Contents lists available at ScienceDirect

Physics and Imaging in Radiation Oncology

Original Research Article

Comparison of 3D and 4D robustly optimized proton treatment plans for non-small cell lung cancer patients with tumour motion amplitudes larger than 5 mm

Saskia Spautz [a], Leon Haase [a], Maria Tschiche [b], Sebastian Makocki [b], Christian Richter [a,b,c,d], Esther G.C. Troost [a,b,c,d,e], Kristin Stützer [a,c,*]

[a] *OncoRay – National Center for Radiation Research in Oncology, Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, Helmholtz-Zentrum Dresden, Rossendorf, Fetscherstraße 74, PF 41, 01307 Dresden, Germany*

[b] *Department of Radiotherapy and Radiation Oncology, Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, Fetscherstraße 74, PF 50, 01307 Dresden, Germany*

[c] *Helmholtz-Zentrum Dresden – Rossendorf, Institute of Radiooncology – OncoRay, Bautzner Landstraße 400, 01328 Dresden, Germany*

[d] *German Cancer Consortium (DKTK), Partner Site Dresden, and German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69192 Heidelberg, Germany*

[e] *National Center for Tumor Diseases (NCT), Partner Site Dresden, Germany: German Cancer Research Center (DKFZ), Heidelberg, Germany; Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany, and; Helmholtz Association / Helmholtz-Zentrum Dresden - Rossendorf (HZDR), Dresden, Germany; Im Neuenheimer Feld 280, 69192 Heidelberg, Germany*

ARTICLE INFO

ABSTRACT

Background and purpose: There is no consensus about an ideal robust optimization (RO) strategy for proton therapy of targets with large intrafractional motion. We investigated the plan robustness of 3D and different 4D RO strategies.

Materials and methods: For eight non-small cell lung cancer patients with clinical target volume (CTV) motion >5 mm, different RO approaches were investigated: 3DRO considering the average CT (AvgCT) with a target density override, 4DRO considering three/all 4DCT phases, and 4DRO considering the AvgCT and three/all 4DCT phases. Robustness against setup/range errors, interplay effects based on breathing and machine log file data for deliveries with/without rescanning, and interfractional anatomical changes were analyzed for target coverage and OAR sparing.

Results: All nominal plans fulfilled the clinical requirements with individual CTV coverage differences <2pp; 4DRO without AvgCT generated the most conformal dose distributions. Robustness against setup/range errors was best for 4DRO with AvgCT (18% more passed error scenarios than 3DRO). Interplay effects caused fraction-wise median CTV coverage loss of 3pp and missed maximum dose constraints for heart and esophagus in 18% of scenarios. CTV coverage and OAR sparing fulfilled requirements in all cases when accumulating four interplay scenarios. Interfractional changes caused less target misses for RO with AvgCT compared to 4DRO without AvgCT (≤42%/33% vs. ≥56%/44% failed single/accumulated scenarios).

Conclusions: All RO strategies provided acceptable plans with equally low robustness against interplay effects demanding other mitigation than rescanning to ensure fraction-wise target coverage. 4DRO considering three phases and the AvgCT provided best compromise on planning effort and robustness.

1. Introduction

Proton therapy offers conformal dose distributions with superior organs at risk (OAR) sparing compared to conventional photon-based radiotherapy [1]. However, the finite proton range is sensitive to variations in the beam path, which leads to inherent challenges when treating moving targets, like lung cancer patients. Respiratory intrafractional motion combined with the density heterogeneities in the thorax can lead to geometrical misses during irradiation. Interference between target motion and dynamic dose delivery with pencil beam

* Corresponding author.
  E-mail address: kristin.stuetzer@oncoray.de (K. Stützer).

scanning (PBS), called interplay effect, can result in unpredictable deviations from the planned dose distribution [2,3]. Motion mitigation strategies such as gating, rescanning, or abdominal compression reduce but might not completely compensate motion-induced dose distortions [4–10]. 4D log file-based dose reconstructions (4DLogReco) can fraction-wise assess the interplay-affected dose distributions [11,12].

Robust plan optimization (RO) can account for range and setup uncertainties, but may also enhance plan robustness against motion-related variations in anatomy by including multiple images of a time-resolved computed tomography (4DCT) scan. The PTCOG Thoracic and Lymphoma Subcommittee guidelines recommend 4DRO to improve plan robustness [13]. There are controversial findings regarding the potential benefit of 4DRO. Some studies reported superior robustness of target coverage and homogeneity with 4DRO, whereas others observed no clear benefit compared to 3DRO, and contrariwise found worse robustness with 4DRO in single cases [14–19]. Including a limited number of 4DCT phases [19,20] could reduce the additional contouring effort and computation time associated with 4DRO.

There is no clear recommendation for generating an optimal robust treatment plan [21], especially for targets with large motion amplitudes, as most studies focused on smaller motions [12,17,22–24]. Furthermore, some clinical workflows consider the average CT (AvgCT) as planning image to generate the necessary digitally reconstructed radiographs for daily X-ray based patient positioning, but the inclusion of the AvgCT during 4DRO and the impact thereof has not been studied yet.

For a patient cohort with respiration-related motions of more than 5 mm, we investigated the impact of different RO strategies on delivered dose distributions when taking intra- and interfractional changes and fractionation effects into account. The 4DLogReco was applied to simulate realistic interplay-affected dose distributions using patient breathing and machine log file data from beam deliveries with/without layered rescanning.

2. Materials and methods

2.1. Patient data

Data from eight non-small cell lung cancer (NSCLC) patients with target motions >5 mm were included in this retrospective study, approved by the Ethics Committee of the Technische Universität Dresden (BO-EK-85022022). These patients had been treated either with photons or passively scattered proton beams mainly due to the large

target motion. The planning CT (pCT) consisted of an amplitude-based reconstructed 4DCT with eight phases and the 3D reconstructed AvgCT. Up to two control CTs (cCT) acquired at approximately fraction 9 and 21 of a modestly accelerated treatment (6 fractions/week) were available for five patients and rigidly registered to the pCT based on grey levels (Supplement S1). Breathing signals measured with a pressure belt (AZ-733V, Anzai Medical Co., Ltd., Tokyo, Japan) during the CT acquisitions were extracted. Deformable image registration (DIR) (ANACONDA [25]) was used to register all 4DCT phases within each CT group onto the reference phase (end-exhalation [EE]). Target volumes were delineated by an experienced radiation oncologist (ET). The primary gross tumor volumes (GTVp) of the 4DCT were combined to an internal GTVp (iGTVp) on the AvgCT. The nodal GTV (GTVn) was delineated on the AvgCT. The clinical target volumes (CTV) were defined by an 8 mm uniform expansion of the respective GTV (confirming that the CTVn encompassed the GTVn in all 4DCT phases) and non-infiltrated anatomical structures were excluded. Target volumes and motion amplitudes are summarized in Table 1 for the pCT and cCTs. Target locations in the pCT are shown in Fig. 1. The OARs spinal cord, heart, esophagus, brachial plexuses and lungs were manually delineated on the AvgrCT and on three 4DCT phases (EE, end-inhalation [EI] and mid-inspiration [MI]), deformed to the remaining 4DCT phases and adjusted if necessary.

