spot weights.

2.A.2. Phase occupancy

A specification of the time spent in each phase p in the set $\mathcal{P}$ of phases constitutes a scenario, denoted by s. The time spent in each phase is subject to uncertainty, and a treatment optimized to handle a single scenario can be highly sensitive to variations in the phase occupancy.[15] We, therefore, define and handle a set $\mathcal{S}$ of scenarios, spanning a range of different phase occupancies. Two different methods—method 1 and method 2—are used for generating the set $\mathcal{S}$.

For both methods, all phases have the same phase occupancy within each breathing cycle. In method 1, the breathing cycle time is constant within each scenario. Scenarios are generated where the breathing cycle time, and thus the time spent in each phase, has been scaled up and down by 3.75, 7.5, 11.25, and 15% compared to the mean breathing cycle time measured before treatment. In method 2, the breathing cycle time is varied between consecutive breathing cycles in each scenario. Scenarios are generated by sampling the breathing cycle length from a normal distribution with mean and variance from

the pretreatment measurements. For both methods, optimization is performed as if the delivery were the same in all fractions. This is a conservative approximation to refrain the optimization algorithm from trying to compensate underdosage in one fraction with overdosage in a fraction with a more favorable breathing pattern. Moreover, the starting phase is assumed to be fixed, and the nominal scenario, with a constant breathing cycle time according to the mean of the measured breathing cycle times, is included in the optimization.

2.A.3. Partial beam dose calculation

For a given time structure t and scenario s, the partial beam dose delivered to phase p is calculated as the dose from the spots that have delivery start times in phase p. Formally, let $T_p(s)$ be the union of time intervals in which the patient is in phase p under scenario s. In general, $T_p(s)$ is given by

$$T_p(s) = \bigcup_{k=1}^{\infty} [t_{p,k}^{\text{start}}(s), t_{p,k}^{\text{stop}}(s)],$$

where $t_{p,k}^{\text{start}}(s)$ and $t_{p,k}^{\text{stop}}(s)$ denote, respectively, the start and the stop time of phase p in breathing cycle k under scenario s.

The partial beam dose $d_p(x;s,t)$ delivered to phase p under scenario s and time structure t given the spot weight vector x is then calculated as

$$d_p(x;s,t) = D_p \tilde{x}(s,t), \quad \text{where}$$

$$\tilde{x}_i(s,t) = \begin{cases} x_i & \text{if } t_i \in T_p(s) \\ 0 & \text{otherwise} \end{cases}, \quad i = 1, \ldots, |x|,$$

where D_p denotes the dose influence matrix of phase p, and $|x|$ is the number of spots.

2.A.4. Total beam dose calculation

Deformable image registration is used to map the partial beam doses of all phases onto a reference phase. We use the ANAtomically CONstrained Deformation Algorithm (ANACONDA) of RayStation 7.[37] It is a hybrid approach utilizing a combination of image intensities and controlling structures from contoured image sets. For this study, no controlling structures are used.

The total beam dose under scenario s and time structure t for the spot weights x is then calculated as

$$d(x;s,t) = \sum_{p \in \mathcal{P}} R_p d_p(x;s,t),$$

where R_p is the deformation matrix — resulting from the deformable image registration — that maps the dose in phase p to the reference phase.

2.A.5. Robust optimization

Robust optimization taking the set $\mathcal{S}$ of scenarios into consideration is implemented by means of minimax optimization.[18] For n objectives $f_1,\ldots,f_n$ with associated importance weights $w_1,\ldots,w_n$ and m constraints $g_1,\ldots,g_m$, the optimization problem is formulated as

$$\begin{aligned} &\underset{x \in \mathcal{X}}{\text{minimize}} \quad \max_{s \in \mathcal{S}} \left\{ \sum_{i=1}^{n} w_i f_i(d(x;s,t)) \right\} \\ &\text{subject to} \quad g_i(d(x;s,t)) \le 0, \qquad i = 1,\ldots,m, \quad s \in \mathcal{S}, \end{aligned}$$

where $\mathcal{X}$ is the set of feasible spot weights.

2.B. Patients

Three NSCLC patients with different tumor sizes and motion amplitudes are selected for the study from The Cancer Imaging Archive (TCIA) 4D Lung database.[38,39] The patient dataset includes high-quality 4DCT images with targets and organs at risk (OARs) delineated by a single radiation oncologist. For treatment planning, a CT-to-density table based on table IV in Hugo et al.[39] was used. Audio-visual biofeedback was used for all patients to provide steady breathing patterns. Table I summarizes the breathing cycle lengths, as well as tumor characteristics of the selected patients.

The target motion amplitude presented in Table I was assessed through the mean of the deformation vector lengths (DVLs) of the deformable registration within the internal target volume (iCTV): $DVL_i = \sqrt{x_i^2 + y_i^2 + z_i^2}$, where x_i, y_i and z_i are the components in voxel i of the deformation vector field between maximum expiration and inspiration.

