Towards the clinical implementation of intensity-modulated proton therapy for thoracic indications with moderate motion

Ribeiro, Cássia O; Visser, Sabine; Korevaar, Erik W; Sijtsema, Nanna M; Melissa Anakotta, R; Dieters, Margriet; Both, Stefan; Langendijk, Johannes A; Wijsman, Robin; Muijs, Christina T

Published in:
Radiotherapy and Oncology

DOI:
10.1016/j.radonc.2021.01.014

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

Radiotherapy and Oncology 157 (2021) 210–218

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

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

Cássia O. Ribeiro [a,*], Sabine Visser [a], Erik W. Korevaar [a], Nanna M. Sijtsema [a], R. Melissa Anakotta [a], Margriet Dieters [a], Stefan Both [a], Johannes A. Langendijk [a], Robin Wijsman [a], Christina T. Muijs [a], Arturs Meijers [a], Antje Knopf [a,b]

[a] Department of Radiation Oncology, University Medical Center Groningen, University of Groningen, The Netherlands; [b] Division for Medical Radiation Physics, Carl von Ossietzky University Oldenburg, Germany

ARTICLE INFO

Article history:
Received 13 February 2020
Received in revised form 9 December 2020
Accepted 6 January 2021
Available online 3 February 2021

Keywords:
Intensity-modulated proton therapy
Lung cancer
Oesophageal cancer
3D robust optimisation
4D robust optimisation
4D robustness evaluation

ABSTRACT

Purpose: Compared to volumetric modulated arc therapy (VMAT), clinical benefits are anticipated when treating thoracic tumours with intensity-modulated proton therapy (IMPT). However, the current concern of plan robustness as a result of motion hampers its wide clinical implementation. To define an optimal protocol to treat lung and oesophageal cancers, we present a comprehensive evaluation of IMPT planning strategies, based on patient 4DCTs and machine log files.

Materials and methods: For ten lung and ten oesophageal cancer patients, a planning 4DCT and weekly repeated 4DCTs were collected. For these twenty patients, the CTV volume and motion were assessed based on the 4DCTs. In addition to clinical VMAT plans, layered rescanned 3D and 4D robust optimised IMPT plans (IMPT_3D and IMPT_4D respectively) were generated, and approved clinically, for all patients. The IMPT plans were then delivered in dry runs at our proton facility to obtain log files, and subsequently evaluated through our 4D robustness evaluation method (4DREM). With this method, for each evaluated plan, fourteen 4D accumulated scenario doses were obtained, representing 14 possible fractionated treatment courses.

Results: From VMAT to IMPT_3D, nominal D_{mean}(lungs-GTV) decreased 2.75 ± 0.56 Gy_{RBE} and 3.76 ± 0.92 Gy_{RBE} over all lung and oesophageal cancer patients, respectively. A more pronounced reduction was verified for D_{mean}(heart): 5.38 ± 7.36 Gy_{RBE} (lung cases) and 9.51 ± 2.25 Gy_{RBE} (oesophagus cases). Target coverage robustness of IMPT_3D was sufficient for 18/20 patients. Averaged dose in critical structures over all 4DREM scenarios changed only slightly for both IMPT_3D and IMPT_4D. Relative to IMPT_3D, no gain in IMPT_4D was observed.

Conclusion: The dosimetric superiority of IMPT over VMAT has been established. For most thoracic tumours, our IMPT_3D planning protocol showed to be robust and clinically suitable. Nevertheless, accurate patient positioning and adapting to anatomical variations over the course of treatment remain compulsory.

Conformal and highly precise radiotherapy techniques are required for thoracic indications due to the organs-at-risk (OARs) surrounding the tumour, such as the lungs and the heart. Volumetric modulated arc therapy (VMAT) is capable of reducing the number of monitor units and treatment delivery time, and proved to result at least in similar target coverage and OARs dose sparing for lung and oesophageal cancer patients, compared to conventional photon modalities [1,2]. In terms of treatment related toxicities, even more clinical benefits are anticipated with pencil beam scanned proton therapy (PBS-PT) for these patients. Superior planned dose distributions can be achieved due to the shape of the proton depth-dose curve (low entrance dose, high peak dose and sharp distal dose fall-off) [3–6]. Within the more recent optimisation developments of PBS-PT, intensity-modulated proton therapy (IMPT) is able to improve dose conformality to the target, while reducing the dose to the OARs [7–9].

* Corresponding author at: Department of Radiation Oncology, University Medical Center Groningen, PO Box 30001, 9700 RB Groningen, The Netherlands.
E-mail address: c.oraboni.ribeiro@umcg.nl (C.O. Ribeiro).

https://doi.org/10.1016/j.radonc.2021.01.014
0167-8140/

Despite the anticipated advantages of IMPT for thoracic tumours, the concern of plan robustness to possible treatment uncertainties (machine delivery imperfections, patient setup variations, CT number conversions into proton stopping power, anatomical changes, and intra-fractional motion) hampers its wide clinical implementation [10–12]. Especially for moving indications, dose inhomogeneities caused by the interaction between patient respiratory motion and the delivered pencil beams (interplay effects) can occur [13]. Rescanning (i.e. delivering the planned proton spots multiple times) has been shown to be effective in mitigating interplay effects [14]. Increasing the spot size has also demonstrated to preserve plan robustness against inter- and intra- fractional uncertainties and interplay effects for thoracic indications [15,16]. Additionally, sophisticated planning approaches, such as 4D robust optimisation, have been proposed to moderate respiratory-induced dosimetric impacts for IMPT [17–19].

Comprehensive robustness evaluations of novel robust optimisation techniques in IMPT treatment planning for thoracic indications are crucial before their clinical deployment. For oesophageal cancer patients, an already clinically implemented IMPT planning protocol was assessed using a robustness evaluation method for setup and range errors, and independently, also for breathing motion [20]. Anatomy changes (provided by repeated patient CT imaging) were also investigated for robust optimised IMPT plans [21]. Additionally, a robustness comparison between 3D and 4D robust optimisation strategies has been performed by Liu et al. [18], by examining the combined influence of setup and range uncertainties, and also, but separately, breathing motion and interplay effects. Using a similar evaluation method, other optimisation parameters (spot size and spacing) for lung robust optimised IMPT have been reported [22]. A robustness evaluation study by Inoue et al. [23] incorporated setup and range errors and breathing motion, and individually, interplay (per energy layer and not per spot), in robust optimised IMPT plans for stage III non-small cell lung cancer (NSCLC) patients. A more complete robustness verification tool, considering the smearing effect of fractionation in the combined impact of interplay, setup and range errors, and breathing motion (simulating the variability in period and amplitude), was later released by Souris et al. [24].

Furthermore, a comprehensive 4D robustness evaluation method (4DREM), accounting for all the formerly mentioned PBS-PT uncertainties for thoracic indications simultaneously, with the inclusion of machine errors, was published by our team [25]. Here we apply this method to perform IMPT robustness evaluation analysis of different optimisation approaches using log files (acquired from specific treatment-plan dry runs) and extensive 4DCT imaging. To our knowledge, this is the first study to use patient and machine specific data to define an optimal clinical planning protocol for IMPT for a representative number of lung and oesophageal cancer patients.