2.2. Robust optimization approaches

For each patient, normo-fractionated single field uniform dose plans (66 Gy to the CTV, 2 Gy/fraction) with three individually selected beam angles (Fig. 1) were generated in RayStation v.7.99.3 (RaySearch Laboratories AB, Stockholm, Sweden) (Supplement S2). Five different RO approaches were investigated, all aiming for robust CTV coverage and robust maximum dose (D_{max}) in spinal cord, esophagus and brachial plexuses. The RO strategies consider the anatomy of the following CT images: (3DRO) the AvgCT, but applying a density override with muscle (1.05 g/cm^3) in the iGTVp [26,27]; (4DRO3) three 4DCT phases of EE, EI and MI; (4DRO8) all eight 4DCT phases; (4DRO3A) the AvgCT and three 4DCT phases of EE, EI and MI; (4DRO8A) the AvgCT and all eight 4DCT phases.

A plan was considered acceptable if the calculated dose on the AvgCT as well as the EE and EI phase fulfilled the target coverage and OAR dose criteria of the PRONTOX protocol [28] (Supplement S2). All plans were delivered in dry runs at the clinical gantry room (Proteus®PLUS, Ion

Table 1

Tumor locations, volumes and motions for the investigated cohort on the planning and up to two control 4DCTs per patient. Mean/STD/max motion amplitudes were calculated from the deformation vector field between the inspiration and expiration phase of the respective 4DCT.

ID	Localisation	Fx(s) of cCT	CTV	pCT Volume [cm^3]	pCT Motion [mm] Mean ± STD	pCT Motion [mm] Max	cCT1 Volume [cm^3]	cCT1 Motion [mm] Mean ± STD	cCT1 Motion [mm] Max	cCT2 Volume [cm^3]	cCT2 Motion [mm] Mean ± STD	cCT2 Motion [mm] Max
1*	Upper left	6; 21	CTVp	52.6	3.2 ± 0.8	5.6	51.3	2.4 ± 0.5	5.0	36.3	2.0 ± 0.7	3.9
			CTVn	57.7	5.2 ± 1.4	9.7	51.5	4.8 ± 1.0	6.9	52.4	4.0 ± 1.0	8.2
2	Upper right	8; 20	CTVp	11.5	8.9 ± 1.8	11.9	9.1	6.3 ± 1.9	11.4	10.8	12.5 ± 1.1	15.2
			CTVn	79.9	0.9 ± 0.8	5.8	79.6	3.9 ± 1.8	8.1	51.2	1.3 ± 1.7	11.4
3	Lower left	9; 21	CTVp	129.2	7.0 ± 1.6	12.9	132.5	6.0 ± 0.9	9.8	122.6	6.2 ± 1.4	9.4
			CTVn	13.0	7.4 ± 1.2	12.5	12.3	7.3 ± 0.7	9.2	12.3	6.4 ± 0.9	9.6
4	Lower right	14	CTVp	127.5	11.1 ± 4.0	16.0	114.2	12.4 ± 5.2	18.6	–	–	–
			CTVn	38.0	7.4 ± 2.5	13.8	37.2	7.9 ± 3.0	15.1	–	–	–
5	Upper left	10; 22	CTVp	236.8	6.3 ± 1.6	12.4	308.5	3.4 ± 1.4	7.5	275.0	6.7 ± 2.5	12.9
6	Lower left	–	CTVp	100.8	8.2 ± 2.2	11.9	–	–	–	–	–	–
			CTVn	255.1	9.5 ± 2.2	17.3	–	–	–	–	–	–
7	Lower right	–	CTVp	218.2	13.8 ± 6.9	24.5	–	–	–	–	–	–
			CTVn	90.8	8.0 ± 3.5	16.8	–	–	–	–	–	–
8	Lower left	–	CTVp	182.7	11.0 ± 2.7	17.9	–	–	–	–	–	–
			CTVn	82.8	9.8 ± 1.3	13.0	–	–	–	–	–	–

* Patient with large motion of the ribs in the beam path.

Fx = fraction; pCT = planning CT; cCT = control CT; CTV = clinical target volume; CTVp = primary CTV; CTVn = nodal CTV; STD = Standard deviation; Max = Maximum.
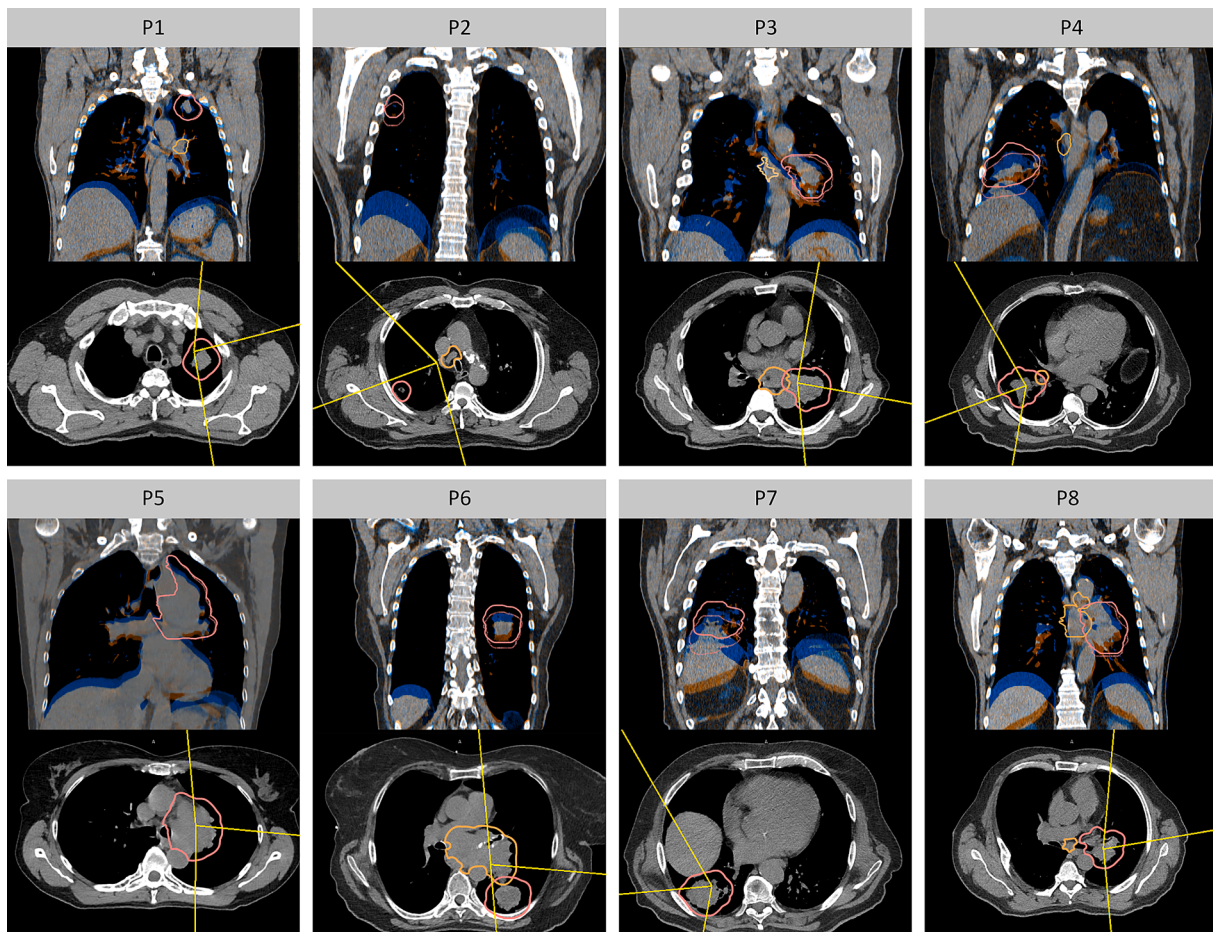
Fig. 1. Coronal fusion view of the end-exhale (blue) and end-inhale (orange) 4DCT phase and a transversal view of the end-exhale phase with indicated beam directions (yellow lines) for all patients (P1-P8). Primary and nodal target (if depicted on the same slice) are delineated in pink and orange, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Beam Applications SA, Louvain-La-Neuve, Belgium), once with and once without five layered rescans. The spot-wise records were extracted from the machine log files.