2.C. Treatment planning

Treatment plans were created in a research version of the treatment planning system RayStation 7. All plans had the

TABLE I. Patient and plan characteristics for the studied patients: the first five columns display tumour volume and motion in terms of the mean breathing amplitude and its standard deviation (in parenthesis), the mean CTV volume, *V*, averaged over the 10 different phases, the mean breathing cycle length, $\overline{T}$, and the standard deviation of the breathing cycle length, σ (in seconds and percent). The last two columns show the number of layers in the nominal 4D plan as well as in the rescanned 4D plan. The number of layers in the rescanned plan is dictated by a limit on the maximum weight of 1 MU of each layer. (Note that the number of layers in the 4D plans and IPR plans with the same rescanning setting are very close to each other, but could differ by a few layers.). Resc., Rescanning

Patient	Mean (std) ampl. [cm]	Mean CTV V[cm^3]	$\overline{T}$ [s]	σ [s]	σ [%]	# layers	
						No resc.	Resc.
P111	1.22 (0.18)	73.8	3.2	0.16	4.9	64	625
P104	0.60 (0.13)	44.3	3.5	0.25	7.1	80	676
P115	0.37 (0.10)	6.5	6.6	0.30	4.6	38	169

same beam configuration with three beams each, arranged to avoid traversing critical structures and having critical structures directly behind the target. The separation between the beams was in general 30°. 4D robust optimization (in which the worst case objective value over the phases is minimized) was performed with the CTV as a target in all phases. No additional margins were applied and no uncertainties for range or setup errors were taken into consideration in the robust optimization. While not resulting in clinically realizable plans, this will limit the problem to intrafractional motion only, making the evaluation of the interplay effect cleaner. For each patient, a number of different plans were created for comparison:

1. a standard 4D robustly optimized plan,
2. a rescanned version of plan (1),
3. a 4D robustly optimized plan with time structures according to method 1, described in Section 2.A.2, with 9 scenarios,
4. a rescanned version of plan (3),
5. two 4D robustly optimized plans with time structures according to method 2, described in Section 2.A.2, with 10 and 40 scenarios, respectively, and
6. rescanned versions of the plans in (5).

Hereinafter, we refer to the 4D robustly optimized plans ((3)–(6)) with respect to time structure as Inter Play Robustness (IPR) plans. Standard 4D robustly optimized plans are simply referred to as 4D plans.

The treatments were planned for 30 fractions and the optimization functions were identical for all patients and plans: for the CTV there was a minimum dose constraint of 60 Gy_{RBE} and a maximum dose objective of 60 Gy_{RBE} (weight 60), and for the External a maximum dose of 63 Gy_{RBE} (weight 4), as well as a dose fall-off from 60 Gy_{RBE} to 10 Gy_{RBE} over 1 cm (weight 10). For simplicity, in this proof of concept study, no OAR other than the External was included in the optimization. The interplay evaluation was, therefore, also made with respect to the dose in the CTV only. The dose engine used in the optimization was the RayStation Monte Carlo dose engine.[40,41] All final doses were calculated to 0.5% statistical uncertainty and the dose grid resolution was $3 \times 3 \times 3$ mm^3.

2.D. Rescanning

The rescanning strategy used for the rescanned plans is layered rescanning, where each energy layer is rescanned a number of times before switching to next energy. The number of rescans for each energy layer is dictated by a limit on each delivered layer weight of maximum 1 MU. This will create a larger number of rescans for the energy layers with higher weight. The last two columns of Table I show the number of layers for the 4D plans with and without rescanning. The minimum spot weight of the treatment machine of 0.02 MU was included both in the 4D optimization and in the creation of rescanned layers, to ensure that all spots are directly deliverable. In the optimization process this is achieved through spot filtering after a preset number of iterations, when spots below the minimum MU are removed. For the standard 4D plans the spot filtering is performed after iteration 40. The maximum number of iterations is set to 200. For the IPR plans the spot filtering process is performed in steps, since the redistribution of spots over phases after the time structure updates every 10 iterations can significantly change the dose distribution. Therefore, spots with weight below the minimum spot weight before the time structure update might be needed after the time structure update to ensure target coverage. The stepwise filtering is performed according to the following: after 40 iterations, spots with weight below 25% of the minimum spot weight are removed. The next removal is made after 40 more iterations for spots with weight below 50% of the minimum spot weight, and finally spots with weights below the minimum spot weight are removed after yet 40 more iterations. Such a stepwise filtering could in some cases also be needed for standard plans, but for the studied patients and plans the differences between unfiltered and filtered plans are minimal, and therefore we just employed common filtering with one removal after iteration 40 as described above.

For the 4D plans, the rescanning is applied as a postprocessing step after the optimization. In the absence of motion, this results in the same dose distribution for the 4D plan with and without rescanning. However, for the IPR plans, such an approach would deteriorate the plan quality, since the time structure of delivery will be completely changed when rescanning is introduced. Therefore, the creation of rescanned IPR plans needs to be included in the optimization with the corresponding unfiltered 4D optimized rescanned plans used as starting point. The time structure of the delivery is applied to each individual spot in all rescans and the subsequent optimization is performed in the same way as the standard IPR plans with stepwise spot filtering and time structure update every 10 iterations. This means that the spot weights in the rescans of one energy level are allowed to change independently of each other, and the spot weights in the rescans of each energy will not be multiples of the original spot weights as in a traditional rescanned plan. The number of layers of the IPR plans for a chosen rescanning setting are very close to the corresponding 4D plan (as shown in Table I). Small differences of a few layers could arise, since spot weights will vary differently in the two optimization methods and spot filtering can result in the removal of complete energy layers.

2.E. Interplay evaluation

The interplay effect is evaluated in a separate framework from the robust optimization, but based on the same principles using a 4D dose calculation. It is implemented in the scripting environment of RayStation 7 and has been experimentally verified by Pfeiler et al.[42] The method distributes the spots over the different breathing cycle phases, computes the partial beam doses, and maps the doses back to a

reference phase utilizing deformable registrations. The organ motion is described by the phases in the 4DCT and by different realistic breathing signals in each evaluation scenario. These breathing signals are generated with varying consecutive breathing cycle lengths through sampling from a normal distribution with mean and variance taken from the breathing signal obtained in conjunction with CT acquisition before treatment (see Table I). The interplay effect is evaluated over 20 such evaluation scenarios that are independent from the scenarios used in the optimization.

