



Technical note

Evaluation of alternative parameter settings for dose restoration and full plan adaptation in IMPT for prostate cancer

Thyrza Z. Jagt^{a,*}, Sebastiaan Breedveld^a, Mischa S. Hoogeman^{a,b}

^a Department of Radiotherapy, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

^b Department of Medical Physics & Informatics, HollandPTC, Delft, The Netherlands



ARTICLE INFO

Keywords:

Intensity-modulated proton therapy
Prostate cancer
Online-adaptive proton therapy
Online treatment planning
Dose restoration
Plan adaptation

ABSTRACT

Background/purpose: Intensity-modulated proton therapy is highly sensitive to anatomical variations. A dose restoration method and a full plan adaptation method have been developed earlier, both requiring several parameter settings. This study evaluates the validity of the previously selected settings by systematically comparing them to alternatives.

Materials/methods: The dose restoration method takes a prior plan and uses an energy-adaptation followed by a spot-intensity re-optimization to restore the plan to its initial state. The full adaptation method uses an energy-adaptation followed by the addition of new spots and a spot-intensity optimization to fit the new anatomy. We varied: 1) The margins and robustness settings of the prior plan, 2) the spot-addition sample size, i.e. the number of added spots, 3) the spot-addition stopping criterion, and 4) the spot-intensity optimization approach. The last three were evaluated only for the full plan adaptation. Evaluations were done on 88 CT scans of 11 prostate cancer patients. Dose was prescribed as 55 Gy(RBE) to the lymph nodes and seminal vesicles with a boost to 74 Gy(RBE) to the prostate.

Results: For the dose restoration method, changing the applied CTV-to-PTV margins and plan robustness in the prior plans yielded insufficient target coverage or increased OAR doses. For the full plan adaptation, more spot-addition iterations and using a different optimization approach resulted in lower OAR doses compared to the default settings while maintaining target coverage. However, the calculation times increased by up to 20 times, making these variations infeasible for online-adaptation.

Conclusion: We recommend maintaining the default setting for the dose restoration approach. For the full plan adaptation we recommend to focus on fine-tuning the optimization-parameters, and apart from this using the default settings.

Introduction

Intensity-modulated proton therapy (IMPT) is highly sensitive to anatomical variations, which can cause differences between the planned and delivered doses. A result of this can be insufficient target coverage during the course of treatment [1–4]. An approach to account for these anatomical variations, ensuring adequate target coverage in all treatment fractions while keeping the doses to the OARs acceptable, is to use adaptive treatment planning.

In previous work we have developed and evaluated two automated online-adaptive approaches, being a straightforward dose restoration approach [5] and a full plan adaptation method [6,7]. The dose restoration aims to restore a prior treatment plan to its intended dose

distribution, by adjusting the energy of each pencil-beam to match the water equivalent path length of the current anatomy. Assuming no differences in position and shape of the targets and OARs, i.e. only accounting for density changes along the pencil-beam paths, the dose-difference between the intended and the adapted plans is then minimized. The full adaptation approach on the other hand also adapts the plan for daily variations in the target shape and location. This is done by adding additional pencil-beams to a new optimization. Both approaches require a prior treatment plan to start the adaptation, hence their outcome depends on:

- (1) the applied CTV-to-PTV margins and plan robustness in the prior plan.

* Corresponding author at: Erasmus MC Cancer Institute, Department of Radiotherapy, Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands.

E-mail addresses: t.jagt@erasmusmc.nl (T.Z. Jagt), s.breedveld@erasmusmc.nl (S. Breedveld), m.hoogeman@erasmusmc.nl (M.S. Hoogeman).

<https://doi.org/10.1016/j.ejmp.2021.11.003>

Received 6 May 2021; Received in revised form 11 November 2021; Accepted 13 November 2021

Available online 23 November 2021

1120-1797/© 2021 Associazione Italiana di Fisica Medica. Published by Elsevier Ltd. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

The outcome of full plan adaptation furthermore depends on:

- (2) the spot-addition sample size prescribing how many spots are added in each iteration,
- (3) the stopping criterion prescribing how many iterations are performed, and
- (4) the optimization method used for the spot-intensity optimization.

In our previously published research [5–7], we have only partly investigated the effect of the choices made in that context. For future clinical implementation, it is important to verify our choices and investigate the effect that changing the settings can have on plan quality and calculation time. In this paper, we will therefore revisit the previously chosen parameter settings by systematically investigating and describing the effect of varying the parameter settings. In this study we address the four above mentioned parameters.

The systematic comparison between the default settings and varied settings will show if the results of our automated online plan adaptation methods can be further improved and at what cost. Using this we will make recommendations for the parameter settings for clinical implementation.

Methods and Materials

Patient data

For this study a patient cohort of 11 prostate cancer patients was used, including 8–10 weekly repeat CT scans (rCTs) per patient. The original planning CT scans were excluded as they contained contrast fluid. We instead used the first rCT as planning CT (pCT), keeping a total of 88 rCTs for evaluation. All patients gave consent after which they were enrolled in a phase II dose-escalation trial delivered with moderately hypo-fractionated IMRT at Haukeland University Hospital, Norway. Enrollment started in 2007. The trial had been approved by the local ethical committee REC West (the western Norway regional committee for medical and health research ethics; 2006-15727).

Treatment planning volumes and prescription

IMPT dose was prescribed according to a simultaneously-integrated boost scheme to be delivered with two laterally opposed beams in 37 fractions. The high-dose region was assigned 74 Gy(RBE) and the low-dose region 55 Gy(RBE), using a Relative Biological Effectiveness (RBE) of 1.1 [8]. To steer dose fall-off, an intermediate dose-region was defined as the 15 mm transition region between the high-dose region and low-dose region. This intermediate target should receive at least 55 Gy(RBE) and at most 74 Gy(RBE). On each scan two clinical target volumes (CTVs) were delineated. For the high-dose region, a CTV_{High} was defined as the prostate, and for the low-dose region a CTV_{Low} was defined as the lymph nodes and seminal vesicles. The rectum, bladder, small and large intestines, and the femoral heads were defined as OARs. Delineations of the target volumes, bladder and rectum were available in all rCTs. For some rCTs no manual delineations were available for the intestines and/or the femoral heads. For those scans, the missing delineations were obtained by rigidly propagating the delineations of the pCT onto the rCT.