2.3. Plan evaluations

We performed robustness evaluations against (a) setup and range errors, (b) interplay effects (intrafractional changes in the pCT + dynamic beam delivery) and (c) additional interfractional changes (interplay effected doses on the cCTs). CTV coverage ($D_{98\%}$) and homogeneity index $\left[\mathrm{HI} = (D_{2\%} - D_{98\%})/D_{\mathrm{prescr}}\right]$ as well as relevant OAR dose-volume histogram (DVH) metrics were compared to their respective nominal value on the pCT EE phase.

(a) Setup and range errors

For each plan, 16 perturbed dose scenarios combining a range error of $\pm(3.5\%+2\mathrm{mm})$ and a 5 mm setup error were evaluated on both, the EE and EI phase of the pCT. The number of failed scenarios and the voxel-wise worst case (VWWC) DVH values were analyzed.

(b) Interplay effects

4D dynamic doses (4DDD) with/without five layered rescans were calculated and accumulated on the pCT EE phase using the 4DLogReco [11] to simulate realistic interplay scenarios based on machine log files and patient breathing curves. Irrespective of the local breathing pattern, the three beam delivery logs per plan were manually synchronized with the first, mid and last part of the breathing curve, respectively. Three additional 4DDD scenarios were created by shifting the synchronized logs by $\left(\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\right) \bullet T$ (T: mean breathing period) to uniformly sample

potential starting phases. To estimate a fractionation effect, the four interplay scenarios per plan and rescanning mode were summed with equal weighting.

(c) Interfractional changes

Plan robustness against additional interfractional changes was investigated for the available cCTs of five patients by 4DDD calculations accumulated on the respective cCT EE phase. Again, deliveries with/without rescanning and the fractionation effect by averaging four interplay patterns were considered.

2.4. Statistics

For the statistical evaluation of differences in RO strategies regarding CTV coverage, Wilcoxon signed-rank tests were performed in SPSS Statistics (v27, IBM Deutschland GmbH, Ehningen, Germany) and p-values <0.05 indicated significance.

3. Results

3.1. Nominal plans

All nominal plans fulfilled target coverage ($D_{98\%} > 95\%$) and OAR dose criteria on the AvgCT as well as EE and EI phases, except for patient 5 in whom lung sparing was compromised to ensure target coverage. All RO strategies offered similar target coverage and homogeneity and individual differences remained below 2pp and 6%, respectively (Table 2, Supplement S3). Significant differences, although clinically irrelevant, were only observed between 4DRO3A (favorable $D_{98\%}[\mathrm{CTVn}]$) and

Table 2

Median difference in target coverage $D_{98\%}$ compared to the nominal plan dose (computed on the planning CT EE phase) for each planning strategy, influenced by setup and range errors (reported as voxel-wise worst-case), by interplay effects for single and four accumulated scenarios, for deliveries with and without rescanning and additional interfractional changes in the control CTs. Eight and five patient cases were evaluated on their planning CT and control CT data, respectively.

| | CTV | RO strategy | Median (min, max) of nominal target coverage [%] | Median (min, max) change of target coverage [pp] | | | | |
| | | | | VWWC | 4DDD single scenario | | 4D accumulated scenario | |
					No rescan	5 rescans	No rescan	5 rescans
pCT	CTVp	3DRO	97.9 (96.4, 98.7)	-3.0 (-6.7, -1.7)	-3.1 (-8.1, -0.3)	-1.7 (-7.0, 1.1)	-0.8 (-1.7, 1.5)	-0.0 (-1.4, 1.3)
		4DRO3	98.3 (96.3, 98.6)	-3.9 (-9.7, -1.9)	-3.2 (-10.2, 0.0)	-2.5 (-6.8, -0.6)	-0.8 (-3.3, 0.7)	-0.2 (-2.2, 0.4)
		4DRO8	98.2 (96.5, 98.6)	-3.9 (-10.1, -1.9)	-2.5 (-12.3, 0.4)	-2.0 (-5.4, 1.5)	-0.4 (-2.0, 1.0)	-0.0 (-2.4, 1.0)
		4DRO3A	97.9 (95.4, 98.8)	-3.0 (-7.0, -1.4)	-2.7 (-9.3, 2.5)	-1.6 (-6.2, 1.6)	-0.5 (-1.5, 2.2)	-0.1 (-1.3, 1.6)
		4DRO8A	98.0 (95.5, 98.8)	-3.2 (-7.7, -1.3)	-2.7 (-8.3, 0.8)	-1.7 (-6.9, 0.9)	-0.5 (-1.5, 1.6)	-0.1 (-1.5, 1.1)
	CTVn	3DRO	97.3 (96.8, 98.6)	-4.0 (-7.0, -2.2)	-2.0 (-5.7, -1.0)	-1.8 (-3.1, 0.1)	-0.5 (-1.0, 0.3)	+0.1 (-1.3, 0.7)
		4DRO3	97.7 (97.4, 98.8)	-3.5 (-6.7, -2.3)	-2.4 (-7.9, 0.3)	-1.7 (-3.7, 1.0)	-0.3 (-2.5, 0.3)	+0.1 (-1.8, 0.7)
		4DRO8	97.8 (97.4, 98.7)	-3.7 (-6.9, -2.2)	-2.3 (-5.1, 0.6)	-1.6 (-6.6, 1.4)	-0.4 (-1.4, -0.1)	-0.3 (-2.2, 0.6)
		4DRO3A	97.8 (97.1, 98.5)	-3.1 (-5.6, -2.1)	-1.9 (-5.6, 1.0)	-1.5 (-4.3, 1.1)	-0.6 (-1.0, 0.4)	-0.1 (-0.9, 0.8)
		4DRO8A	97.8 (97.1, 98.5)	-3.1 (-5.8, -2.2)	-2.2 (-8.7, -0.3)	-1.5 (-6.2, 1.1)	-0.6 (-1.3, 0.3)	-0.3 (-0.8, 0.8)
cCT	CTVp	3DRO	95.5 (90.2, 98.6)	–	-2.8 (-7.8, 0.4)	-2.8 (-9.5, 0.3)	-1.8 (-7.3, 0.4)	-1.0 (-6.4, 0.9)
		4DRO3	95.6 (80.1, 98.5)	–	-4.4 (-17.1, 0.1)	-3.8 (-17.1, 0.4)	-2.7 (-16.8, 0.1)	-2.5 (-16.3, 0.2)
		4DRO8	95.9 (80.6, 98.8)	–	-4.0 (-15.8, 0.1)	-3.5 (-15.3, 1.1)	-2.5 (-15.1, 0.3)	-2.6 (-14.8, 0.7)
		4DRO3A	96.7 (88.4, 99.2)	–	-2.7 (-12.8, 0.1)	-2.3 (-8.6, 0.6)	-1.4 (-8.4, 0.5)	-0.9 (-8.1, 1.0)
		4DRO8A	97.0 (89.2, 99.1)	–	-2.8 (-11.1, 0.3)	-2.1 (-8.2, 0.6)	-0.8 (-8.4, 0.5)	-0.8 (-7.5, 1.0)
	CTVn	3DRO	98.0 (96.3, 98.8)	–	-2.3 (-5.3, 0.2)	-1.9 (-4.9, 0.6)	-0.7 (-2.7, -0.3)	-0.8 (-1.9, 0.5)
		4DRO3	98.4 (96.5, 98.9)	–	-1.1 (-6.0, 0.1)	-1.0 (-6.7, 1.1)	-0.0 (-3.1, 0.3)	+0.3 (-3.3, 0.6)
		4DRO8	98.5 (96.2, 98.7)	–	-1.3 (-5.6, 0.3)	-1.1 (-8.8, 0.8)	+0.2 (-3.6, 0.6)	+0.3 (-3.4, 0.7)
		4DRO3A	98.7 (96.7, 98.8)	–	-1.3 (-5.3, -0.1)	-1.5 (-4.4, 0.7)	-0.2 (-1.7, 0.1)	-0.0 (-1.7, 0.7)
		4DRO8A	98.6 (96.5, 98.8)	–	-1.4 (-5.3, 2.3)	-1.3 (-3.3, 0.5)	+0.1 (-1.6, 0.5)	+0.1 (-1.8, 0.9)