The evaluation is made without any averaging over the treatment fractions, and the evaluation results reflect the delivery as if all fractions were delivered in the same way.

3. RESULTS

Figure 1 gives nominal DVHs for the 4D plans and the IPR plans (with and without rescanning) according to method 1. It should be noted that the nominal DVHs are computed differently for the 4D plans with and without time structure. For the standard 4D plan, the DVH shows the dose from the nominal scenario, which is the dose delivered to the reference phase, whereas for the IPR plans, the DVH shows the dose resulting from an accumulation of partial doses that have been deformed from the different phases to the reference phase. The nominal DVHs serve as a baseline for the quality of the plans in the subsequent interplay evaluation. For the IPR plans, the results from interplay evaluation with 0% variation in breathing cycle are also displayed in Fig. 1 as dashed lines, showing that with the same time structure as in one of the robust optimization scenarios, the nominal DVHs are well reproduced.

The interplay effect is evaluated as described in Section 2.D for the different plans and patients and the results are displayed in Fig. 2. The displayed IPR plans are optimized according to method 1 (upper panel), and method 2 with 10 different scenarios (lower panel). For patients P111 and P104, the interplay effect gives rise to severe cold spots for the 4D plans both with and without rescanning. The IPR plans overcome these problems and the results are especially good for the plan with rescanning, where the spread of the DVHs is small and the maximum dose is lower. For patient P115, the differences to the 4D plans are not as expressed as for patients P111 and P104. However, the rescanned IPR plan is very close to the corresponding original plan. It could also be noted that the rescanned 4D plan is less robust to interplay effects than the normal 4D plan. This is not the expected result for rescanned plans, but could occur for certain plan configurations and motion patterns in combination with realistic energy switching times (1–2 s), resulting in a disadvantageous distribution of energy layers.

Comparing method 1 and 2, there is a slight improvement for the IPR plans generated according to method 2 compared to method 1. This can also be observed in Fig. 3, which summarizes the evaluations over the different methods displaying dose metrics (D5, D50, D95, and the homogeneity index HI, here defined as D5-D95). The performance over all dose metrics considered simultaneously is in general slightly better for method 2 than for method 1. Moreover, rescanned IPR plans are in general better than normal IPR plans with respect to D5, D50 and homogeneity, but worse with respect to target coverage (D95). This can also be seen in the nominal DVHs for the IPR plans. One explanation is that in the rescanned IPR plans there are more spots with low weights with risk of being filtered out. The low weight spots do not contribute significantly to the overall dose distribution, but can be important for complete target coverage. One way to leverage the improved homogeneity of the rescanned plans without losing target coverage would be to set the dose level of the

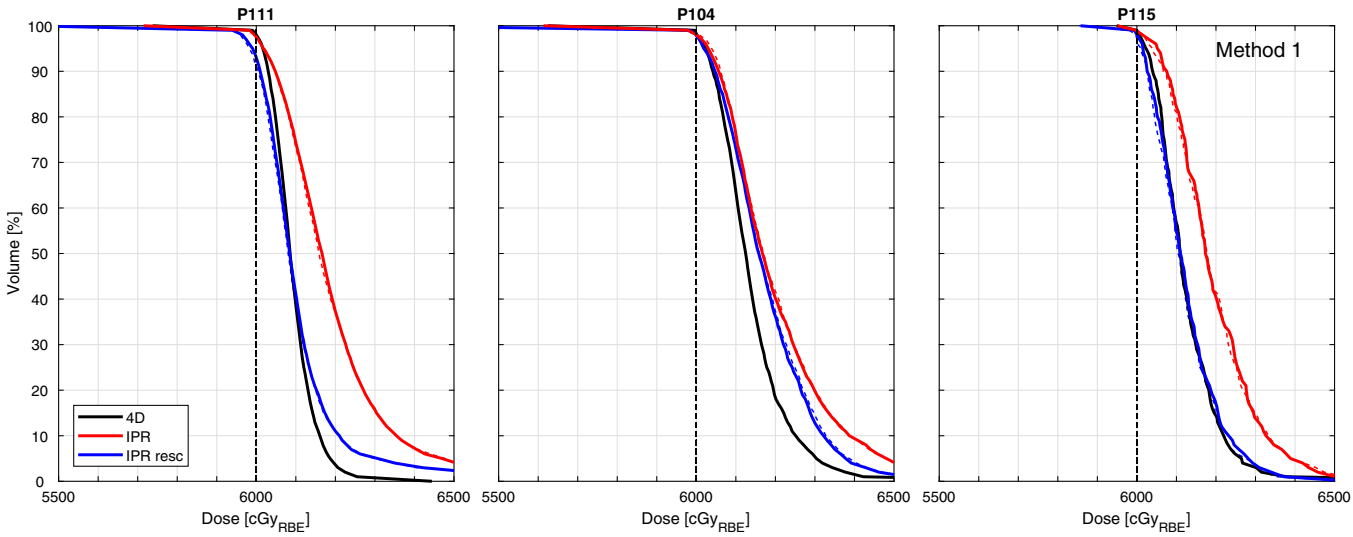

FIG. 1. Nominal DVHs for the 4D plans (solid black), the IPR plans (solid red), and the rescanned IPR plans (solid blue) for patients P111, P104, and P115. The IPR plans are created according to method 1. For the IPR plans the interplay evaluated plans with 0% variation in breathing cycle are shown for comparison (dashed red and blue lines). The vertical dashed line represents the minimum dose constraint on the target of 60 Gy_{RBE}, used in the optimization of all plans.