Materials and methods

Patient data

A prospective study approved by the medical ethics review committee of the UMCG (ClinicalTrials.gov NCT03024138) included, by written informed consent, patients with thoracic malignancies to undergo a planning 4DCT and weekly repeated 4DCTs during the course of treatment. Twenty patients (ten stage III NSCLC and ten stage IB-IVA oesophageal cancer patients [see Table 1]), with a planning 4DCT and five or six weekly repeated 4DCTs, were retrospectively randomly selected for our study (Suppl. 1). Each 4DCT was reconstructed into ten respiratory phases (determined using an Anzai belt [Anzai Medical, Tokyo, Japan]) and an averaged CT. All 4DCTs were inspected, and only major-artefact-free scans were used. The internal clinical target volume (iCTV) and planning target volume (PTV) of the planning 4DCT were defined for all patients, taking into account the respective breathing phases (Suppl. 2.1) [26–29]. Clinical target volumes (CTVs) were delineated on all breathing phases (Suppl. 2.2). Additionally, OARs such as the heart, the spinal cord, the oesophagus (exclusively for lung cancer patients), the lungs-minus-gross tumour volume (lungs-GTV), among others, were delineated on the averaged planning CT (av_pCT) and on the end-of-exhalation planning CT phase (EE_pCT). All these delineations were approved by radiation oncologists. As specified in Suppl. 3, the CTV volume and motion throughout the treatment course were assessed based on the available weekly 4DCTs (Table 1).

Treatment planning

VMAT and IMPT plans were produced for all patients in the RayStation 6.99 (RaySearch Laboratories, Stockholm, Sweden) treatment planning system (TPS). Dose parameters are reported in terms of relative biological effectiveness (RBE) corrected dose, assuming RBE values of 1.0 and 1.1 for VMAT and IMPT, respectively [30]. Prescribed doses (PDs) were 60.00 Gy_{RBE} (2.40 Gy_{RBE} in 25 fractions) and 41.40 Gy_{RBE} (1.80 Gy_{RBE} in 23 fractions) for the lung and oesophageal cancer patients, respectively. Besides the strict planning criteria for both VMAT and IMPT (Suppl. 4) [28], all plans were individually thoroughly revised in several meetings within a multidisciplinary team of treatment planners, radiation oncologists, and medical physicists (regarding optimal beam arrangements, necessary overrides, adequate target coverage, and minimisation of OARs dose), until a clinically acceptable plan was achieved.

VMAT

As currently performed in our photon clinical workflow, PTV optimised VMAT plans (with unique contribution from VMAT modality 6 MV fields [two half arcs]) were created for all patients on the av_pCT. The dose was prescribed to the PTV and the collapsed cone dose engine was used [31]. For all finalized VMAT plans: V_{95}(PTV) in the nominal dose distribution $\geq$ 98 %.

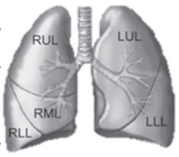
IMPT

IMPT plans were generated for all patients using the Monte Carlo dose engine (Suppl. 5.1) [32,33]. The minimax robust optimisation approach was used [34], aiming for robustness against $\pm$ 3% range uncertainties, and setup uncertainties of 6.0 mm and 8.0 mm (equivalent to the iCTV to PTV margin in our photon treatments) for lung and oesophageal cancer patients [35–37], respectively.

Three beams were used for all lung cases. The individual patient field directions were chosen based on tumour location, OARs involvement, plan robustness, and compliance with planning criteria (Table S.5.1). For the oesophagus indications, typically two fields (posterior-anterior and right-posterior oblique) were selected, except for one case (patient 19), for whom target volume extended in cranial direction, requiring an additional anterior field. For each patient, both 3D and 4D robust optimised IMPT plans (IMPT_3D and IMPT_4D respectively) were generated, sharing exactly the same beam arrangement. The IMPT_3D plan was created on the av_pCT. The IMPT_4D plan was created on the EE_pCT, using all planning 4DCT phases during the optimisation [18]. The nominal dose was prescribed to the iCTV in the IMPT_3D plan and to the CTV in the IMPT_4D plan. To ensure a fair plan comparison, the difference in fulfilled mean dose to the target structure (iCTV or CTV) between IMPT_3D and IMPT_4D plans, was assured

Table 1
Primary tumour location and CTV characteristics (mean ± SD volume and motion over the weeks of treatment) for the lung and oesophageal cancer patients included in this study.

Lung			
Patient	**Location**	**Volume [cm³]**	**Motion [mm]**
1	RLL	372 ± 21 (339 - 391)	5.7 ± 1.3 (4.5 - 7.9)
2	LLL	55 ± 6 (49 - 64)	4.6 ± 1.4 (2.6 - 6.8)
3	RUL	55 ± 3 (53 - 59)	3.6 ± 1.0 (2.9 - 5.7)
4	RUL	71 ± 4 (66 - 78)	1.8 ± 0.3 (1.5 - 2.3)
5	RUL	46 ± 9 (29 - 55)	3.8 ± 1.0 (2.8 - 5.6)
6	RLL	108 ± 16 (88 - 129)	3.9 ± 0.7 (3.3 - 4.8)
7	RUL	145 ± 13 (133 - 171)	2.0 ± 0.3 (1.7 - 2.4)
8	RUL	66 ± 2 (64 - 69)	3.1 ± 0.4 (2.9 - 3.8)
9	RUL	95 ± 22 (67 - 125)	2.9 ± 1.0 (1.8 - 4.5)
10	LUL	51 ± 1 (50 - 52)	4.8 ± 0.6 (4.1 - 5.7)
Oesophagus			
Patient	**Location**	**Volume [cm³]**	**Motion [mm]**
11	D	367 ± 14 (356 - 393)	6.4 ± 0.4 (6.2 - 7.2)
12	D	329 ± 17 (303 - 347)	6.0 ± 0.9 (5.1 - 7.4)
13	M	117 ± 7 (110 - 128)	4.7 ± 1.2 (3.6 - 6.3)
14	D	158 ± 10 (149 - 173)	9.1 ± 1.5 (7.3 - 11.1)
15	D	134 ± 10 (117 - 147)	6.4 ± 0.7 (5.5 - 7.2)
16	D	210 ± 3 (207 - 215)	3.1 ± 0.6 (2.2 - 3.7)
17	D	232 ± 6 (223 - 239)	7.6 ± 1.2 (6.3 - 9.8)
18	D	411 ± 22 (379 - 439)	3.4 ± 0.3 (2.9 - 3.7)
19	PMD	466 ± 39 (404 - 507)	8.3 ± 2.2 (4.2 - 10.4)
20	D	439 ± 27 (410 - 479)	5.9 ± 1.2 (4.4 - 7.6)



Abbreviations: RUL = Right Upper Lobe; RML = Right Middle Lobe; RLL = Right Lower Lobe; LUL = Left Upper Lobe; LLL = Left Lower Lobe; P = Proximal; M = Middle; D = Distal.

to be within ± 0.50 Gy_{RBE}. A density override to muscle tissue (1.050 g/cm^3) was applied within the iCTV (excluding bone) for the IMPT_3D plan. The override to the iCTV was removed for final dose evaluation.