CTV = clinical target volume; RO = robust optimization; min = minimum; max = maximum; VWWC = voxel-wise worst-case; 4DDD = 4D dynamic dose; pCT = planning CT; cCT = control CTCTVp = primary CTV; CTVn = nodal CTV.

4DRO8A (favorable $D_{98\%}$[CTVp]). 4DRO without AvgCT provided lower mean lung dose and lower maximum heart dose, whereas 3DRO offered highest sparing of contralateral lung, although differences were not clinically relevant (Supplement S4).

3.2. Robustness against setup and range errors

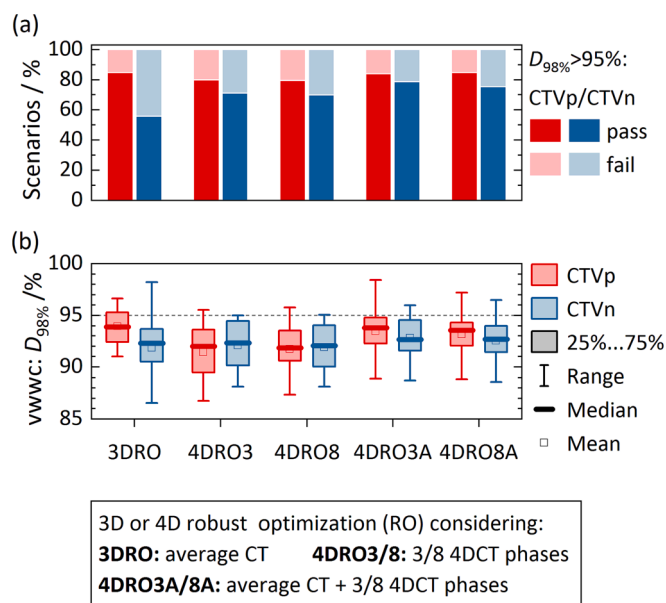Target coverage robustness against setup and range errors was worse

Fig. 2. Coverage robustness against setup and range errors for both primary (CTVp; red) and nodal (CTVn; blue) clinical target volumes, represented by (a) the ratio of scenarios that passed/failed the criterion $D_{98\%} > 95\%$ and (b) the coverage ($D_{98\%}$) of voxel-wise worst-case scenarios (VWWC). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

for the CTVn compared to the CTVp in terms of percentage of passed error scenarios and VWWC coverage (Fig. 2, Table 2). CTVp coverage robustness was higher for RO with AvgCT compared to 4DRO without AvgCT with improved percentage of passed error scenarios (85% vs. 80%) and a significantly lower median VWWC reduction of the $D_{98\%}$ (–3pp vs. –4pp). CTVn coverage robustness was best for plans optimized by 4DRO with AvgCT and worst for 3DRO plans with significant differences in the percentage of passed scenarios (77% vs. 59%) and the median VWWC coverages (92% vs. 91%). In the presence of setup and range errors, all OAR dose values increased by a comparable amount for all plans and were predominantly below the constraints, except the D_{max} in spinal cord, heart and esophagus in single scenarios, mainly for the 3DRO plans.

3.3. Robustness against interplay effects

Without rescanning, no RO strategy provided sufficient target coverage robustness for single fractions, at least for those patients with mean target motions of more than 10 mm (Fig. 3, Table 2). Overall, CTVp/CTVn coverage was failed in 53%/36% of the 4DDD scenarios, respectively. The median loss in CTVp/CTVn coverage relative to the nominal plan was –3pp/–2pp, with highest single deviations of –12pp (4DRO8)/-9pp(4DRO8A), respectively. Only a significantly higher CTVn coverage robustness was observed for 4DRO8 compared to 4DRO8A. The median increase of interplay-affected target coverage by rescanning was <1pp. It reduced the number of 4DDD scenarios with failed CTVp/CTVn coverage to 33%/22% and the worst scenario difference to the nominal target coverage to –7pp. Improved target coverage led also to smaller HI (Supplement S3). 4DRO3 performed significantly worse than 3DRO and 4DRO3A regarding CTVp coverage robustness.