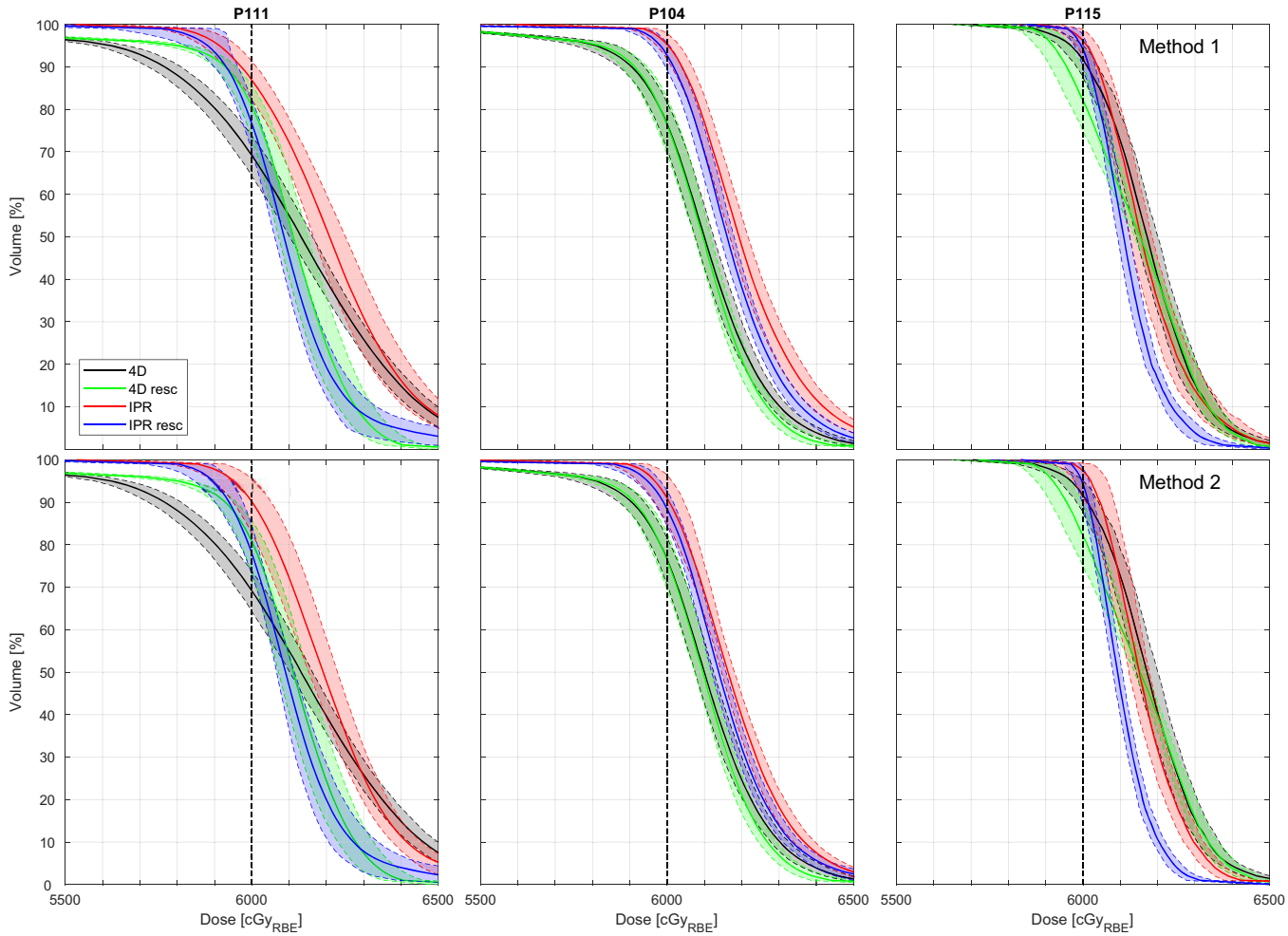

FIG. 2. Interplay evaluation of 4D plans without (black) and with rescanning (green), and IPR plans without (red) and with rescanning (blue) according to method 1 (upper panel) and method 2 with 10 optimization scenarios (lower panel). All plans are evaluated over 20 different scenarios with randomly generated realistic breathing cycle signals based on the pretreatment measured mean breathing cycle length and standard deviation. The solid lines represent the mean DVHs and the area around is bounded by $\pm 1\sigma$.

target optimization function slightly higher. Another alternative could be to construct a more elaborate filtering method where low-weighted spots instead of being removed are redistributed to a layer with the same nominal energy. Furthermore, it can be noted that increasing the number of scenarios to 40 in method 2 leads to no evident improvement for the studied patients.

The highest benefit of time structure robustness is seen for patient P111, which by far has the highest tumor motion amplitude, with an increase in mean D95 of 6.4% (from 56.2 Gy_{RBE} to 59.8 Gy_{RBE}) for the best IPR plan compared to the standard 4D plan and for the homogeneity index, the maximum mean value reduction is 50% (from 9.3 Gy_{RBE} to 4.7 Gy_{RBE}). In terms of volumes receiving underdosage and overdosage (expressed as V95% and V107%, respectively), mean V95% is increased from 92.7% to 99.2% and mean V107% is decreased from 13.1% to 3.65% for the best rescanned IPR plan as compared to the standard 4D plan.

With respect to dose to normal tissue the IPR plans perform on equal standards or slightly better than 4D plans. Also in this context the IPR plans with rescanning produce the highest quality plans. Table II displays mean nominal doses to the external, heart, esophagus, and the lung containing the tumor for the 4D plans and the rescanned IPR plan according to method 2 (40 scenarios). The doses to OARs could probably be further reduced by applying appropriate objectives in the optimization.