Online-adaptive methods and default configuration

Treatment plan optimization of the prior plans

Both online-adaptive methods require a prior treatment plan created on the pCT to start the adaptation on the rCT. All prior treatment plans were generated using ‘Erasmus-iCycle’, our in-house developed treatment planning system (TPS) for fully automated plan generation [9,10], combined with the ‘Astroid’ dose engine [11]. Pencil-beams with energies ranging from 68.6 MeV to 250 MeV were available for planning,

with respective spot sizes (sigma in air at isocenter) ranging from 6 to 3.8 mm. The optimization applies the 2-phase ϵ -constraint method (2pec), which performs a prioritized optimization of the objectives, obtaining a Pareto optimal plan. The pencil-beams are selected through resampling, in which each iteration adds 3000 random spots to the optimization, and non-contributing spots get removed [10]. Prior plans were generated on the pCTs, applying margins not optimized for use throughout the treatment course but selected to provide an optimal start for the adaptation techniques. PTV_{Wide} structures were generated by expanding the CTV_{High} and CTV_{Low} structures of the pCT with margins of 7 and 10 mm, respectively. A sketch of the target delineations can be found in Fig. 1. Note that the PTVs used to generate the prior plans are used only as a starting point for the adaptation methods. These large default margins were selected to ensure sufficient spot coverage for most inter-fraction target deformations seen in the rCTs [6]. The prior plans were optimized non-robustly to obtain clinically acceptable target coverage for the PTV_{Wide} structures ($V_{95\%} \geq 98\%$) while at the same time aiming for $V_{107\%} \leq 2\%$. Here $V_{95\%}$ and $V_{107\%}$ are the percentages of the volumes receiving 95% and 107% of the prescribed dose, respectively. Dose to the OARs was further minimized according to the objectives shown in Table 1. More details on Erasmus-iCycle and the 2pec method can be found in [9,10,12–14].

The dose restoration method

The first method is a straightforward adaptation approach, which for each treatment fraction takes a prior plan optimized on the pCT and structure-wise restores that dose distribution on the new anatomy. Instead of requiring daily delineations of the targets and OARs, the method assumes no changes in shape and location of the targets and OARs, and hence uses the delineations of the pCT rigidly projected onto the rCT. To restore the dose distribution to its initial state, the restoration method adjusts the energy of the pencil-beams of the prior plan to the water equivalent path length of the current anatomy. This is followed by a fast spot-intensity re-optimization using the BOXCQP algorithm with a higher focus on the PTVs than on the OARs. Target coverage of the restored plans was evaluated on the CTV structures of the rCT. Further details on the BOXCQP algorithm and the restoration method can be found in [5,15,16].

The full plan adaptation method

The second method is a more elaborate approach, which for each treatment fraction takes the prior plan optimized on the pCT and re-optimizes this to fit the new daily anatomy. Using the delineations of the rCT, the method uses the prior treatment plan as a starting point to generate a new plan for the treatment fraction. The full adaptation aims

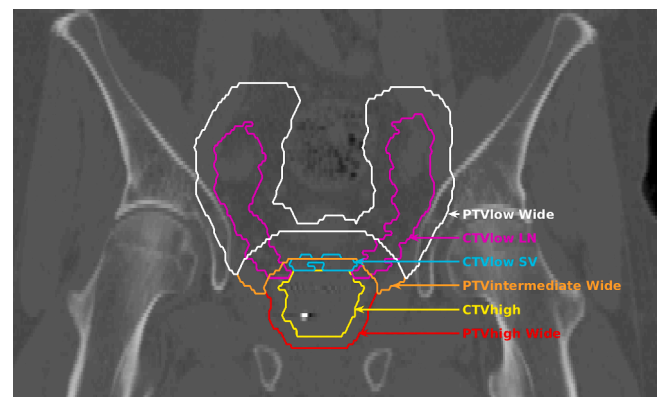


Fig. 1. A sketch of the different target definitions of the default prior plan in a coronal view.

Table 1

The ‘wishlist’ with planning constraints and objectives used for automated IMPT plan generation for prostate cancer patients. Constraints will always be met. The priority numbers of the objectives indicate the order in which objectives are to be optimized. A low number corresponds to a high priority. An artificial PTV-intermediate was generated as the 15 mm transition between the high-dose PTV and the low-dose PTV, included to steer the dose fall-off. Conformity rings were created around the PTV-high and the PTV-full (=PTV-high + PTV-low), to steer the dose-fall-off around the targets. The ring around the PTV-high was 15 mm, excluding the PTV-intermediate.

Constraints	Structure	Type	Limit
	PTV-high	Minimum	0.97×74 Gy (RBE)
	PTV-intermediate	Minimum	0.99×55 Gy (RBE)
	PTV-low	Minimum	0.99×55 Gy (RBE)
Objectives			
Priority	Structure	Type	Limit
f_1	PTV-high	Maximum	1.06×74 Gy (RBE)
f_2	PTV-intermediate	Maximum	1.06×74 Gy (RBE)
f_3	PTV-low	Maximum	1.06×55 Gy (RBE)
f_4	Conformity ring PTV-high	Maximum	1.07×74 Gy (RBE)
f_5	Conformity ring PTV-full 0–10 mm	Maximum	1.07×55 Gy (RBE)
f_6	Conformity ring PTV-full 10–15 mm	Maximum	0.90×55 Gy (RBE)
f_7	Rectum	Maximum	1.02×74 Gy (RBE)
f_8	Bladder	Maximum	1.02×74 Gy (RBE)
f_9	Body	Maximum	1.06×74 Gy (RBE)
f_{10}	Femoral heads	Maximum	50 Gy(RBE)
f_{11}	Rectum	Mean	0 Gy(RBE)
f_{12}	Small and large intestines	Mean	0 Gy(RBE)
f_{13}	Bladder	Mean	0 Gy(RBE)
f_{14}	Femoral heads	Mean	0 Gy(RBE)
f_{15}	All conformity rings	Mean	0 Gy(RBE)
f_{16}	All conformity rings	Maximum	0 Gy(RBE)
f_{17}	Total spot weight	Sum	0 Gp