Accumulating the four 4DDD scenarios without/with rescanning led to median improvement of CTVp coverage by 2pp/2pp and CTVn coverage by 1pp/2pp, respectively, and fulfilled $D_{98\%} > 95\%$ in all cases. Overall, target coverage robustness against interplay effects was comparable for all RO strategies, with a slight preference for RO with AvgCT that was also significant for CTVp robustness compared to 4DRO3 for the rescanned accumulated scenarios. Considering only three 4DCT phases
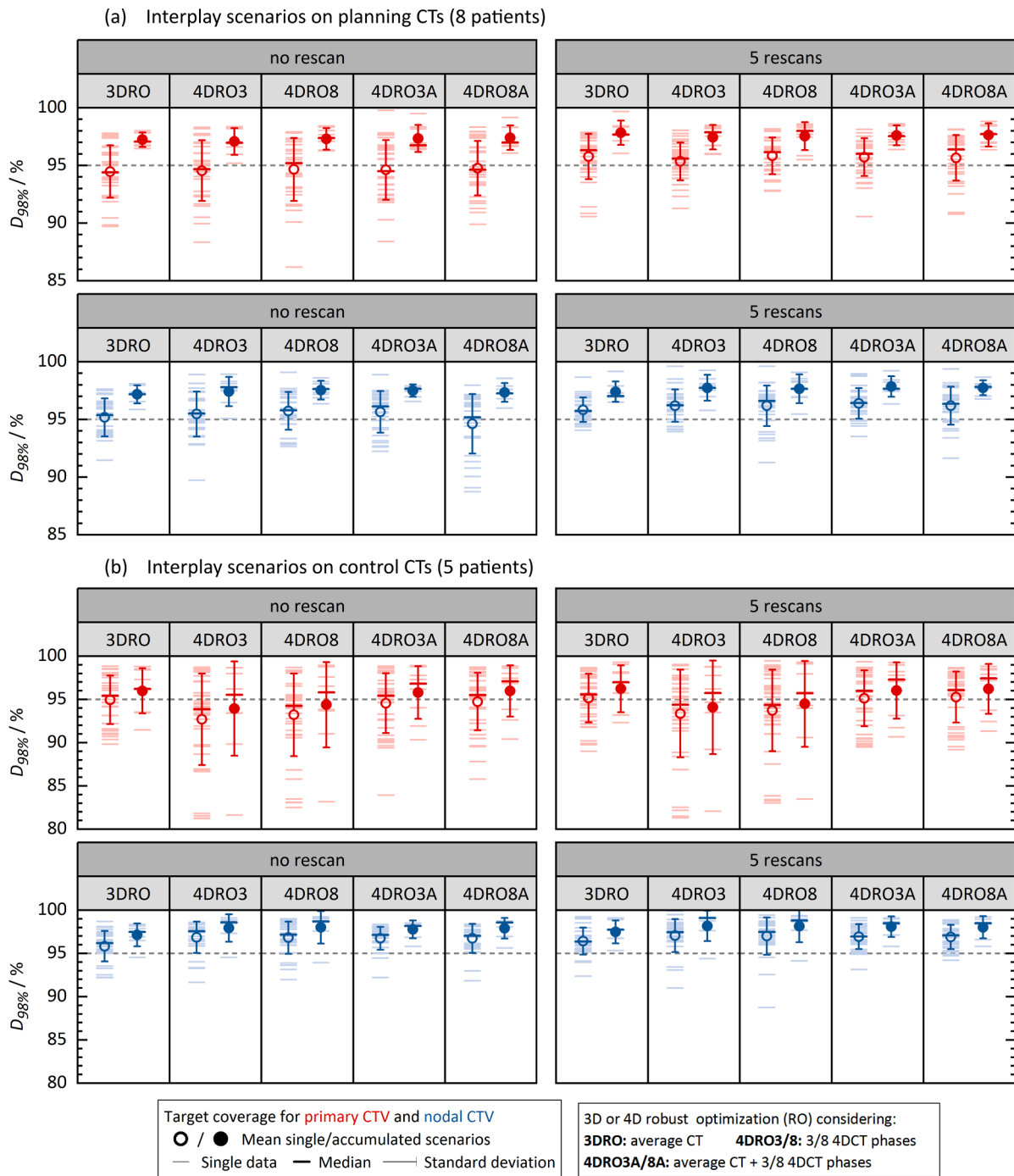
Fig. 3. Target coverage ($D_{98\%}$) of the primary (CTVp; red) and nodal clinical target volume (CTVn; blue) in the planning CTs (a) and the control CTs (b) for interplay-affected dose distributions considering plan deliveries with/without (right/left) layered rescanning in single scenarios (cohort mean: open circle, cohort median: thick tick) and the four accumulated scenario per patient (cohort mean: solid circle, cohort median: thick tick). Target coverage is clinically acceptable when fulfilling $D_{98\%}>95\%$ (dashed line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

during 4DRO did not decrease target coverage robustness.

DVH values for OARs were basically patient-geometry-specific, i.e. similar for different interplay scenarios per patient and for the different RO strategies. Differences in e.g. the spinal cord D_{max} (Fig. 4) were related to deviating nominal plan values but not to different interplay robustness of certain plans (Supplement S4). Accumulating four scenarios had relevant effects for the D_{max} values of OARs close to the target region, i.e. heart and esophagus: Constraints were not fulfilled in about 18% of single 4DDD scenarios, but were met in the accumulated cases, irrespective of rescanning and RO strategy. Largest DVH parameter

deviations between nominal and accumulated scenarios were observed for the contralateral lung (0pp $< \Delta V_{5Gy} <$ 5pp) and the heart (–16 Gy $< \Delta D_{max} <$ 2 Gy).