4. DISCUSSION

We have introduced a new method aimed at being robust against interplay effects in actively scanned proton treatments. Uncertainties in the delivery and organ motion are handled by the inclusion of multiple scenarios in a 4D robust optimization. The number of scenarios in such an optimization could potentially be very high and the method has to be combined with means to reduce the uncertainties. The rationales for reducing the number of scenarios are to achieve reasonable computational times and even more importantly to restrict the dose delivered to normal tissue. The optimizer will strive to achieve target coverage in all scenarios according to the robust optimization functions. A large number of

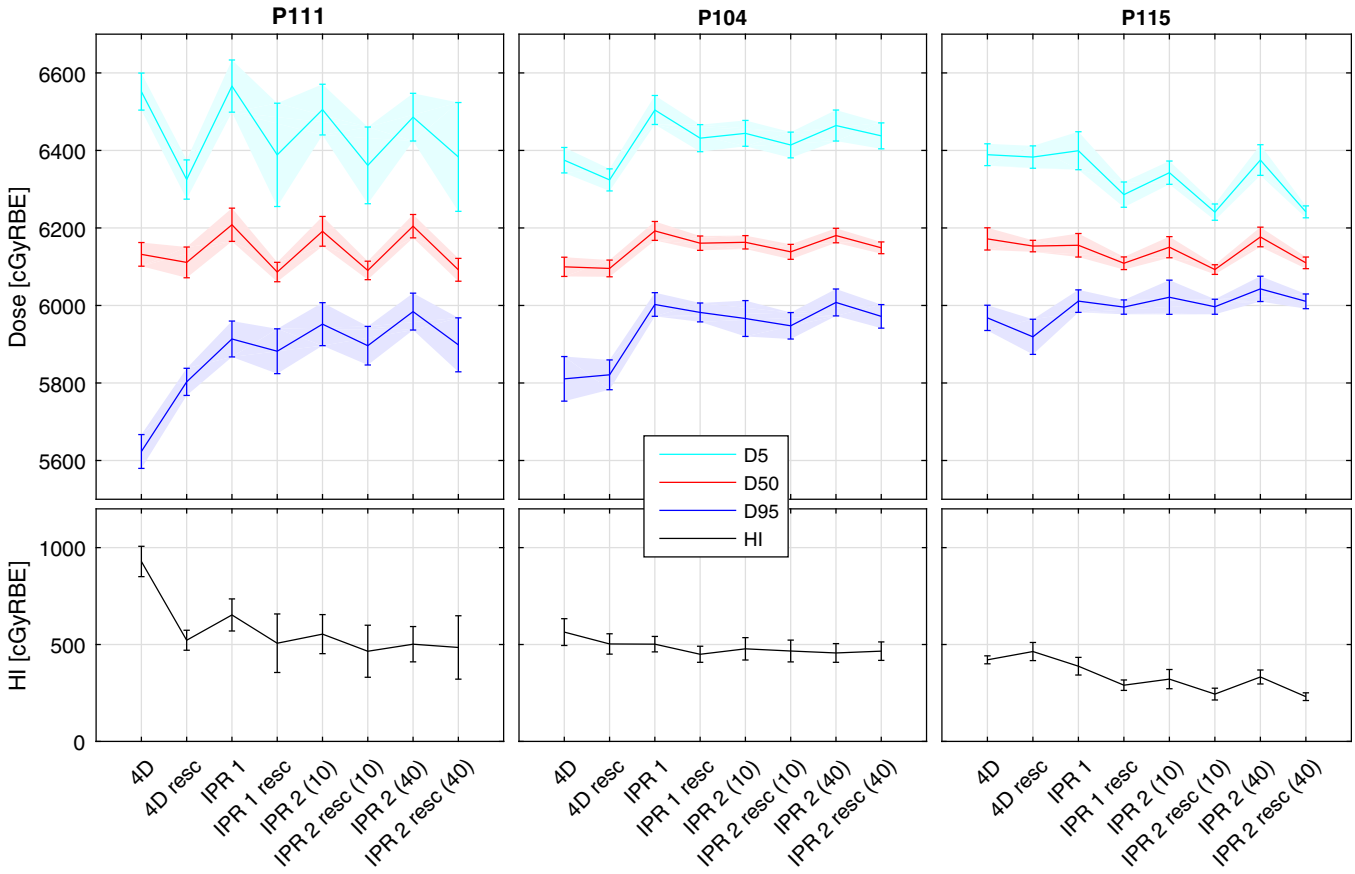

FIG. 3. Comparison of dose metrics (D5, D50, and D95 (upper panel), and homogeneity index, HI ≡ D5-D95 (lower panel)) for interplay evaluation of different robust optimized plans. [4D = 4D plan; IPR 1 = IPR optimized according to method 1; IPR 2 (10) = IPR optimized according to method 2 over 10 scenarios; IPR 2 (40) = IPR optimized according to method 2 over 40 scenarios; rescanned plans are denoted with the suffix "resc"]. All plans are evaluated over 20 different scenarios with randomly generated realistic breathing cycle signals. The dose metrics are shown with mean value and error bands based on the standard deviation.