Preliminary robustness evaluation of all IMPT plans (IMPT_3D and IMPT_4D) was then performed on the av_pCT [38]. This 3D robustness evaluation method (3DREM), which is part of our clinical protocol for proton treatment planning, accounts for several disturbing scenarios, simulating different patient setup and range errors. Setup errors were modelled by shifting the planning isocentre in fixed translations in fourteen directions (with magnitudes of 6.0 mm and 8.0 mm for lung and oesophagus indications, respectively). Range errors were considered by applying density perturbations of ± 3%. The robustness of the plans was then evaluated in voxel-wise worst-case minimum (Vw_{min}) and voxel-wise worst-case maximum (Vw_{max}) dose distributions, which score the minimum and maximum dose per voxel over all calculated scenarios, respectively. If all robustness criteria were met (Suppl. 5.2), and after final clinical acceptance, the plans were delivered in dry runs at our proton facility to obtain log files. The spot sizes at our beam line range from 6.5 mm to 3.0 mm for proton energies from 70 MeV to 230 MeV in air (sigma at isocentre). Five times layered rescanning was used as motion mitigation technique [39].

4DREM for IMPT plans

To account for the impact of the disturbing effects occurring when treating moving targets with PBS-PT, all IMPT plans were subsequently evaluated through our 4DREM [25]. Using sub-plans (derived from the machine log files, assuming constant breathing cycles) and all phases of all available patient 4DCTs, the 4DREM assesses the plan robustness for the combination of (1) setup and range errors, (2) machine errors, (3) patient anatomy changes, (4) breathing motion, and (5) interplay effects. At first, rigid registrations were conducted from the planning to each repeated 4DCT (Suppl. 6). Then, sub-plan doses were calculated on all individual 4DCT phases, considering setup and range errors. Setup and range errors were simulated similarly as in the 3DREM, with the inclusion of the dose-fraction-smoothening effect of eight fractions per scenario [40], and a decrease in the shifts magnitude to 2.0 mm. The 2.0 mm remaining setup uncertainty in the 4DREM (calculated internally) has been established due to the use of the available repeated imaging, which led to the disregard of the patient inter-fractional setup error in the magnitude of the simulated shifts (see Suppl. 7) [25,28,29,41–43]. A fraction dose was calculated by applying the same setup and range errors to all sub-plan doses of that fraction specific 4DCT. For each fraction calculation, the 4DCT starting phase of the delivery was randomly selected,

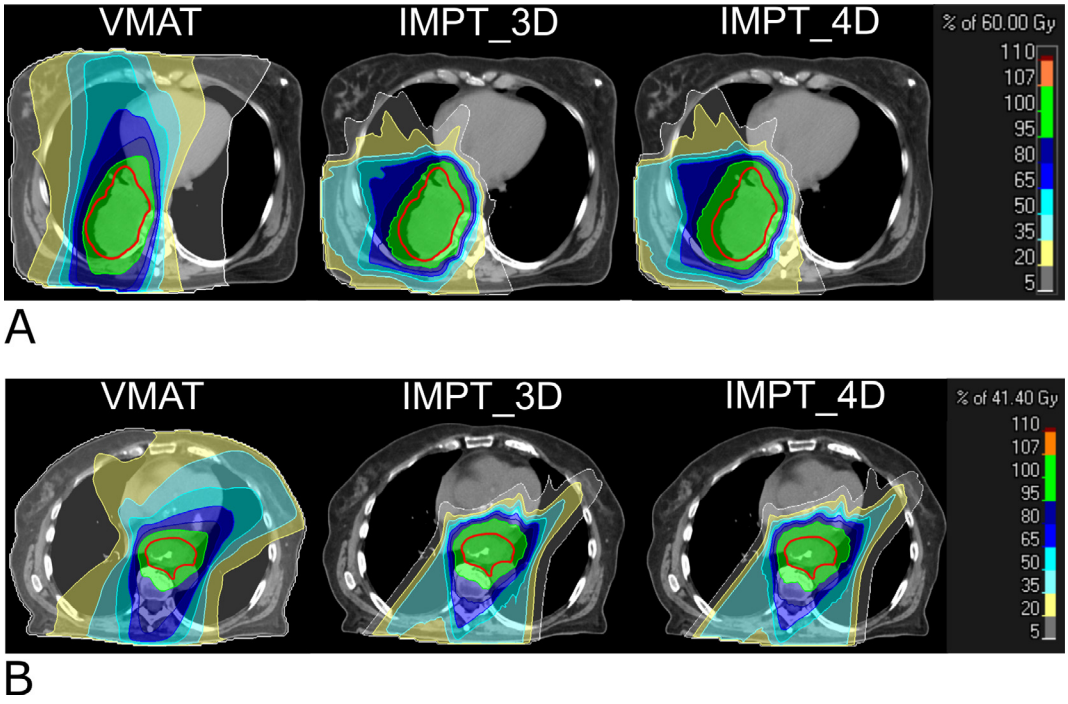


Fig. 1. Nominal dose distributions for sample **A:** lung and **B:** oesophagus cases (patients 1 and 20 respectively) planned with VMAT, IMPT_3D, and IMPT_4D. The red contour represents the iCTV.

and for all fractions of the same scenario, the range error was randomly designated (0/± 3%) [11,25,34]. Finally, the entire treatment course dose distribution was obtained by performing 4D dose accumulation of eight fraction doses based on different 4DCTs onto the EE_pCT.

For each plan scenario, the available 4DCTs per patient were distributed and equally weighted through the eight evaluated fractions. For the first two fractions, 4D dose accumulation of sub-plan doses was performed on the planning 4DCT. For the subsequent two fractions, the first repeated 4DCT, or consecutively the first and second repeated 4DCTs if seven 4DCTs were available in total, were used. For the last four fractions, the remaining repeated 4DCTs were successively selected. The ANACONDA deformable image registration (DIR) method [44] available in the TPS was used for the 4D dose accumulation, with the CTV as controlling region of interest (ROI).

With the 4DREM, for each evaluated plan, 14 4D accumulated scenario dose distributions were obtained, representing 14 possible fractionated treatment courses of the nominal plan [25]. IMPT plan robustness was evaluated on the EE_pCT through the obtained scenario doses. The dose-volume histogram (DVH) of the CTV and respective metric V_{95} were examined in the Vw_{min} dose distribution of the 4D accumulated scenario doses ($4DVwa_{min}$) [9,38]. Additionally, the OAR DVH indices D_{mean}(lungs-GTV), D_{mean}(heart), and D_1(spinal cord) (MLD, MHD, and D_1(spine) respectively) were averaged over all scenarios resulting from the execution of the 4DREM, and extracted for all plans.