3.4. Robustness against additional interfractional changes

The cCTs for 5 patients revealed interfractional changes in the target volumes and their maximal motion amplitude of up to 36% and 6 mm, respectively (Table 1). The interfractional change of the mean breathing period was between 0.2 s and 0.6 s. The 4DDD CTVp coverage was more
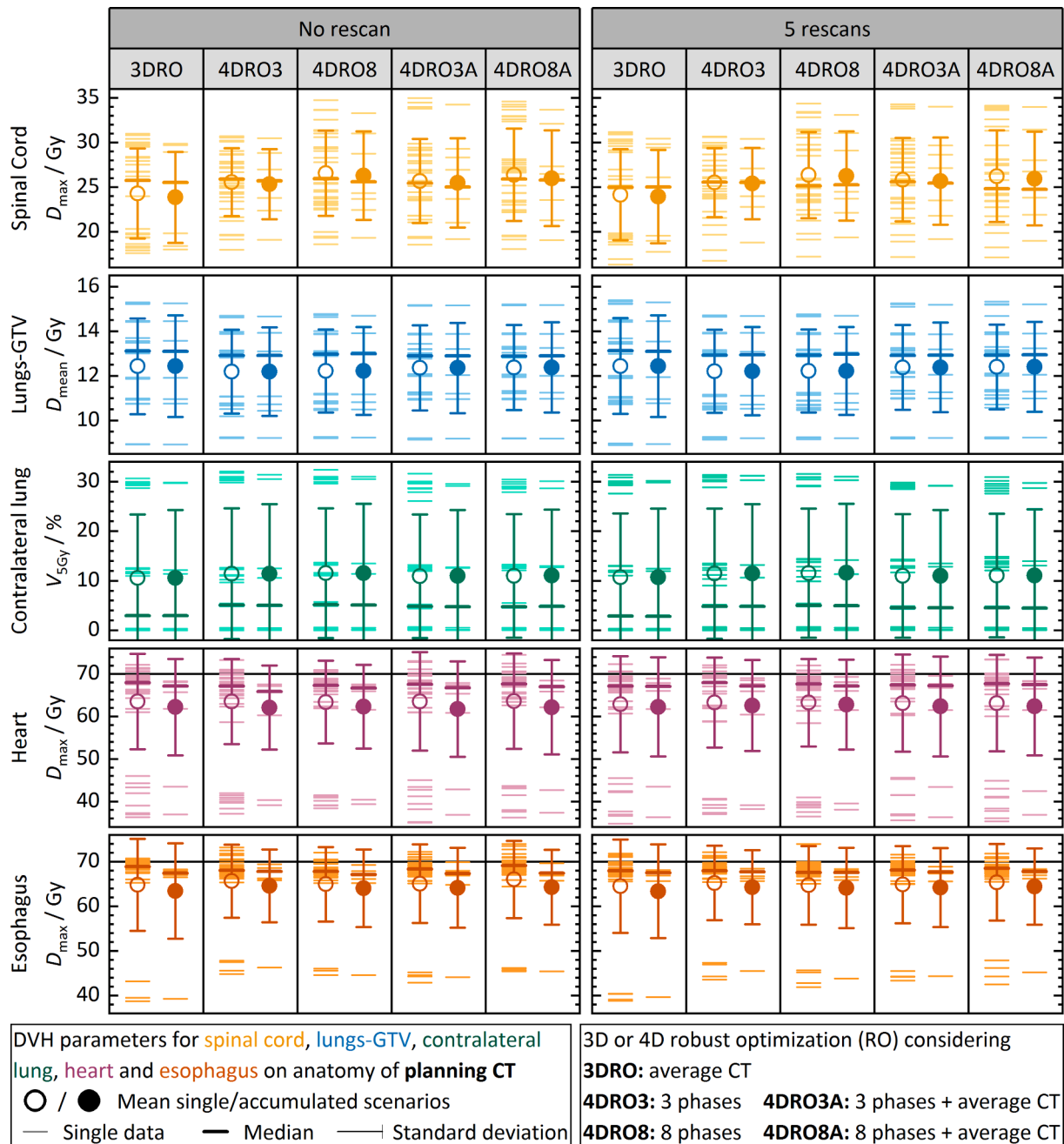
Fig. 4. Organ at risk DVH parameters extracted from interplay-affected dose distributions on the planning CTs from the different robust optimization strategies and considering plan deliveries with/without (right/left) layered rescanning in single scenarios (4 per patient, cohort mean: open circle) and the four accumulated scenario per patient (cohort mean: solid circle, cohort median: thick tick). The DVH parameters for spinal cord and lungs are well below the constraints for all patients while the maximum dose constraints in the heart and esophagus (70 Gy, horizontal line) are violated in several single scenarios.

sensitive to interfractional changes than the CTVn coverage (Fig. 3) with 48% vs. 11% of failed interplay scenarios and a maximum/median reduction of D$_{98\%}$ with respect to the nominal plan of −17pp/−3pp vs. −6pp/−1pp (Table 2). The median improvement of target coverage by rescanning was marginal (0.2 pp). A superior robustness in CTVp coverage (significant in single scenarios and in rescanned accumulated scenarios) was observed for RO with AvgCT (≤42%/33% failed single/accumulated scenarios) compared to 4DRO without AvgCT (>56%/44% failed single/accumulated scenarios). For relevant anatomical changes (esp. cCT1 of patient 4 and 5 and cCT2 of patient 2 and 3), the majority of single 4DDD scenarios of all RO strategies failed CTVp coverage. Accumulating four 4DDD scenarios did not restore CTVn coverage in 9% of the cases.

Regarding OARs, differences between the RO strategies were negligible. Due to the proximity to the target, D$_{max}$ in heart/esophagus was not fulfilled in 8%/3% of the single 4DDD scenarios. For patient 5, even the accumulation of four scenarios failed the heart constraint. Compared to the nominal plans, the largest changes of DVH parameters in accumulated scenarios was again found for the contralateral lung (−4pp < ΔV$_{5Gy}$ < 12pp) and heart (−21 Gy < ΔD$_{max}$ < 16 Gy).

4. Discussion

We compared 3D and 4DRO strategies for proton treatment planning of NSCLC patients with large breathing motion amplitudes (>5 mm) and investigated their robustness against dose impairing factors encountered

throughout patient treatment, including realistic interplay scenarios based on real patient-specific breathing data and machine log files. Besides conventional average-CT-based 3D and 4DCT-based 4DRO approaches, we additionally considered the AvgCT during 4DRO and hereby present the first comprehensive plan comparison study of the aforementioned RO strategies and found improved robustness when including the AvgCT in 4DRO. Furthermore, we investigated the influence on 4DRO results when considering only the three CT phases of extremal and mid elongation compared to all 4DCT phases and found no clinically relevant differences, but a reduced computation time (scaling linearly with the number of considered CT phases) and contouring workload.

All RO strategies yielded clinically acceptable plans and no clinically relevant distinction regarding robustness of target coverage, homogeneity and OAR sparing against motion effects was found, which is in agreement with Ribeiro et al. [17]. Contrary to the PTCOG Thoracic and Lymphoma Subcommittee guideline, the 3DRO provided similarly good results. Conventional 4DRO (without AvgCT) offered best nominal plan conformity, similar to Feng et al. [15], but the target coverage robustness against setup and range uncertainties was highest for 4DRO with AvgCT.

As 4DRO is not designed to optimize for interplay mitigation, a severe loss of target coverage was observed in single 4DDD scenarios for all RO strategies (Fig. 3). This emphasizes the need for either clinically available RO strategies, that include the interplay effect in the optimization process [29], or at least a fraction-wise assessment of the interplay-affected dose throughout patient treatment [11,30]. Interestingly, even for the largest motion amplitudes and irrespective of rescanning, target coverage was restored in all cases when considering fractionation effect by accumulating just four interplay scenarios. However, homogeneous target coverage is pursued in each fraction as it is unknown in how far treatment effectiveness could be impaired by fraction-wise dose distortions, and thus additional motion mitigation strategies should be considered.

Interfractional changes introduced an additional loss in interplay-affected target coverage of more than 10pp in single patients. All plans failed to meet target coverage for relevant anatomical changes, even when fractionation was taken into account (Fig. 3), though 4DRO without AvrgCT performed significantly worse. Therefore, regular control imaging to monitor anatomical changes throughout the treatment course and correspondingly fast plan adaptations [31] are crucial to ensure fulfilled target coverage. Unfortunately, no cCTs were available for the patients with the largest motion amplitudes and thus, strongest interplay-affected target coverage within our cohort.

The beneficial robustness against setup and range uncertainties and against interfractional changes of 4DRO with AvgCT comes at a relatively low price: The plans resulted in an average increase of the mean lung dose of just 0.4 Gy due to a higher number of spots per plan, and thus, slightly longer delivery times compared to 3DRO (+491 spots / +4 s) and standard 4DRO plans (+285 spots / +2 s). For comparison, including three instead of all 4DCT phases in 4DRO resulted in 25 less spots per plan and no noticeable change of delivery time whereas five-times layered rescanning increased the delivery time by approximately 52%.