TABLE II. Mean nominal doses to the OARs for the 4D plan and the rescanned IPR plan according to method 2 (40 scenarios). The lung dose refers to the mean dose to the lung containing the tumor

	Mean nominal doses [cGY_{RBE}]					
	P111		P104		P115	
	4D plan	IPR resc	4D plan	IPR resc	4D plan	IPR resc
External	190	174	208	176	53	45
Heart	152	142	245	209	0	0
Esophagus	534	436	769	713	8	6
Lung	1290	1190	162	122	234	208

scenarios will inherently increase the integral dose compared to an ideal plan with no uncertainties taken into account and is comparable to increasing target margins in conventional radiotherapy planning. In the current study, we have limited the uncertainties to the breathing pattern, and assumed that the other parameters are constant. For example, even though the current implementation allows for variation in the phase in which the delivery starts, we assume that the delivery is synchronized with the patient's breathing and starts in a specific phase. Given that motion monitoring equipment without delay in signal processing is available at the clinic, the start of delivery is easier to control than uncertainties in the breathing pattern. On the delivery machine side, it is highly desirable that the time structure is stable, so that any uncertainties related to the delivery machine could be neglected. For this sake, it is of importance to have a realistic delivery simulation of the machine reflecting the actual delivery, for example, from the ScanAlgo simulation tool or from experimentally determined time structures.[42,43] If an interlock occurs in the machine, the delivery should be resumed in the same position in the breathing cycle to ensure that robustness is maintained.

Furthermore, range and setup errors must be included in the robust optimization, if the method is to be used in a clinical setting. This is especially important in the inhomogeneous lung region, where a small shift of the densities in the patient relative the beam could induce large shifts in range. In the IPR method, this would translate to time-domain effects through changes in spot weights to account for the range shifts in different scenarios.

Even if the uncertainties in the breathing pattern are included in the robust optimization, the use of audio-visual

biofeedback or other types of breath coaching is recommended to limit these variations and thus the number of scenarios needed. If large changes in the tumor motion occur during treatment, for example, phase shifts induced by the patient coughing, the delivery might need to be interrupted and restarted in the correct phase.

Other potential problems with IPR relate to the quality of the 4DCT dataset, as well as differences in the patient geometry and organ motion between planning and delivery. These uncertainties might in some cases require being taken into account in the robust optimization. Base line shifts, where the mid-position of the target is shifted between planning and delivery, could serve as one example. Another related problem is a changing patient geometry during the course of treatment fractions, for example, tumor shrinkage or weight loss of the patient. If the changes are substantial, this would require adaptive replanning, in the same way as in the static case. To assess the importance of the interplay effect in general and understand when replanning is needed in particular, it is important to evaluate the actual delivered dose through 4D dose calculation after each fraction with the use of actual log files from the treatment machine.

The methods for both robust optimization and interplay evaluation rely on correct deformable image registrations. The quality of the deformable registration is especially important for the robust optimization, since errors in the deformation fields could drive the optimizer in the wrong direction. Future studies should investigate sensitivity with respect to changing patient geometry and quality of 4DCT, as well as errors in deformable image registrations between the different phases in the breathing cycle.

The rescanning method used in this study is standard layered rescanning, where the rescans of each energy layer are delivered directly after each other. Layered rescanning with splitting of each original energy layer into a preset number of rescans is a standard method used at clinics.[44] Therefore, we employed layer rescanning in this study, but with a slightly more elaborate method for splitting the energy layers: the weight of each layer is limited by a maximum MU, resulting in more rescans for high-weighted layers. Even with a more advanced splitting method, the main problem of layered rescanning remains with the rescans being delivered in a limited part of the breathing cycle. Volumetric rescanning or other more elaborate rescanning regimes, such as phase-controlled or breath sampled layered rescanning,[9,45,46,47,48] where the rescans are distributed throughout the respiration period, have the potential to reduce the interplay effects to a much larger extent than standard layered rescanning, and future studies should include several rescanning techniques for intercomparison.

In addition to rescanning, IPR could also be combined with gating, breath hold, and abdominal compression. The combination of IPR and gating could be used to reduce the number of scenarios, leading to just a few of the breathing phases being included in the delivery and thus in the optimization. However, the method could also be seen as an alternative to gating, providing a homogeneous dose distribution even under motion, but without resulting in the prolonged treatment time of gated treatments. Breath hold and abdominal compression would just as audio-visual biofeedback limit the motion uncertainties. Complementary ways to increase plan quality under motion is to make educated choices of planning parameters, such as beam angles[44] and spot size.[49,50]

Averaging effects over fractions are not taken into account in the interplay evaluation. For moderate motion and normally fractionated treatments, these effects could mitigate the interplay effects to a fairly large extent.[21,51] However, the errors may not always be brought down to a clinically acceptable level by the effect of averaging over multiple fractions.[1,3] For example, for patient P111, who exhibits the largest motion amplitude, it is evident from the different evaluation scenarios that fractionation would not achieve sufficient target coverage for the standard 4D plans (both with and without rescanning). In addition, for hypo-fractionated treatments, which is considered for NSCLC,[52] it is necessary to achieve a homogeneous dose distribution in each fraction. Furthermore, the combination of low doses in parts of the target for some of the fractions and high doses in other fractions, might not result in the same biological effect as more homogeneous dose distributions over all fractions, even if the total doses are the same in the two cases.[2,53] To fully understand this effect, biological models related to TCP and NTCP need to be included in the evaluation, and such an analysis is beyond the scope of this paper. A conservative approach with respect to biological effects is to implement motion-mitigating techniques resulting in small deviations between planned and delivered dose distributions in each fraction. We have, therefore, chosen to evaluate the interplay effect for one single fraction, in the same way as, for example, Zenklusen et al.,[10] Knopf et al.[44] and Bernatowicz et al.[54]

5. CONCLUSIONS

The current study serves as a proof of concept for the possibility to reduce interplay effects in active proton treatments by the use of 4D robust optimization including time structures. The conclusions can be summarized as follows:

1. Inclusion of the time of delivery and breathing motion in the 4D robust optimization achieves time structure robustness that has proven to have the potential to overcome interplay effects. In the present study, this was especially pronounced for the patient with largest tumor motion, where the investigated layered rescanning strategy alone could not mitigate the interplay effects.
2. Time structure robustness can be combined with rescanning. This combination provides the most efficient plans with respect to interplay effect mitigation.
3. The time structure robust plans can be optimized to be robust for moderate changes in breathing cycle pattern. The best robustness is achieved when the scenarios are based on realistic breathing cycle patterns with varying consecutive breathing cycle lengths.