Results

IMPT_4D plans were computed and analysed on the av_pCT for nominal treatment plan comparisons with IMPT_3D or VMAT. Planned dose distributions obtained with VMAT, IMPT_3D, and IMPT_4D for two sample cases (a lung and an oesophageal cancer patients) are shown (Fig. 1). From the axial slice, for both patients a clear gain for IMPT relative to VMAT can be seen in terms of reduction of low-dose deposition in OARs. Additionally, minor differences in conformality between IMPT_3D and IMPT_4D can be observed.

For all patients, the OAR DVH parameters obtained with VMAT, IMPT_3D, and IMPT_4D were computed (Fig. 2, Suppl. 8, and Fig. S.9.1). For the lung cancer patients, mean ± SD differences in MLD, MHD, and D_1(spine) between VMAT and IMPT_3D were 2.75 ± 0.56 Gy_{RBE}, 5.38 ± 7.36 Gy_{RBE}, and 17.71 ± 8.59 Gy_{RBE}, respectively. For the oesophageal cancer patients, these differences were 3.76 ± 0.92 Gy_{RBE}, 9.51 ± 2.25 Gy_{RBE}, and −0.51 ± 4.85 Gy_{RBE} for MLD, MHD, and D_1(spine), respectively. Maximum differences between VMAT and IMPT_3D reached up to 3.52 Gy_{RBE} (MLD), 20.58 Gy_{RBE} (MHD), and 30.70 Gy_{RBE} (D_1(spine)) for the lung, and 5.09 Gy_{RBE} (MLD), 13.55 Gy_{RBE} (MHD), and 5.20 Gy_{RBE} (D_1(spine)) for the oesophageal cancer patients. Relative to the planned dose differences between VMAT and IMPT_3D, the differences between IMPT_3D and IMPT_4D remained almost indiscernible. Concerning OARs, there was no pronounced dosimetric benefit of IMPT_4D over IMPT_3D. For the lung cases, mean ± SD differences in MLD, MHD, and D_1(spine) between IMPT_3D and IMPT_4D were 0.34 ± 0.37 Gy_{RBE}, 0.39 ± 0.53 Gy_{RBE}, and 1.68 ± 4.19 Gy_{RBE}, respectively. For the oesophagus cases, these differences were −0.04 ± 0.22 Gy_{RBE}, −0.40 ± 0.62 Gy_{RBE}, and 0.05 ± 0.34 Gy_{RBE} for MLD, MHD, and D_1(spine), respectively.

V_{95}(CTV) of the $4DVwa_{min}$ dose, and averaged MLD, MHD, and D_1(spine) over all scenarios considered within the 4DREM were plotted (Fig. 3). For most IMPT_3D / IMPT_4D plans, V_{95}(CTV) values in the $4DVwa_{min}$ dose were sufficient (≥98 %), except for three lung cases (patients 8, 9, and 10), and one oesophagus case (patient 19) (Fig. 3A). For patients 9 and 10, unlike IMPT_4D, an acceptable robustness was obtained for IMPT_3D regarding target coverage (Suppl. 10). In general, concerning relevant OARs (Fig. 3B and Fig. S.9.2), there was no gain in IMPT_4D over IMPT_3D. In some cases, there was a slight benefit from IMPT_4D, and in others the opposite was observed. The variations obtained in the investigated OAR doses between different disturbing scenarios were not prominent. These differences were indication and patient specific. As

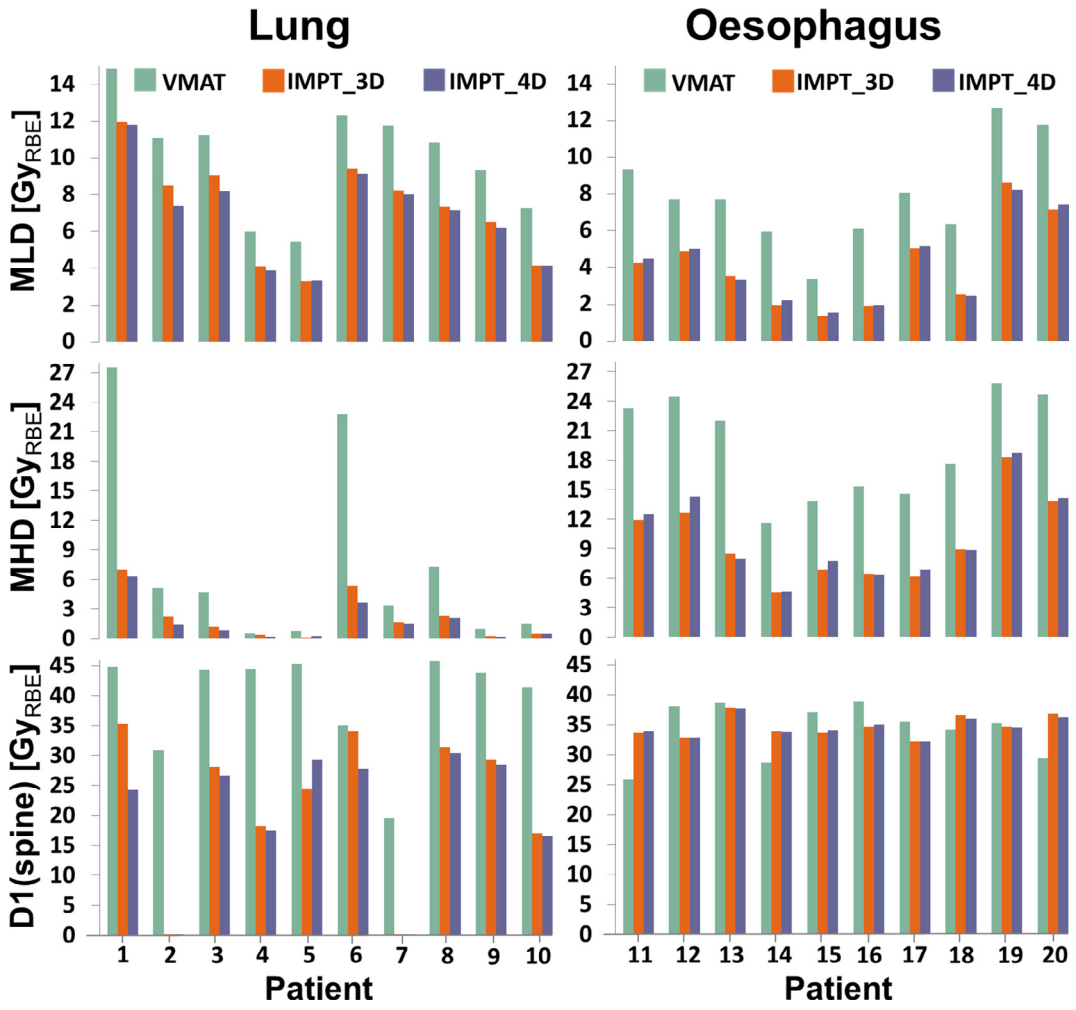


Fig. 2. Nominal treatment plan MLD, MHD, and D_1(spine) obtained with VMAT, IMPT_3D, and IMPT_4D for all lung and oesophageal cancer patients.

expected, the OARs closer to the CTV were particularly more affected by the different scenario variations simulated with the 4DREM. The largest variations were observed in IMPT_3D plans for D_1(spine) for lung cancer patient 1 (42.67 $\pm$ 2.38 Gy_{RBE}), and MHD for oesophageal cancer patient 11 (14.43 $\pm$ 1.51 Gy_{RBE}).