Abbreviations: PTV = planning target volume; Gp = Gigaprotons.

to obtain adequate target coverage on PTV_{OAPT} structures (Online-Adaptive Proton Therapy). These structures were created by expanding the CTV_{High} and CTV_{Low} structures delineated on the rCT by 2 and 3.5 mm, respectively [6]. The PTV_{OAPT} is designed only to account for uncertainties in the daily delineations and intra-fraction uncertainties. As the idea behind the online plan adaptation is that the plan is adapted on each treatment day, other variations are eliminated, and thereby smaller margins can be used. Adaptation starts with an energy-adaptation of the pencil-beams of the prior plan followed by the addition of 2500 new spots to the target area using a latin hypercube sampling. A spot-intensity optimization based on the rCT delineations is then performed using the Reference Point Method. More details on the adaptation method can be found in [6,7].

The Reference Point Method (RPM)

The RPM is an optimizer which obtains a Pareto optimal treatment plan in a single optimization, making it a simpler and faster method than the 2pec optimizer [17–21]. Parameter tuning for the RPM was done using training plans generated with Erasmus-iCycle. Tuning aims to minimize the differences between the prior plans and the plans obtained

when optimizing the same pencil-beams with the RPM, thereby training the RPM to make similar trade-offs as the 2pec optimizer.

In the tuning of the RPM-parameters the differences in the objective values between predefined benchmark plans and plans obtained using the RPM with the tuned parameters is iteratively minimized. Tuning was stopped after 20 iterations or when all differences complied the demands of the tuning protocol in Table 2. The values of the total spot-weights were scaled down by a factor $1e^5$ to avoid too much focus on that objective.

Tuning was done using three-fold cross validation, where one third of the patients was used for tuning and the remaining two thirds for testing. Evaluation was done on all folds simultaneously, i.e. 176 plans (two per scan). Target coverage was evaluated on the PTV_{OAPT} structures.

Variations on the default parameter settings

In this study we have investigated the effects of varying four of the default parameter settings. While altering one of these settings, all others were kept at default, as shown in Table 3.

1. *Parameter 1 - The margin settings of the prior plan:* Both adaptation methods require a prior plan to start. In previous work we have generated these prior plans with the inclusion of additional margins around the target structures. These margins were added to ensure sufficient spot coverage for most target deformations seen in the rCTs. For the first setting we investigated what happens if we do not include such additional margins. Instead of using these ‘wide’ prior plans generated based on the PTV_{Wide} structures, ‘small’ prior plans were generated using the margin definition of the PTV_{OAPT} structures. Second, we investigated what happens if we account for even more possible anatomical and positional variations. To this end, ‘robust’ prior plans were generated based on the CTV structures of the pCTs. Safety margins of 4 and 8 mm were added to the CTV_{High} and CTV_{Low}, respectively [7], and robust optimization was applied on these expanded regions using the ‘minimax’ worst-case approach

Table 2

The ‘protocol’ which was used in the RPM-parameter tuning. Includes the user preferences for the population-based differences between the benchmark plans and the RPM generated plans (benchmark – RPM). The PTV conformity rings include the conformity ring PTV-high, the conformity ring PTV-full 0–10 mm and the conformity ring PTV-full 10–15 mm. Depending on the unit of the plan objective, the lower bounds is either in Gy or in %-point.

Plan objective	Measure	Lower bound
PTV D_{max} (Gy)	Median	–0.5
	1st quartile	–0.3
	5th percentile	–0.1
PTV conformity rings D_{max} (Gy), D_{mean} (Gy)	Median	–1
	1st quartile	–0.8
	5th percentile	–0.5
Rectum D_{max} (Gy), D_{mean} (Gy), $V_{75Gy(RBE)}$ (%), $V_{60Gy(RBE)}$ (%), $V_{45Gy(RBE)}$ (%)	Median	–0.8
	1st quartile	–0.5
	5th percentile	–0.3
Bladder D_{max} (Gy), D_{mean} (Gy), $V_{65Gy(RBE)}$ (%)	Median	–0.8
	1st quartile	–0.5
	5th percentile	–0.3
Bladder $V_{45Gy(RBE)}$ (%)	Median	–1
	1st quartile	–0.5
	5th percentile	–0.3
Small and large intestines D_{mean} (Gy)	Median	–0.8
	1st quartile	–0.5
	5th percentile	–0.3
Femoral heads D_{max} (Gy), D_{mean} (Gy)	Median	–1.5
	1st quartile	–1
	5th percentile	–0.5
Total spot-weight (MU)	Median	–2
	1st quartile	–1.5
	5th percentile	–1

Table 3

A summary of the varied parameter settings, including default values and investigated variations.

Parameter	Default	Variations
Margin settings of the prior plan	Wide prior plans, optimized non-robustly on the PTV _{Wide} structures. PTV _{Wide} : CTV _{High} + 7 mm, CTV _{Low} + 10 mm	<ol style="list-style-type: none"> Use small prior plans, optimized non-robustly on the PTV_{OAPT} structures. PTV_{OAPT}: CTV_{High} + 2 mm, CTV_{Low} + 3.5 mm Use robust prior plans, optimized robustly on expanded structures. Robust: CTV_{High} + 4 mm, CTV_{Low} + 8 mm optimized robustly using 6 mm setup robustness and 3% range robustness.
Spot-addition sample size	2500 spots.	Investigate 500, 1500, 3500 and 4500.
Stopping criterion	Stop after 1 iteration of spot addition and optimization.	Continue adding new spots until no improvement of greater than 3% is observed for the cost functions.
Optimization approach	RPM	2pec

[22–24]. Nine scenarios were optimized simultaneously (nominal, \pm setup, \pm range). A setup robustness of 6 mm and a range robustness of 3% were used. In previous work, these settings were found to yield adequate target coverage without online adaptation for all rCTs of at least 90% of the patients [7].