4. Even if all uncertainties related to the delivery dynamics and the organ motion in theory could be included in the robust optimization, the number of scenarios needs to be reduced in practice. The current study was limited to variations in the breathing pattern, but sensitivity to other variations, such as changes in patient geometry between planning and delivery, should also be investigated in future work.

ACKNOWLEDGMENTS

The Cancer Imaging Archive is acknowledged for making high-quality 4DCT data of NSCLC patients available to the community. Geoffrey Hugo at Washington University School of Medicine in St. Louis is thanked for providing additional information on the dataset, including corresponding breathing signals. The access to the ScanAlgo system was kindly arranged by Bartek Wlodarczyk, Sofie Gillis, and the beam data team at IBA.

CONFLICT OF INTEREST

The authors have no conflicts to disclose.

[a]Author to whom correspondence should be addressed. Electronic mail: erik.engwall@raysearchlabs.com.

REFERENCES

1. Lambert J, Suchowerska N, McKenzie DR, Jackson M. Intrafractional motion during proton beam scanning. *Phys Med Biol*. 2005;50:4853.
2. Seco J, Sharp GC, Turcotte J, Gierga D, Bortfeld T, Paganetti H. Effects of organ motion on IMRT treatments with segments of few monitor units. *Med Phys*. 2007;34:923–934.
3. Bert C, Grözinger SO, Rietzel E. Quantification of interplay effects of scanned particle beams and moving targets. *Phys Med Biol*. 2008;53:2253.
4. Bert C, Durante M. Motion in radiotherapy: particle therapy. *Phys Med Biol*. 2011;56:R113.
5. Knopf A-C, Stutzer K, Richter C, et al. Required transition from research to clinical application: Report on the 4D treatment planning workshops 2014 and 2015. *Phys Med*. 2016;32:874–882.
6. Bert C, Saito N, Schmidt A, Chaudhri N, Schardt D, Rietzel E. Target motion tracking with a scanned particle beam. *Med Phys*. 2007;34:4768–4771.
7. Zhang Y, Knopf A, Tanner C, Lomax AJ. Online image guided tumour tracking with scanned proton beams: a comprehensive simulation study. *Phys Med Biol*. 2014;59:7793.
8. Phillips MH, Pedroni E, Blattmann H, Boehringer T, Coray A, Scheib S. Effects of respiratory motion on dose uniformity with a charged particle scanning method. *Phys Med Biol*. 1992;37:223.
9. Seco J, Robertson D, Trofimov A, Paganetti H. Breathing interplay effects during proton beam scanning: simulation and statistical analysis. *Phys Med Biol*. 2009;54:N283.
10. Zenklusen SM, Pedroni E, 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.
11. Ohara K, Okumura T, Akisada M, et al. Irradiation synchronized with respiration gate. *Int J Radiat Oncol Biol Phys*. 1989;17:853–857.
12. Hashimoto T, Tokuuye K, Fukumitsu N, et al. Repeated proton beam therapy for hepatocellular carcinoma. *Int J Radiat Oncol Biol Phys*. 2006;65:196–202.
13. Iwata H, Murakami M, Demizu Y, et al. High-dose proton therapy and carbon-ion therapy for stage I nonsmall cell lung cancer. *Cancer*. 2010;116:2476–2485.
14. Gorgisyan J, Rosenschold PM, Perrin R, et al. Feasibility of pencil beam scanned intensity modulated proton therapy in breath-hold for locally advanced non small cell lung cancer. *Int J Radiat Oncol Biol Phys*. 2017;99:1121.
15. Bernatowicz K, Zhang Y, Perrin R, Weber DC, Lomax AJ. Advanced treatment planning using direct 4D optimisation for pencil-beam scanned particle therapy. *Phys Med Biol*. 2017;62:6595.
16. Pflugfelder D, Wilkens J, Oelfke U. Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy. *Phys Med Biol*. 2008;53:1689–1700.
17. Unkelbach J, Bortfeld T, Martin BC, Soukup M. Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning. *Med Phys*. 2009;36:149–163.
18. Fredriksson A, Forsgren A, Hårdemark B. Minimax optimization for handling range and setup uncertainties in proton therapy. *Med Phys*. 2011;38:1672–1684.
19. Liu W, Zhang X, Li Y, Mohan R. Robust optimization of intensity modulated proton therapy. *Med Phys*. 2012;39:1079–1091.
20. Stuschke M, Kaiser A, Pöttgen C, Lübcke W, Farr J. Potentials of robust intensity modulated scanning proton plans for locally advanced lung cancer in comparison to intensity modulated photon plans. *Radiother Oncol*. 2012;104:45–51.
21. Inoue T, Widder J, Dijk LV, et al. Limited impact of setup and range uncertainties, breathing motion, and interplay effects in robustly optimized intensity modulated proton therapy for stage III non-small cell lung cancer. *Int J Radiat Oncol Biol Phys*. 2016;96:661–669.
22. Liu W, Liao Z, Schild SE, et al. Impact of respiratory motion on worst-case scenario optimized intensity modulated proton therapy for lung cancers. *Pract Radiat Oncol*. 2015;5:e77–e86.
23. Heath E, Unkelbach J, Oelfke U. Incorporating uncertainties in respiratory motion into 4D treatment plan optimization. *Med Phys*. 2009;36:3059–3071.
24. Graeff C. Motion mitigation in scanned ion beam therapy through 4D-optimization. *Phys Med*. 2014;30:570–577.
25. Li H, Zhang X, Park P, et al. Robust optimization in intensity-modulated proton therapy to account for anatomy changes in lung cancer patients. *Radiother Oncol*. 2015;114:367–372.
26. Liu W, Schild SE, Chang JY, et al. Exploratory study of 4D versus 3D robust optimization intensity modulated proton therapy for lung cancer. *Int J Radiat Oncol Biol Phys*. 2016;95:523–533.
27. Graeff C, Luchtenborg R, Eley JG, Durante M, Bert CA. 4D optimization concept for scannedion beam therapy. *Radiother Oncol*. 2013;109:419–424.
28. Eley JG, Newhauser WD, Luchtenborg R, Graeff C, Bert C. 4D optimization of scanned ion beam tracking therapy for moving tumors. *Phys Med Biol*. 2014;59:3431.
29. Chang JY, Zhang X, 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.
30. Zhang X, Li Y, 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.
31. Nichols RC, Huh SN, Henderson RH, et al. Proton radiation therapy offers reduced normal lung and bone marrow exposure for patients receiving dose-escalated radiation therapy for unresectable stage III non-small-cell lung cancer: a dosimetric study. *Clin Lung Cancer*. 2011;12:252–257.
32. Lomax A. J. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: the potential effects of calculational uncertainties. *Phys Med Biol*. 2008;53:1027.
33. Lomax AJ. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: the potential effects of inter-fraction and inter-field motions. *Phys Med Biol*. 2008;53:1043.