The explanation for insufficient target coverage in both IMPT_3D and IMPT_4D (for patients 8 and 19) was carefully explored (Fig. 4). As can be seen, variability in patient positioning (shoulder position) and anatomical changes (diaphragm position) during the course of treatment were the causes of these robustness failures.

Discussion

A treatment plan comparison between VMAT and different IMPT optimisation strategies for lung and oesophageal cancer patients is presented in this paper. Moreover, IMPT robustness evaluation was conducted, using longitudinal patient and machine specific information. The obtained results and the experience gained while conducting this research will be essential for the definition of an optimal IMPT planning protocol for patients with thoracic tumours. Retrospective patient data from our institute (and therefore representative of the patient population that will be treated in our proton therapy centre) was included. A considerable number of patients (20), with extensive numbers of 4DCTs were analysed.

As reported in previous studies, the rationale for IMPT has been demonstrated through the superior dosimetry obtained for OARs, when compared to VMAT, for both lung and oesophageal cancer patients [4–9]. Especially for MLD and MHD, a benefit of IMPT was confirmed for all patients [23,28,29]. Since we aim here to define a planning strategy for lung and oesophagus indications who might benefit from IMPT, for VMAT, only nominal dose distributions were analysed as reference, and any further VMAT robustness evaluation was out of the scope of this article [6,28,29,45,46]. Comprehensive robustness studies, such as the one presented in this paper for IMPT, have not yet been performed for VMAT.

Only 2/20 lung and oesophageal cancer patients revealed robustness shortcomings in the 4DREM for IMPT_3D. Both the planning protocol and subsequent delivery of IMPT_3D plans were clinically suitable for most patients. The cause of target coverage failure for both IMPT_3D and IMPT_4D plans of patients 8 and 19 were mostly due to a misposition of the shoulder and diaphragm baseline shifts relative to the planning situation, respectively. These deviations were consistent throughout all repeated 4DCTs. Shoulder position may be adjusted during verification of the patient positioning. However, the anatomical variation observed in the oesophagus case would require repeated 4D imaging, and a subsequent plan adaptation.

Especially for lung cancer patients, there was no direct correlation found between target characteristics (motion or volume) and plan robustness, whereas for the oesophagus indications, the most elongated volume (presenting a large CTV motion variation as well) was the one presenting coverage inadequacy. For the characteristics of the patient population included in this study, position

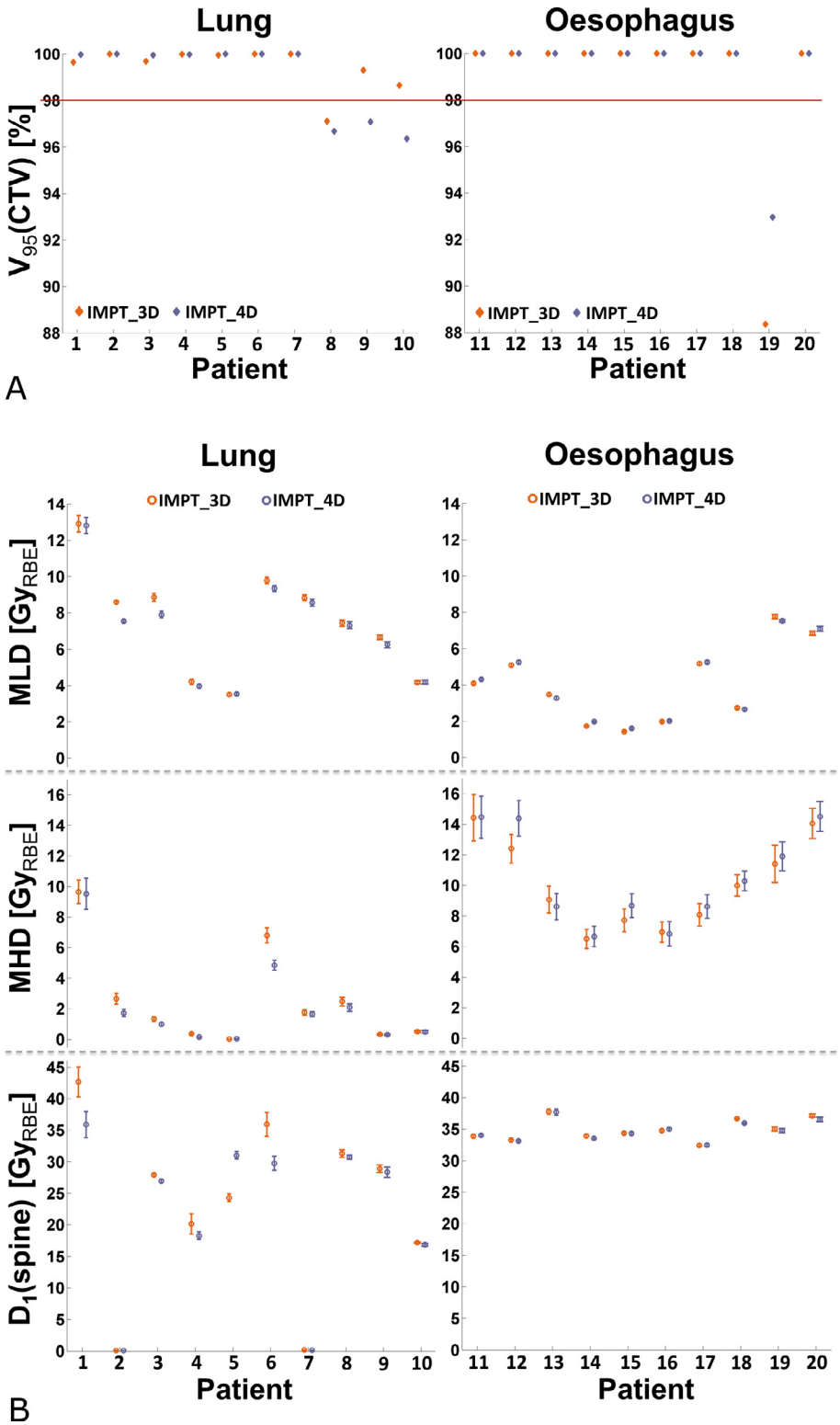


Fig. 3. Results of the 4DREM for the IMPT_3D and IMPT_4D plans created for all lung and oesophageal cancer patients. **A:** Target coverage (V_{95}(CTV)) on the 4DVwa$_{min}$ dose distribution. **B:** Mean ± SD MLD, MHD, and D_1(spine) over all 14 simulated treatment scenarios with the 4DREM.

and anatomical changes proved to influence the 4DREM results more pronouncedly than CTV motion and volume.