- Parameter 2 – The spot-addition sample size:** In the full adaptation method new spots are iteratively added to the target region. Previously, we found for the prostate cancer patient dataset used in this study that adding 2500 spots and optimizing the intensities once, adequate target coverage could be obtained for all rCTs [6]. For this parameter, we investigated the effect of varying the spot-addition sample size on calculation time, target coverage and OAR sparing. The spot-addition sample size was varied from 500 to 4500 in steps of 1000.
- Parameter 3 – Stopping criterion:** For the prostate cancer dataset used in this study, adequate target coverage was obtained for all rCTs using only a single iteration [6]. By varying the stopping criterion we evaluated the effect of allowing the full adaptation to continue to iteratively add new spots and optimize the spot-weights until the improvements of the cost functions are all below 3%. With this approach no time constraints were given, thereby allowing for an evaluation of the final trade-offs made by the RPM optimization. In between spot-addition iterations non-contributing spots were removed.
- Parameter 4 – Optimization approach:** To evaluate the effect of the RPM, we replaced this optimization by the 2pec optimization. The 2pec method was applied while using only a single iteration of spot-addition, and while running the adaptation until no improvement larger than 3% was observed (like parameter 3).

A summary of the variations on the parameter settings we have investigated is provided in Table 3.

For all PTV definitions, the PTV_{Low} consists of the expanded CTV_{Low} excluding the 15 mm PTV_{intermediate} transition region around the PTV_{High}. From here on we will denote the combination of the CTV_{High} and the CTV_{Low} as the CTV.

Evaluation of the investigated variations

For all variations of the parameter settings, we report the $V_{95\%}$ success rates (percentage of plans achieving $V_{95\%} \geq 98\%$ for all targets) and

the $V_{107\%}$ success rates (percentage of plans achieving $V_{107\%} \leq 2\%$ for all targets). To compare sparing of the healthy tissues we evaluated dose to the rectum, bladder and whole body. For all investigated variations we show the rectum $V_{75Gy(RBE)}$, $V_{60Gy(RBE)}$, $V_{45Gy(RBE)}$, D_{mean} and $D_{2\%}$, the bladder $V_{65Gy(RBE)}$, $V_{45Gy(RBE)}$, D_{mean} and $D_{2\%}$, and the whole body $V_{10Gy(RBE)}$ and $D_{2\%}$. Here $V_{xGy(RBE)}$ is the percentage of the volume receiving x Gy(RBE), D_{mean} is the average dose, and $D_{2\%}$ is the dose delivered to 2% of the volume.

All calculations were performed on a dual Intel Xeon E5-2690 server with 128 GB RAM.

Statistical analysis

Wilcoxon signed-rank tests were performed using MATLAB (Mathworks version 2017a) to evaluate the differences in OAR dose between the varied parameter values and the default parameter value. A p-value < 0.05 was considered to be statistically significant.

Results

Parameter 1: Changing the margin settings of the prior plan

All prior plans achieved adequate target coverage for their respective target regions. Table 4 shows the number of energy layers and pencil-beams in the different prior plans.

Parameter 1a: Changing the margin settings of the prior plan in the restoration method

Applying plan restoration using the ‘wide’ (default), ‘small’ and ‘robust’ prior plan resulted in $V_{95\%}$ success rates of 100%, 64.8% and 100%, respectively. For the $V_{107\%}$ the success rates were 76.1%, 67.0% and 0.0%, respectively. With the wide prior, $V_{107\%}$ values up to 3.7% were obtained for the CTV_{Low}. With the small prior, $V_{107\%}$ values up to 5.4% were obtained for the CTV_{High} of the rCTs, and values up to 3.4% were obtained for the CTV_{Low} of the rCTs. Using the robust prior resulted in $V_{107\%}$ values for the CTV_{Low} between 30.5% and 53.8% (average value of 42.7%).

Fig. 2 shows the achieved values of the evaluation criteria for the three methods, with the first two columns depicting the target coverage success rate.

Largest differences between the restorations obtained with the different priors were observed for the $V_{45Gy(RBE)}$ bladder and $V_{45Gy(RBE)}$ rectum. Using the small prior instead of the wide default prior resulted in a decrease in the median values of 22.6%-point and 20.6%-point, respectively. Changing from the wide prior to the robust prior increased these median values by 7.7%-point and 9.7%-point for bladder and rectum, respectively. Overall, the differences in obtained values were larger between the small and wide prior than between the wide and robust prior. For all evaluation criteria of the OARs, the differences between using the small and default wide prior, as well as the differences between using the wide and robust prior were statistically significant.

Parameter 1b: Changing the margin settings of the prior plan in the full adaptation method

Full adaptation resulted in a success rate of 100% for all three

Table 4

Average number of energy layers and pencil-beams in the different prior plans. The range is given in the parenthesis.