We found a higher coverage robustness against interplay effects and interfractional changes for CTVn than for CTVp. This is most likely caused by the more central location with homogenous tissue densities and the often smaller motion of the CTVn. However, we observed a higher sensitivity against setup and range errors. This might artificially originate from our CTVn delineation concept of identical contours on the 4DCT phases (since CTVn shifts in the 4DCT are mainly caused by heartbeat rather than respiration).

There are no standards for a reliable motion-related robustness analysis of proton therapy plans. We split our analysis into two parts: robustness against range and setup errors as known for non-moving targets and the motion-related robustness. For the latter, several studies simulated non-realistic interplay patterns by considering

simplified regular breathing patterns and/or arbitrary or model-based spot sorting [3,32–35] while we included both the patient-specific breathing curves from the 4DCTs and machine log-files from dry runs. In total, we considered the combined effects of machine errors, real irregular breathing, inter- and intrafractional patient anatomy changes and fractionation. We would expect further deteriorated target coverage for interplay calculations combined with setup and range errors (see [33]). However, due to the limited number of simulated scenarios, the results do not contain statistical power, i.e. they do not represent the whole spread of possible dose distortions (see [36]). Nevertheless, observed differences in interplay-affected target coverage between the RO strategies were small and not clinically relevant. In an additional evaluation, the influence of interplay effects were excluded by analyzing 4D dose distributions [37] (Supplement S5), allowing for a more objective plan comparison. Despite the fact that 4D dose CTVp/CTVn coverage was in median 3pp/2pp higher than those values observed in the interplay-affected scenarios, the results generally confirmed our findings for the 4DDD doses, i.e. worst robustness against interfractional changes was found for 4DRO without AvgCT.

A general study limitation is the simplified representation of the 4DCT, which cannot map changes in breathing motion amplitude although such intrafractional changes are present in the breathing signals. Moreover, the DIR within the 4DLogReco introduced uncertainties [38,39], but affected all 4DDD dose distributions equally as the same deformation vector fields per 4DCT were used for all RO strategies.

In summary, all investigated RO strategies yielded clinically acceptable nominal plans with comparable target coverage and OAR sparing robustness, especially in the presence of interplay-effects due to large motion amplitudes. We found benefits for 4DRO with AvgCT regarding the target coverage robustness against setup and range errors as well as interfractional changes that need to be monitored throughout the treatment course. Including only three instead of all 4DCT phases during optimization did not impair plan robustness but considerably reduced the computation time and overall workload. We can recommend a 4DRO considering three 4DCT phases and the AvgCT (4DRO3A).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi. org/10.1016/j.phro.2023.100465.

References

[1] Chang JY, Li H, Zhu XR, Liao Z, Zhao L, Liu A, et al. Clinical implementation of intensity modulated proton therapy for thoracic malignancies. Int J Radiat Oncol Biol Phys 2014;90:809–18. https://doi.org/10.1016/j.ijrobp.2014.07.045.

[2] 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;37:223–33. https://doi.org/10.1088/0031-9155/37/1/016.

[3] Bert C, Grözinger SO, Rietzel E. Quantification of interplay effects of scanned particle beams and moving targets. Phys Med Biol 2008;53:2253–65. https://doi.org/10.1088/0031-9155/53/9/003.

[4] Lu H-M, Brett R, Sharp G, Safai S, Jiang S, Madden T, et al. The development and commissioning of a respiratory-gated treatment system for proton therapy. In: Magjarevic R, Nagel JH, editors. World Congr. Med. Phys. Biomed. Eng. 2006, vol. 14. Berlin, Heidelberg: Springer Berlin Heidelberg; 2007. p. 2215–8. https://doi.org/10.1007/978-3-540-36841-0_559.

[5] Grassberger C, Dowdell S, Sharp G, Paganetti H. Motion mitigation for lung cancer patients treated with active scanning proton therapy. Med Phys 2015;42:2462–9. https://doi.org/10.1118/1.4916662.

[6] Schätti A, Zakova M, Meer D, Lomax AJ. The effectiveness of combined gating and re-scanning for treating mobile targets with proton spot scanning. An experimental and simulation-based investigation. Phys Med Biol 2014;59:3813–28. https://doi.org/10.1088/0031-9155/59/14/3813.

[7] Furukawa T, Inaniwa T, Sato S, Shirai T, Mori S, Takeshita E, et al. Moving target irradiation with fast rescanning and gating in particle therapy. Med Phys 2010;37: 4874–9. https://doi.org/10.1118/1.3481512.

[8] Schätti A, Zakova M, Meer D, Lomax J. Experimental verification of motion mitigation of discrete proton spot scanning by re-scanning. Phys Med Biol 2013;58: 8555–72. https://doi.org/10.1088/0031-9155/58/23/8555.

[9] Heinzerling JH, Anderson JF, Papiez L, Boike T, Chien S, Zhang G, et al. Four-dimensional computed tomography scan analysis of tumor and organ motion at varying levels of abdominal compression during stereotactic treatment of lung and liver. Int J Radiat Oncol Biol Phys 2008;70:1571–8. https://doi.org/10.1016/j.ijrobp.2007.12.023.

[10] Mampuya WA, Nakamura M, Matsuo Y, Ueki N, Iizuka Y, Fujimoto T, et al. Interfraction variation in lung tumor position with abdominal compression during stereotactic body radiotherapy. Med Phys 2013;40:091718. https://doi.org/10.1118/1.4819940.

[11] Spautz S, Jakobi A, Meijers A, Peters N, Löck S, Knopf A, et al. Experimental validation of 4D log file-based proton dose reconstruction for interplay assessment considering amplitude-sorted 4DCTs. Med Phys 2022;49:3538–49. https://doi.org/10.1002/mp.15625.

[12] Meijers A, Knopf A-C, Crijns APG, Ubbels JF, Niezink AGH, Langendijk JA, et al. Evaluation of interplay and organ motion effects by means of 4D dose reconstruction and accumulation. Radiother Oncol 2020;150:268–74. https://doi.org/10.1016/j.radonc.2020.07.055.

[13] Chang JY, Zhang X, Knopf A, Li H, Mori S, Dong L, et al. Consensus guidelines for implementing pencil-beam scanning proton therapy for thoracic malignancies on behalf of the PTCOG thoracic and lymphoma subcommittee. Int J Radiat Oncol Biol Phys 2017;99:41–50. https://doi.org/10.1016/j.ijrobp.2017.05.014.

[14] Liu W, Schild SE, Chang JY, Liao Z, Chang YH, Wen Z, et al. Exploratory Study of 4D versus 3D robust optimization in intensity modulated proton therapy for lung cancer. Int J Radiat Oncol Biol Phys 2016;95:523–33. https://doi.org/10.1016/j.ijrobp.2015.11.002.