Motion amplitudes were not substantial for the lung targets (as high as 5.7 ± 1.3 mm). For the oesophageal cancer patients, these did not exceed 9.1 ± 1.5 mm. By chance, this population of lung and oesophageal cancer patients did not include 'big movers', which are not the most representative thoracic patients of the admissions in our radiotherapy clinic [43]. The target motion amplitudes reported in this study were quantified by the mean of all deformation vector lengths from DIR within the whole CTV (CTV of primary tumour and CTV of [multiple] pathological lymph nodes). Therefore, the motion amplitudes of different regions of the

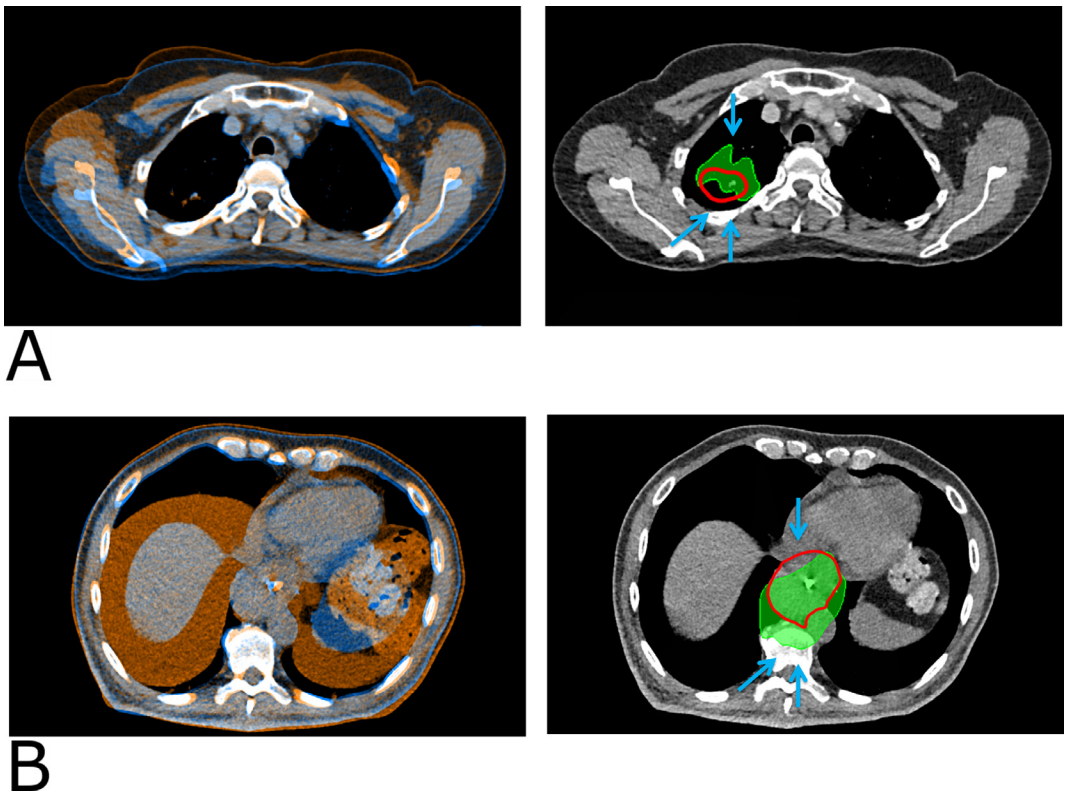


Fig. 4. A: Lung and **B:** oesophagus cases of failure in target coverage with the 4DREM (patients 8 and 19 respectively). Observed patient variation throughout the treatment course (left) and worst-case scenario dose distribution (in terms of V_{95}(CTV)) obtained with IMPT_3D (right). In red is the delineated CTV, the green line shows the 95 % isodose, and the light blue arrows represent the beam directions.

CTV, which can significantly differ, were not specifically considered. Naturally, also the motion values reported would change if other quantitative metrics would be used, such as maximum motion [13,42,47]. Our motion quantification can be considered underrating, reporting amplitudes that are exceeded at least in parts of the investigated volumes. Additionally, we report the mean CTV motion amplitudes over all weeks of treatment, and as verified by the SD, these values can pronouncedly vary between different 4DCTs, leading to higher or lower numbers. Enhancements in the motion evaluation procedure, as well as a follow-up study including patients with larger motion amplitudes, are work in progress, in order to confirm and translate our conclusions to the entire lung and oesophageal patient population.

Contrasting previous literature suggesting the supremacy of IMPT_4D over IMPT_3D for lung and oesophageal cancer patients [17–19], obtained differences in target coverage and OAR dose statistics for nominal and successive robustness evaluations between these optimisation approaches were minimal for our patients. Therefore, even when using a comprehensive tool for plan assessment for IMPT, such as the 4DREM, IMPT_4D did not show an advantage over IMPT_3D for patients with no considerable motion. The 4D robust optimisation method used in this work is the standard one implemented in RayStation [48], which can be considered quite simple, when compared to more advanced ones [49,50]. Specifically the method developed by Engwall et al. [50] for PBS-PT, besides incorporating in the optimisation process the organ motion, it also includes the delivery time structure. Unexpectedly, for two lung cancer patients (patients 9 and 10), superior target coverage robustness was found for IMPT_3D. For patient 9, besides displacement and shrinkage of the target, also the appearance of fluid in the right lung was verified in repeated 4DCTs acquired along the treatment. For patient 10, the body contour is altered in relation to the planning 4DCT, which is most likely caused by the arm position, and this is especially affecting the anterior-lateral beam. The exclusive use of an iCTV density override approach for IMPT_3D, accounting for the motion effects on the av_pCT within that target structure [51], proved to be beneficial to avoid underdosage in the beam path for these two lung cases. Doses were prescribed to the iCTV (on the av_pCT) and to the CTV (on the EE_pCT) for the IMPT_3D and the IMPT_4D plans, respectively. It is pertinent to point out that the iCTV was not defined by a union of the CTVs from all planning 4DCT phases, but instead was delineated by the radiation oncologist, following clinical protocol. For this reason, and to avoid any biased comparisons due to inconsistency in target volumes between IMPT_3D and IMPT_4D, the difference between the volume of the union of the CTVs of all planning 4DCT phases and the volume of the iCTV had to be inspected for all the twenty lung and oesophageal cancer patients included in our study. If the deviation between the iCTV and CTV union structures was pronounced, we asked an experienced radiation oncologist to re-delineate (if necessary) the iCTV and the CTV of the individual phases of the planning 4DCT. After these corrections, we obtained a new deviation between iCTV and CTV union of about 10 % or smaller, which is also a plausible threshold for inevitable inter-observer variability [52].