	Wide prior (default)	Small prior	Robust prior
# energy layers	84 (65–96)	73 (59–80)	84 (65–95)
# spots	1697 (1508–1856)	1268 (1129–1358)	1682 (1560–1838)

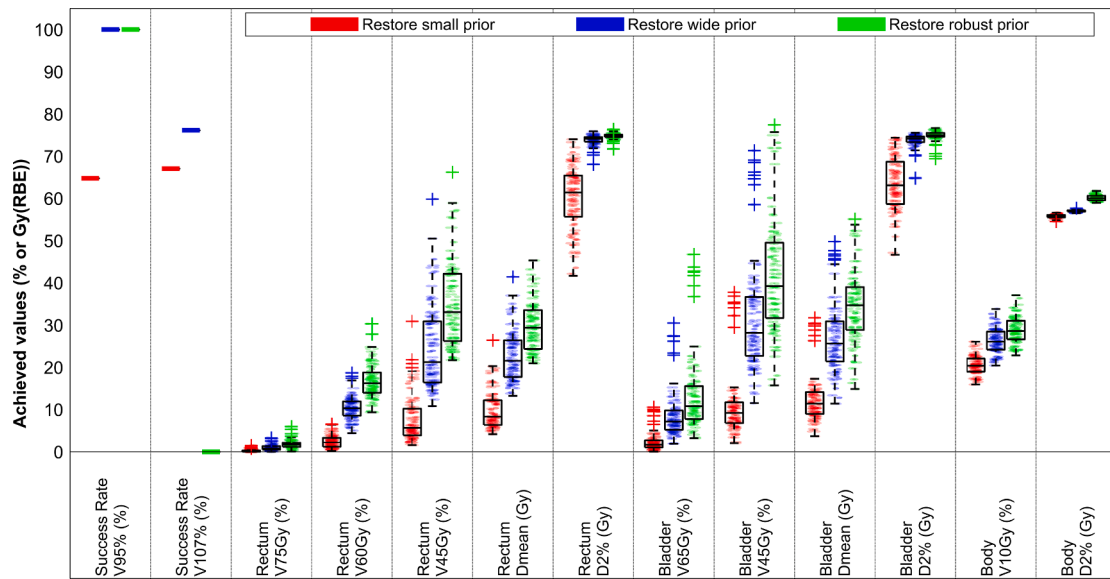


Fig. 2. Boxplot depicting the obtained dosimetric parameter values for target and OARs using the dose restoration method with the three different margin settings in the priors. The first two columns use horizontal lines to denote the percentage of plans that meet the target constraints for all target regions.

evaluated priors when adding 1 iteration of spots. Taking the wide and robust priors always resulted in $V_{107\%} \leq 2\%$. When using the small prior, 172/176 plans obtained a $V_{107\%} \leq 2\%$.

Fig. 3 shows the achieved values of the evaluation criteria after full adaptation using the three different priors. The first columns depict the target coverage success rate.

Using the small prior plans mostly resulted in lower dosimetric parameter values than taking the wide (default) prior plan, with a maximum improvement of the median value of the bladder $V_{45Gy(RBE)}$ of 5.2%-point. For all evaluation criteria of the OARs the differences between using the small and wide prior were statistically significant.

Using the robust instead of the wide prior resulted in smaller differences, with a maximum increase of 3.2%-point in median value of both the bladder $V_{45Gy(RBE)}$ and the bladder D_{mean} . Except for the rectum $V_{75Gy(RBE)}$ and $V_{60Gy(RBE)}$ and bladder $D_{2\%}$, differences between using the wide and robust prior were statistically significant.

Parameter 2: Changing the spot-addition sample size

Running the full adaptation with a spot-addition sample size of 500 yielded adequate target coverage for all scans, and a $V_{107\%} \leq 2\%$ for 155/176 plans. Running the full adaptation with a single iteration of spot addition using sample sizes of 1500, 2500, 3500, or 4500 yielded adequate target coverage for all scans, as well as yielding $V_{107\%} \leq 2\%$ for all scans.

Fig. 4 shows the achieved values of the evaluation criteria after full adaptation with a single iteration of spot addition, using the different sample sizes. High dose values were reduced when using a larger sample size, but the mean dose to the bladder remained similar. The differences between using a spot-addition sample size of 500, 1500, 3500 or 4500 and taking the default sample size of 2500 were almost all statistically significant.

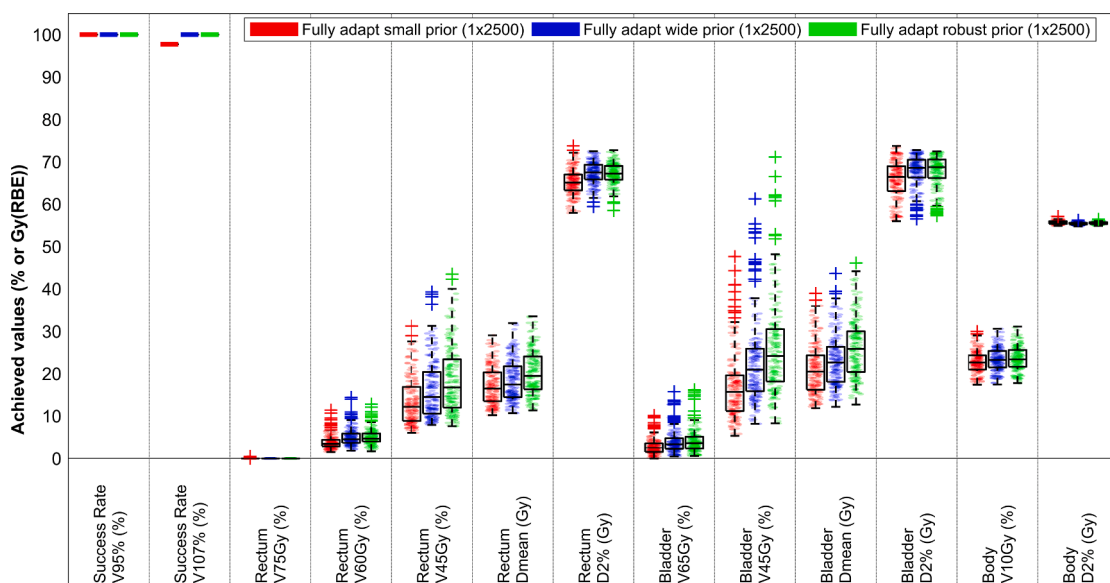


Fig. 3. Boxplot depicting the obtained dosimetric parameter values using the full adaptation with the three different margin settings in the priors. Adaptation included one iteration of spot addition and spot-intensity optimization. The first two columns use horizontal lines to denote the percentage of plans that meet the target constraints for all target regions.