[15] Feng H, Shan J, Ashman JB, Rule WG, Bhangoo RS, Yu NY, et al. Technical Note: 4D robust optimization in small spot intensity-modulated proton therapy (IMPT) for distal esophageal carcinoma. Med Phys 2021;48:4636–47. https://doi.org/10.1002/mp.15003.

[16] Ge S, Wang X, Liao Z, Zhang L, Sahoo N, Yang J, et al. Potential for improvements in robustness and optimality of intensity-modulated proton therapy for lung cancer with 4-dimensional robust optimization. Cancers 2019;11:35. https://doi.org/10.3390/cancers11010035.

[17] Ribeiro CO, Visser S, Korevaar EW, Sijtsema NM, Anakotta RM, Dieters M, et al. Towards the clinical implementation of intensity-modulated proton therapy for thoracic indications with moderate motion: Robust optimised plan evaluation by means of patient and machine specific information. Radiother Oncol 2021;157: 210–8. https://doi.org/10.1016/j.radonc.2021.01.014.

[18] Wolf M, Anderle K, Durante M, Graeff C. Robust treatment planning with 4D intensity modulated carbon ion therapy for multiple targets in stage IV non-small cell lung cancer. Phys Med Biol 2020;65:215012. https://doi.org/10.1088/1361-6560/aba1a3.

[19] Cummings D, Tang S, Ichter W, Wang P, Sturgeon JD, Lee AK, et al. Four-dimensional plan optimization for the treatment of lung tumors using pencil-beam scanning proton radiotherapy. Cureus 2018;10(8):e3192. https://doi.org/10.7759/cureus.3192.

[20] Mastella E, Molinelli S, Pella A, Vai A, Maestri D, Vitolo V, et al. 4D strategies for lung tumors treated with hypofractionated scanning proton beam therapy: Dosimetric impact and robustness to interplay effects. Radiother Oncol 2020;146: 213–20. https://doi.org/10.1016/j.radonc.2020.02.025.

[21] Knopf AC, Czerska K, Fracchiolla F, Graeff C, Molinelli S, Rinaldi I, et al. Clinical necessity of multi-image based (4DMIB) optimization for targets affected by respiratory motion and treated with scanned particle therapy – A comprehensive review. Radiother Oncol 2022;169:77–85. https://doi.org/10.1016/j.radonc.2022.02.018.

[22] Liu W, Liao Z, Schild SE, Liu Z, Li H, Li Y, et al. Impact of respiratory motion on worst-case scenario optimized intensity modulated proton therapy for lung cancers. Pract Radiat Oncol 2015;5:e77–86. https://doi.org/10.1016/j.prro.2014.08.002.

[23] Fracchiolla F, Dionisi F, Giacomelli I, Hild S, Esposito PG, Lorentini S, et al. Implementation of proton therapy treatments with pencil beam scanning of targets with limited intrafraction motion. Phys Med 2019;57:215–20. https://doi.org/10.1016/j.ejmp.2019.01.007.

[24] Inoue T, Widder J, van Dijk LV, Takegawa H, Koizumi M, Takashina M, 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;96:661–9. https://doi.org/10.1016/j.ijrobp.2016.06.2454.

[25] Weistrand O, Svensson S. The ANACONDA algorithm for deformable image registration in radiotherapy. Med Phys 2014;42:40–53. https://doi.org/10.1118/1.4894702.

[26] Kang Y, Zhang X, Chang JY, Wang H, Wei X, Liao Z, et al. 4D Proton treatment planning strategy for mobile lung tumors. Int J Radiat Oncol Biol Phys 2007;67: 906–14. https://doi.org/10.1016/j.ijrobp.2006.10.045.

[27] Hoppe BS, Flampouri S, Henderson RH, Pham D, Bajwa AA, D'Agostino H, et al. Proton therapy with concurrent chemotherapy for non – small-cell lung cancer: technique and early results. Clin Lung Cancer 2012;13:352–8. https://doi.org/10.1016/j.cllc.2011.11.008.

[28] Zschaeck S, Simon M, Löck S, Troost EGC, Stützer K, Wohlfahrt P, et al. PRONTOX - proton therapy to reduce acute normal tissue toxicity in locally advanced non-small-cell lung carcinomas (NSCLC): Study protocol for a randomised controlled trial. Trials 2016;17. https://doi.org/10.1186/s13063-016-1679-4.

[29] Engwall E, Fredriksson A, Glimelius L. 4D robust optimization including uncertainties in time structures can reduce the interplay effect in proton pencil beam scanning radiation therapy. Med Phys 2018;45:4020–9. https://doi.org/10.1002/mp.13094.

[30] Meijers A, Jakobi A, Stützer K, Guterres Marmitt G, Both S, Langendijk JA, et al. Log file-based dose reconstruction and accumulation for 4D adaptive pencil beam scanned proton therapy in a clinical treatment planning system: implementation and proof-of-concept. Med Phys 2019;46:1140–9. https://doi.org/10.1002/mp.13371.

[31] Borderías Villarroel E, Geets X, Sterpin E. Online adaptive dose restoration in intensity modulated proton therapy of lung cancer to account for inter-fractional density changes. Phys Imaging Radiat Oncol 2020;15:30–7. https://doi.org/10.1016/j.phro.2020.06.004.

[32] Liu C, Schild SE, Chang JY, Liao Z, Korte S, Shen J, 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;101:479–89. https://doi.org/10.1016/j.ijrobp.2018.02.009.

[33] Ribeiro CO, Meijers A, Korevaar EW, Muijs CT, Both S, Langendijk JA, et al. Comprehensive 4D robustness evaluation for pencil beam scanned proton plans. Radiother Oncol 2019;136:185–9. https://doi.org/10.1016/j.radonc.2019.03.037.

[34] Souris K, Barragan Montero A, Janssens G, Di Perri D, Sterpin E, Lee JA. Technical Note: Monte Carlo methods to comprehensively evaluate the robustness of 4D treatments in proton therapy. Med Phys 2019;46:4676–84. https://doi.org/10.1002/mp.13749.

[35] Dowdell S, Grassberger C, Sharp GC, Paganetti H. 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;58:4137–56. https://doi.org/10.1088/0031-9155/58/12/4137.

[36] Pastor-Serrano O, Habraken S, Lathouwers D, Hoogeman M, Schaart D, Perkó Z. How should we model and evaluate breathing interplay effects in IMPT? Phys Med Biol 2021;66:235003. https://doi.org/10.1088/1361-6560/ac383f.

[37] Taasti VT, Hattu D, Vaassen F, Canters R, Velders M, Mannens J, et al. Treatment planning and 4D robust evaluation strategy for proton therapy of lung tumors with large motion amplitude. Med Phys 2021;48:4425–37. https://doi.org/10.1002/mp.15067.

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

[39] Mastella E, Mirandola A, Russo S, Vai A, Magro G, Molinelli S, et al. High-dose hypofractionated pencil beam scanning carbon ion radiotherapy for lung tumors: Dosimetric impact of different spot sizes and robustness to interfractional uncertainties. Phys Med 2021;85:79–86. https://doi.org/10.1016/j.ejmp.2021.05.004.