All produced IMPT treatment plans were evaluated using our 4DREM, which comprehensively inspects numerous substantial uncertainties (setup and range errors, machine errors, anatomy changes, breathing motion, and interplay effects). This method naturally incorporates a few limitations. First, probabilistic sampling of the setup error simulations, providing therefore unreproducible results. Second, reduced number of fractions (eight) assumed for

each scenario, even though this number already proved to be sufficient to represent the actual clinical delivery [25,40]. Third, besides the high complexity of 4D dose accumulations, with its intrinsic uncertainties, these calculations are also dependent on imperfect DIR algorithms. All DIRs were performed by using the phase specific delineated CTV as controlling ROI in order to drive each registered image pair deformation, and consequently improve 4D dose accumulation accuracy around that targeted area [53]. Fourth, the reliability of the 4DCT reconstruction algorithm. Inherent to the use of 4DCTs (due to their assumption of an average breathing cycle), no irregularity in breathing patterns was considered within one fraction of the 4DREM, although the starting phases within treatment delivery were varied.

Naturally, the ideal 4D evaluation procedure would include patient treatment-fraction specific imaging. In order to include as much anatomical information and variation as possible in the 4DREM, the planning 4DCT was also included in the underlying dose accumulations between the evaluated fractions, with an identical weight as the repeated 4DCTs. This is incoherent with the applied residual setup uncertainty margin of 2.0 mm, and one could argue that this 4DCT should therefore be removed from this evaluation. Anyhow, as already quantified and verified by us, for both IMPT_3D and IMPT_4D, the influence of the presence or not of the planning 4DCT in the 4DREM for the target and OARs is quite minimal. This is mainly due to the fact that all the 4DCTs included (multiple) were equally contributors to the treatment plan scenario dose. To overcome the limited number of 4DCTs available due to imaging dose restrictions, we aim to introduce 4DCBCT imaging soon in our proton clinic [54]. Additionally, a methodology for PBS-PT has been developed by Meijers et al. [55], assessing retrospectively after each delivered treatment fraction the deviations from the planned dose. Besides anatomical information provided by weekly 4DCTs, intra-fractional motion variability is also considered by recording patient breathing patterns and acquiring machine log files for each delivered treatment fraction. This tool has been used for a fraction-wise treatment evaluation of thoracic indications, providing longitudinal treatment course quality control and aiding in the clinical decisions for plan adaptation [56]. As future research, by reducing the applied margins and the magnitude of the IMPT robust optimisation parameters, we will proceed to investigate less conservative, and consequently more conformal planning strategies for thoracic IMPT. This can be done performing further plan comparison studies using our 4DREM.

Our IMPT_3D protocol proved to be adequate for this group of NSCLC and oesophageal cancer patients, as long as correct patient positioning is assured. Additionally, target motion variability, or anatomical changes, in general, along the beam path, throughout the treatment course, remain a concern, emphasizing the importance of daily volumetric imaging. In terms of plan robustness, the need of 4D optimisation, which implies considerably more manual work and optimisation time ($\approx$ 7 h vs. $\approx$ 3 h for 3D optimisation) within clinical workflow, was not justified by our results. This is, however, not completely unforeseen. The followed clinical methods for the application of both IMPT_3D and IMPT_4D approaches (e.g. iCTV and/or iCTV density override in IMPT_3D only) might influence this comparison. Moreover, to some extent, this can be explained by the limited to moderate motion present on the investigated patient cohort, which in most cases does not exceed the magnitude of the robustness settings used for plan optimisation. In any respect, the obtained findings allow us to choose a more efficient plan optimisation strategy for thoracic indications with similar characteristics as the ones included in this study.

Conflicts of interest statement

We have no conflicts of interest to disclose.

Acknowledgements

The authors would like to express special gratitude to the team of the UMCG involved in the delineations for this project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2021.01.014.

References

[1] Benthuysen LV, Hales L, Podgorsak MB. Volumetric modulated arc therapy vs. IMRT for the Treatment of distal esophageal cancer. Med Dosim 2011;36:404–9.
[2] Wijsman R, Dankers F, Troost EGC, Hoffmann AL, van der Heijden EHFM, de Geus-Oei L-F, et al. Comparison of toxicity and outcome in advanced stage non-small cell lung cancer patients treated with intensity-modulated (chemo-) radiotherapy using IMRT or VMAT. Radiother Oncol 2017;122:295–9.
[3] Paganetti H, Bortfeld T. Proton beam radiotherapy - the state of the art. New Technol Radiat Oncol 2005:3–540.
[4] Chuong MD, Hallemeier CL, Jabbour SK, Yu J, Badiyan S, Merrell KW, et al. Improving outcomes for esophageal cancer using proton beam therapy. Int J Radiat Oncol Biol Phys 2016;95:488–97.
[5] Chang JY, Jabbour SK, De Ruysscher D, Schild SE, Simone CB, Rengan R, et al. Consensus statement on proton therapy in early-stage and locally advanced non-small cell lung cancer. Int J Radiat Oncol Biol Phys 2016;95:505–16.
[6] Warren S, Partridge M, Bolsi A, Lomax AJ, Hurt C, Crosby T, et al. An analysis of plan robustness for esophageal tumors: comparing volumetric modulated arc therapy plans and spot scanning proton planning. Int J Radiat Oncol Biol Phys 2016;95:199–207.
[7] Welsh J, Gomez D, Palmer MB, Riley BA, Mayankkumar AV, Komaki R, et al. Intensity-modulated proton therapy further reduces normal tissue exposure during definitive therapy for locally advanced distal esophageal tumors: a dosimetric study. Int J Radiat Oncol Biol Phys 2011;81:1336–42.
[8] 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.
[9] 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.
[10] Lomax A. SFUD, IMPT, and Plan Robustness. Part. Radiother. Emerg. Technol. Treat. Cancer, 2016, p. 169–94. doi:10.1007/978-81-322-2622-2.
[11] Lomax AJ. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: the potential effects of calculational uncertainties. Phys Med Biol 2008;53:1027–42.
[12] 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;53:1043–56.
[13] 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.
[14] Knopf AC, Hong TS, Lomax A. Scanned proton radiotherapy for mobile targets - The effectiveness of re-scanning in the context of different treatment planning approaches and for different motion characteristics. Phys Med Biol 2011;56:7257–71.
[15] Zeng C, Plastaras JP, Tochner ZA, White BM, Hill-Kayser CE, Hahn SM, et al. Proton pencil beam scanning for mediastinal lymphoma: the impact of interplay between target motion and beam scanning. Phys Med Biol 2015;60:3013–29.
[16] Zeng C, Plastaras JP, James P, Tochner ZA, Hill-Kayser CE, Hahn SM, et al. Proton pencil beam scanning for mediastinal lymphoma: treatment planning and robustness assessment. Acta Oncol 2016;55:1132–8.
[17] Graeff C. Motion mitigation in scanned ion beam therapy through 4D-optimization. Phys Medica 2014;30:570–7.
[18] Liu W, Schild SE, Chang JY, Liao Z, Chang Y-H, Wen Z, et al. Exploratory study of 4D versus 3D robust optimization in intensity modulated proton therapy for lung cancer. Int J Radiat Oncol 2016;95:523–33.
[19] Yu J, Zhang X, Liao L, Li H, Zhu R, Park PC, et al. Motion-robust intensity-modulated proton therapy for distal esophageal cancer. Med Phys 2016;43:1111–8.