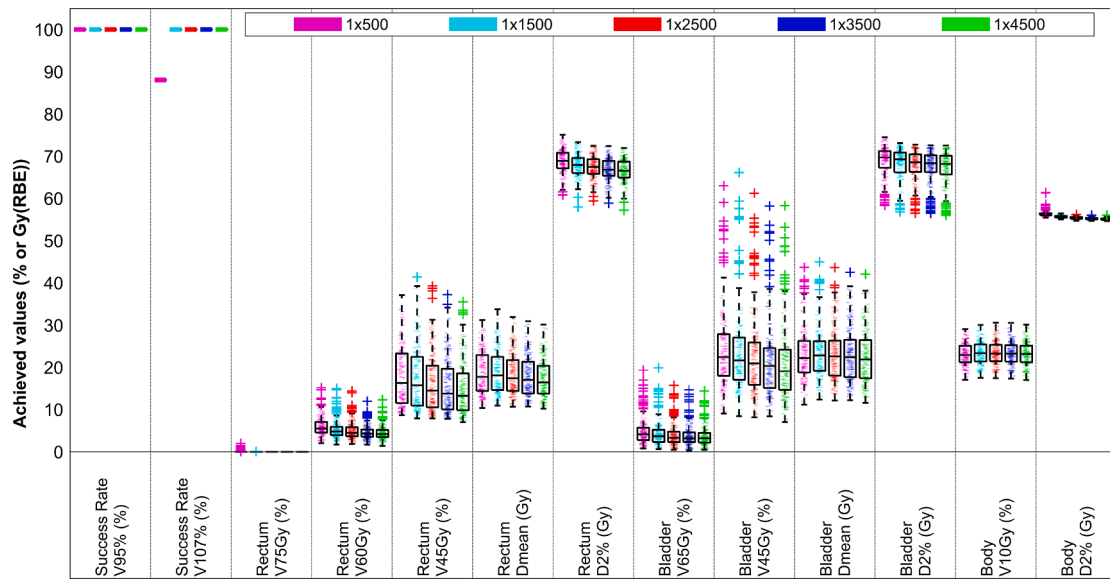


Fig. 4. Boxplot depicting the obtained dosimetric parameter values using the full adaptation when using different spot-addition sample sizes. Adaptation included one iteration of spot addition and spot-intensity optimization. The first two columns use horizontal lines to denote the percentage of plans that meet the target constraints for all target regions.

Parameters 3 & 4 - changing the stopping criterion and the optimization approach

Running the full adaptation method with a single iteration of spot addition and weight optimization resulted in adequate target coverage for all plans both when using the RPM (default) as optimizer and when using the 2pec optimizer. All plans also achieved $V_{107\%} \leq 2\%$. The same was obtained when running the adaptation until no improvements larger than 3% were observed. The RPM required on average 6 iterations of spot addition and spot-weight optimization (range 3–12) to meet this stopping criterion. Using the 2pec optimizer required on average 7 iterations (range 4–10).

Fig. 5 shows boxplots depicting the achieved values for the evaluation criteria when running the adaptation with a single iteration of spot addition for both optimizers (1x2500), as well as running the adaptation

until no improvements larger than 3% were observed for both optimizers (nx2500).

For a single iteration, the 2pec optimizer mostly resulted in lower evaluation values than the RPM optimizer. Improvements in median values of up to 7.5% were observed for the $V_{45Gy(RBE)}$ and D_{mean} of both the rectum and bladder. A higher value for the 2pec optimizer was only observed for the whole body $D_{2\%}$.

The differences between the optimizers for multiple iterations were similar to those observed when running a single iteration. The largest improvement when using 2pec instead of RPM was seen for the mean dose to the bladder, with a reduction of 13.1 Gy(RBE) in median. Again, the whole body $D_{2\%}$ was slightly higher for the 2pec approach.

Comparing the results obtained with a single iteration and multiple iterations, it can be seen that applying the 2pec optimizer with a single iteration already achieved lower doses than applying the RPM optimizer

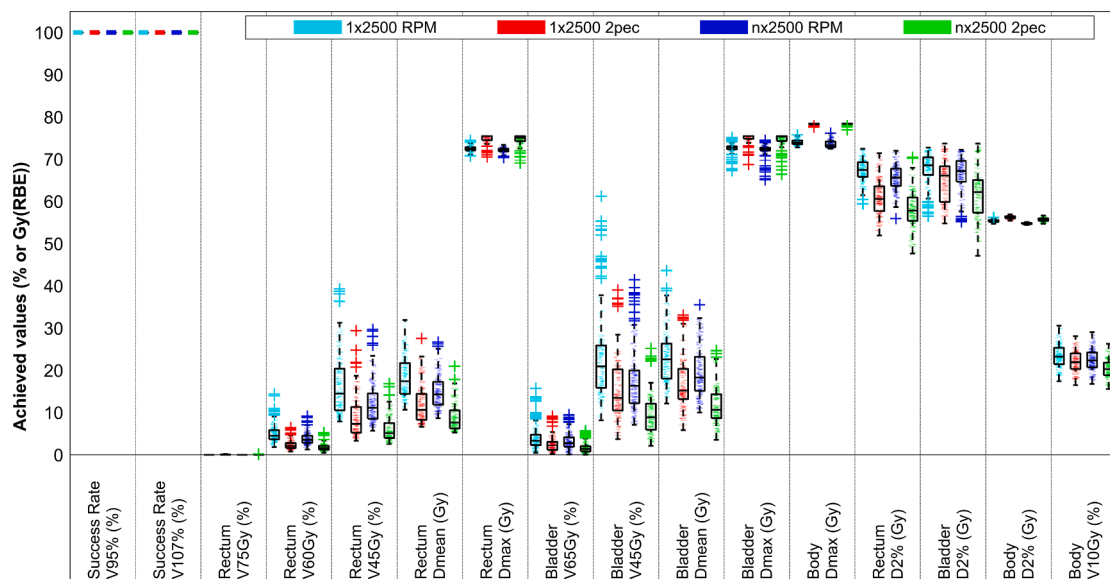


Fig. 5. Boxplot depicting the obtained dosimetric parameter values when using either the 2pec optimizer or the RPM. Results are shown for when running the full adaptation with a single iteration of spot addition (1x2500), and for when running until no improvement larger than 3% is observed (nx2500). The first two columns use horizontal lines to denote the percentage of plans that meet the target constraints for all target regions.

with n iterations. Using the 2p ϵ c optimizer in more iterations further reduced the dosimetric parameter values.