34. Albertini F, Hug EB, Lomax AJ. Is it necessary to plan with safety margins for actively scanned proton therapy? *Phys Med Biol*. 2011;56:4399–4413.
35. Grassberger C, Daartz J, Dowdell S, Ruggieri T, Sharp G, Paganetti H. Quantification of proton dose calculation accuracy in the lung. *Int J Radiat Oncol Biol Phys*. 2014;89:424–430.
36. Taylor PA, Kry SF, Followill S. Pencil beam algorithms are unsuitable for proton dose calculations in lung. *Int J Radiat Oncol Biol Phys*. 2017;99:750–756.
37. Weistrand O, Svensson S. The ANACONDA algorithm for deformable image registration in radiotherapy. *Med Phys*. 2015;42:40–53.
38. Clark K, Vendt B, Smith K, et al. The cancer imaging archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging*. 2013;26:1045–1057.
39. Hugo GD, Weiss E, Sleeman WC, et al. A longitudinal four-dimensional computed tomography and cone beam computed tomography dataset for image-guided radiation therapy research in lung cancer. *Med Phys*. 2017;44:762–771.
40. Sorriaux J, Testa M, Paganetti H, et al. Experimental assessment of proton dose calculation accuracy in inhomogeneous media. *Phys Med*. 2017;38:10–15.
41. Saini J, Maes D, Egan A, et al. Dosimetric evaluation of a commercial proton spot scanning Monte-Carlo dose algorithm: comparisons against measurements and simulations. *Phys Med Biol*. 2017;62:7659–7681.
42. Pfeiler T, Bäumer C, Engwall E, Geismar D, Spaan B, Timmermann B. Experimental validation of a 4D dose calculation routine for pencil beam scanning proton therapy. *Z Med Phys*. 201728:121–133.
43. Shen J, Tryggestad E, Younkin JE, et al. Technical note: using experimentally determined proton spot scanning timing parameters to accurately model beam delivery time. *Med Phys*. 2017;44:5081–5088.
44. Knopf A-C, 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.
45. Furukawa T, Inaniwa T, Sato S, et al. Moving target irradiation with fast rescanning and gating in particle therapy. *Med Phys*. 2010;37:4874–4879.
46. Mori S, Inaniwa T, Furukawa T, et al. Amplitude-based gated phase-controlled rescanning in carbon-ion scanning beam treatment planning under irregular breathing conditions using lung and liver 4DCTs. *J Radiat Res*. 2014;55:948–958.
47. Poulsen PR, Eley J, Langner U, Simone CB, Langen K. Efficient interplay effect mitigation for proton pencil beam scanning by spot-adapted layered repainting evenly spread out over the full breathing cycle. *Int J Radiat Oncol Biol Phys*. 2018;100:226–234.
48. Engwall E, Glimelius L, Hynning E. Effectiveness of different rescanning techniques for scanned proton radiotherapy in lung cancer patients. *Phys Med Biol*. 2018;63:095006.
49. 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.
50. Liu C, Schild SE, Chang JY, et al. Impact of spot size and spacing on the quality of robustly optimized intensity modulated proton therapy plans for lung cancer. *Int J Radiat Oncol Biol Phys*. 2018;101:479–489.
51. Grassberger C, Dowdell S, Sharp G, Paganetti H. Motion mitigation for lung cancer patients treated with active scanning proton therapy. *Med Phys*. 2015;42:2462–2469.
52. Takahashi W, Nakajima M, Yamamoto N, Tsuji H, Kamada T, Tsujii H. Carbon ion radiotherapy in a hypofractionation regimen for stage I non-small-cell lung cancer. *J Radiat Res*. 2014;55:26–27.
53. Bortfeld T, Paganetti H. The biologic relevance of daily dose variations in adaptive treatment planning. *Int J Radiat Oncol Biol Phys*. 2006;65:899–906.
54. Bernatowicz K, Lomax AJ, Knopf A. Comparative study of layered and volumetric rescanning for different scanning speeds of proton beam in liver patients. *Phys Med Biol*. 2013;58:7905.