[20] Prayongrat A, Xu C, Li H, Lin SH. Clinical outcomes of intensity modulated proton therapy and concurrent chemotherapy in esophageal carcinoma: a single institutional experience. Adv Radiat Oncol 2017;2:301–7.
[21] Li H, Zhang X, Park P, Liu W, Chang J, Liao Z, et al. Robust optimization in intensity-modulated proton therapy to account for anatomy changes in lung cancer patients. Radiother Oncol 2015;114:367–72.
[22] 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.
[23] 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 2016;96:661–9.
[24] 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.
[25] 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.
[26] International Commission on Radiation Units and Measurements (ICRU). Prescribing, recording, and reporting photon beam therapy. ICRU Report 62 (Supplement to ICRU Report 50). Bethesda (MD): ICRU Publications; 1999.
[27] Wu AJ, Bosch WR, Chang DT, Hong TS, Jabbour SK, Kleinberg LR, et al. Expert consensus contouring guidelines for intensity modulated radiation therapy in esophageal and gastroesophageal junction cancer. Int J Radiat Oncol Biol Phys 2015;92:911–20.
[28] van der Laan HP, Anakotta RM, Korevaar EW, Dieters M, Ubbels JF, Wijsman R, et al. Organ sparing potential and inter-fraction robustness of adaptive intensity modulated proton therapy for lung cancer. Acta Oncol (Madr) 2019;58:1775–82.
[29] Anakotta RM, van der Laan HP, Visser S, Ribeiro CO, Dieters M, Langendijk JA, et al. Weekly robustness evaluation of intensity-modulated proton therapy for oesophageal cancer. Radiother Oncol 2020;151:66–72.
[30] Paganetti H. Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer. Phys Med Biol 2014;59:R419–72.
[31] Guerreiro F, Zachiu C, Seravalli E, Ribeiro CO, Janssens GO, Ries M, et al. Evaluating the benefit of PBS vs. VMAT dose distributions in terms of dosimetric sparing and robustness against inter-fraction anatomical changes for pediatric abdominal tumors. Radiother Oncol 2019;138:158–65.
[32] 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;89:424–30.
[33] Taylor PA, Kry SF, Followill DS. Pencil beam algorithms are unsuitable for proton dose calculations in lung. Int J Radiat Oncol Biol Phys 2017;99:750–6.
[34] Fredriksson A, Forsgren A, Hårdemark B. Minimax optimization for handling range and setup uncertainties in proton therapy. Med Phys 2011;38:1672–84.
[35] Chang JY. Guidelines and techniques for image-guided radiotherapy for non-small cell lung cancer. Image-Guided Radiother Lung Cancer 2008:25.
[36] Sloth Møller D, Knap MM, Nyeng TB, Khalil AA, Holt MI, Kandi M, et al. Difference in target definition using three different methods to include respiratory motion in radiotherapy of lung cancer. Acta Oncol (Madr) 2017;56:1604–9.
[37] Jin P, Machiels Mélanie, Crama KF, Visser J, van Wieringen N, Bel A, et al. Dosimetric benefits of midposition compared with internal target volume strategy for esophageal cancer radiation therapy. Int J Radiat Oncol 2019;103:491–502.
[38] Korevaar EW, Habraken SJM, Scandurra D, Kierkels RGJ, Unipan M, Eenink MGC, et al. Practical robustness evaluation in radiotherapy – A photon and proton-proof alternative to PTV-based plan evaluation. Radiother Oncol 2019;141:267–74.
[39] 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;58:7905–20.
[40] Lin L, Kang M, Huang S, Mayer R, Thomas A, Solberg TD, et al. Beam-specific planning target volumes incorporating 4D CT for pencil beam scanning proton therapy of thoracic tumors. J Appl Clin Med Phys 2015;16:281–92.
[41] van Herk M, Remeijer P, Rasch C, Lebesque JV. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. Int J Radiat Oncol Biol Phys 2000;47:1121–35.
[42] Sonke J-J, Rossi M, Wolthaus J, van Herk M, Damen E, Belderbos J. Frameless stereotactic body radiotherapy for lung cancer using four-dimensional cone beam CT guidance. Int J Radiat Oncol Biol Phys 2009;74:567–74.
[43] den Otter LA, Anakotta RM, Weessies M, Roos CTG, Sijtsema NM, Muijs CT, et al. Investigation of inter-fraction target motion variations in the context of pencil beam scanned proton therapy in non-small cell lung cancer patients. Med Phys 2020;47:3835–44.
[44] Weistrand O, Svensson S. The ANACONDA algorithm for deformable image registration in radiotherapy. Med Phys 2015;42:40–53.
[45] Chin E, Loewen SK, Nichol A, Otto K. 4D VMAT, gated VMAT, and 3D VMAT for stereotactic body radiation therapy in lung. Phys Med Biol 2013;58:749–70.
[46] Stroom J, Vieira S, Mateus D, Greco C, Fogliata A, Nicolini G, et al. On the robustness of VMAT-SABR treatment plans against isocentre positioning uncertainties. Radiat Oncol 2014;9:196.
[47] Lin L, Souris K, Kang M, Glick A, Lin H, Huang S, et al. Evaluation of motion mitigation using abdominal compression in the clinical implementation of pencil beam scanning proton therapy of liver tumors. Med Phys 2017;44:703–12.
[48] Bodensteiner D. RayStation: External beam treatment planning system. Med Dosim 2018;43:168–76.
[49] Graeff C. Robustness of 4D-optimized scanned carbon ion beam therapy against interfractional changes in lung cancer. Radiother Oncol 2017;122:387–92.
[50] 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.
[51] Botas P, Grassberger C, Sharp G, Paganetti H. Density overwrites of internal tumor volumes in intensity modulated proton therapy plans for mobile lung tumors. Phys Med Biol 2018;63:1–9.
[52] Segedin B, Petric P. Uncertainties in target volume delineation in radiotherapy - Are they relevant and what can we do about them?. Radiol Oncol 2016;50:254–62.
[53] Ribeiro CO, Knopf A, Langendijk JA, Weber DC, Lomax AJ, Zhang Y. Assessment of dosimetric errors induced by deformable image registration methods in 4D pencil beam scanned proton treatment planning for liver tumours. Radiother Oncol 2018;128:174–81.
[54] Niepel K, Kamp F, Kurz C, Hansen D, Rit S, Neppl S, et al. Feasibility of 4DCBCT-based proton dose calculation: An ex vivo porcine lung phantom study. Z Med Phys 2019;29:249–61.
[55] 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.
[56] 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.