Compared to the default setting of a single iteration using the RPM, all variations yielded statistically significant differences for the evaluation criteria of the OARs. The only exception was the difference for the rectum $V_{75\text{Gy}(RBE)}$ between a single and n iterations of using the RPM.

Calculation times

Table 5 shows the average calculation times of the different steps in the restoration method when using the three different prior treatment plans.

Table 6 provides the average calculation times of the different steps in the full adaptation when using the three different prior treatment plans, when using the different spot-addition sample sizes in adapting the default prior plan, and when changing the stopping criterion to more than one iteration and changing the optimization approach. The times required for the energy-adaptation steps are omitted, as these can be found in Table 5.

Discussion

In this study we have investigated and presented variations on the default parameter settings for two previously developed online adaptation methods. The results were compared to those obtained using the previously chosen default settings. We investigated the effect of changing the margins and robustness settings of the prior plan for both the restoration and full adaptation method. For the full adaptation method, we also investigated the effect of changing the spot-addition sample size, removing the optimization time constraints, and changing the optimization approach.

Changing the margins and robustness settings of the prior treatment plan had a large influence on the results of the restoration method in terms of target coverage and OAR sparing. As this method aims to return to the prior dose distribution, this result was expected. Using the wide prior the target coverage success rate was 100%. Applying the small prior resulted in inadequate target coverage in some rCTs. For both the small prior and the wide prior the success rate of the $V_{107\%}$ was below 100%. However, as these high doses were within the CTV, these values were deemed acceptable. Using the robust prior plan led to a $V_{107\%} > 2\%$ for all plans in the low-dose region and higher values for the OARs than using the default prior. The high $V_{107\%}$ values can be explained by the proximity of the optimized high/intermediate-dose region to the evaluated low-dose region. High $V_{107\%}$ values were also observed for the robust treatment plans in the nominal scenario on the pCTs. The extra dose is needed to obtain sufficient target coverage in the error scenarios. Calculation times only varied slightly, decreasing when using a prior plan with fewer energy layers and spots (Tables 4 and 5).

For the full adaptation approach, smaller differences were observed. Adequate target coverage was obtained using all prior plans for all targets. For 4/176 plans using the small prior resulted in $10\% > V_{107\%} > 2\%$. As the high dose was focused in the CTV, these values were still deemed clinically acceptable. Considering the OAR doses, Fig. 3 shows that including smaller margins and less robustness in the prior plan leads to more OAR sparing. An explanation for this is that

pencil-beams placed in a smaller target will have a greater distance to the OARs, thereby increasing OAR sparing.

We found that adding only 500 new spots can already yield adequate target coverage for scans in this dataset. Increasing the spot-addition sample size showed to have most effect on dose conformity; reducing the volumes receiving high doses. Less effect was seen on the areas receiving low dose and on the mean dose. Calculation times proved not to increase linearly with the sample size (Table 6).

Changing the optimizer from RPM to 2p ϵ c showed to have large effects on the quality of the generated treatment plans. Most evaluation criteria obtained lower values with the 2p ϵ c, both when applying a single iteration of spot-addition and when adding spots until no improvement larger than 3% was observed. Only for the high doses the RPM achieved lower values. It should be noted though, that all plans that were generated using the RPM and 2p ϵ c are Pareto optimal. While Fig. 5 seemingly shows that the RPM plans are of lower quality than the 2p ϵ c plans, evaluation of other dosimetric parameters show the opposite. For example, the maximum dose in the rectum, bladder and whole body was lower when using the RPM. We furthermore found that the RPM obtained a lower mean dose in the femoral heads, while the mean dose for rectum and bladder was higher. This shows that the trade-offs between the RPM and the 2p ϵ c optimizer were different. Another consequence of this is that the results of 2p ϵ c cannot be replicated by increasing the number of RPM iterations.

This result indicates that the RPM-parameter tuning, which aimed at forcing the RPM to make similar trade-offs as the 2p ϵ c optimizer, should be improved. Even when optimizing the same set of pencil-beams in the single iteration test, the RPM and 2p ϵ c optimizers still made different trade-offs. Interestingly, when taking the spots selected within a 2p ϵ c optimization (as is done during the tuning) the RPM does yield similar results to the 2p ϵ c optimizer. This indicates that the RPM-parameter tuning may have been too focused on the specific pencil-beam placements selected through resampling in the 2p ϵ c optimization. We conclude that the obtained parameters are not sufficiently robust to handle different pencil-beam sets. This may be improved by performing the RPM-parameter tuning on treatment plans with various pencil-beam sets or by including multiple plans per CT scan. Possibly the adaptation itself could also be included in the tuning, minimizing the difference between adapted plans and fully optimized plans.

Besides these changes, it should be noted that the tuning protocol (Table 2) might not yet be optimal and that training has been done on only a small number of scans. For this study we tuned using three-fold cross validation, training only on a third of the data. The training sets consisted of 30, 34 and 35 scans, respectively. For all variations it should furthermore be noted that the effect they yield on the output plans are likely to be highly dependent on the investigated dataset. Considering the running times, Table 6 shows that using the 2p ϵ c approach takes longer than using the RPM. This was expected, as the 2p ϵ c approach optimizes the objectives sequentially in a prioritized way, while the RPM only requires a single optimization. Running the method until no improvement larger than 3% is observed takes up to 20 times longer. This can be explained by the optimization time depending on the number of pencil beams, which increases with each spot-addition iteration.

While full plan adaptation requires more up-front parameter choices

Table 5

Average calculation times for the simple dose restoration. The total running time is excluding the final dose calculation.

		Initialization (minutes)	Energy- adaptation (seconds)	Calculation of dose deposition matrix (minutes)	Weight optimization (seconds)	Total running times (minutes)
Default	Wide prior	1.1 (0.8–1.6)	6.8 (5.5–8.0)	1.3 (1.1–1.5)	7.5 (5.9–9.9)	2.7 (2.2–3.2)
Change the margin settings of the prior	Small prior	0.9 (0.7–1.1)	6.3 (5.1–7.8)	1.1 (0.9–1.2)	4.4 (3.5–5.6)	2.2 (1.8–2.5)
	Robust	0.9 (0.8–1.1)	6.3 (4.8–8.3)	1.3 (1.0–1.5)	7.1 (6.0–9.9)	2.5 (2.0–2.8)
	prior					

Table 6

Average calculation times for the full adaptation method. The intermediate dose calculation to determine the dose deposition matrices of the new spots is included in the spot addition step. The total running time includes everything starting from the initialization to the optimization, but is excluding the final dose calculation.

		Beam setup (seconds)	Spot addition (minutes)	Optimization (seconds)	Total running time (minutes)
Default	Wide prior + 1x2500	1.3 (0.9–5.8)	2.2 (1.3–5.1)	47.1 (39.6–57.4)	5.6 (4.1–9.1)
Change the margin settings of the prior plan	Small prior + 1x2500	1.2 (0.9–1.8)	2.1 (1.2–5.1)	39.0 (30.0–47.5)	5.1 (3.7–8.4)
	Robust prior + 1x2500	1.2 (0.8–1.8)	2.1 (1.2–5.1)	47.0 (40.2–59.6)	5.5 (3.9–8.7)
Change the spot-addition sample size	Wide prior + 1x500	1.3 (0.9–4.3)	1.2 (0.8–1.8)	25.4 (20.7–31.2)	4.1 (3.1–5.6)
	Wide prior + 1x1500	1.2 (0.9–2.5)	1.6 (1.0–2.4)	34.7 (27.6–42.4)	4.7 (3.5–5.7)
	Wide prior + 1x3500	1.2 (0.8–1.9)	2.3 (1.2–4.1)	66.0 (55.3–80.0)	5.9 (4.3–8.0)
	Wide prior + 1x4500	1.2 (0.9–1.7)	2.6 (1.5–5.9)	90.1 (71.8–106.6)	6.6 (5.0–10.1)
Change the stopping criterion & optimization approach Prior = wide	1x2500 with 2pec	1.2 (0.9–1.9)	2.2 (1.3–5.1)	187.2 (161.6–214.4)	7.8 (6.3–11.1)
	nx2500 with RPM	1.2 (0.9–1.8)	14.3 (6.1–28.6)	1112.8 (372.9–3305.2)	36.2 (16.0–88.0)
	nx2500 with 2pec	1.3 (0.9–6.7)	14.5 (9.1–30.0)	5356.9 (2598.1–10123.9)	107.1 (55.9–193.4)

to be made than dose restoration, it is also more robust against changes in these choices. As the dose restoration method only depends on the margin settings of the prior plan, the outcome of the restoration is greatly affected by a change in input. The individual choices made for the full plan adaptation showed to have less influence. While we have investigated the variations separately, they will probably be correlated in their effect on the adapted plans. For example, we have observed that when using full adaptation with a sample size of 2500, a smaller prior plan can be used while still yielding adequate treatment plans. We have also observed that when taking the default wide prior plan, the sample size can be reduced to 500 while still obtaining adequate target coverage. However, combining a sample size of 500 with a smaller prior plan might not yield adequate results.

In other dose restoration approaches [25,26], new pencil-beams are used in the optimization instead of those from the prior plan as is done in our approach. In those methods the prior plan is only used for the actual dose distribution which is used to guide the restoration. The question which margins and robustness settings to use in those prior plans remains relevant, as not all uncertainties will be accounted for by the dose restoration.

For full plan adaptation, different techniques are described in literature. One method is to quickly generate an entirely new plan for the daily anatomy, optimizing a predefined set of planning objectives using a simplified optimization approach [27]. For such approaches the simplified optimization technique should be thoroughly compared to the original optimizer, as was done in our study by comparing RPM to 2pec. Another technique, described in [28,29], uses deformable image registration to obtain daily delineations. The method first calculates the dose for each pencil-beam on the new anatomy, and then performs a restoration by re-optimizing the intensity of a subset of spots using the planning objectives of the intended plan. As no energy-adaptation is performed and no new spots are added, the spot placement of the prior plan is very important. A study investigating which prior plan to use would therefore be beneficial for this technique. In a similar adaptation approach, one can also change the spot positions using the deformable image registration [30]. In this case important parameter settings will be more linked to the used deformable image registration approach. To our knowledge, no other techniques currently use an iterative addition of new spots in online plan adaptation. However, as it has been shown that using a resampling technique can lead to higher quality plans more quickly compared to using a regular grid [10], current methods using only new spots may still benefit from resampling.

It should be noted that while this study has addressed many variations on the default parameter settings, not all possibilities have been investigated. An example of this is in the generation of the prior plans. While we have varied the amount of robustness included in the prior plans, all prior plans were made using our in-house TPS. The used optimization features a spot-reduction loop to shorten delivery times [10]. On average the default prior plans included around 1700 pencil-beams (Table 4). It is possible that less spot-reduction in generating

the prior plans could reduce the need for new spots in the adaptation. While this could allow a smaller sample size in full plan adaptation, the impact on the overall calculation time would be small as the total number of spots in the adaptation would remain similar.

Conclusion

For the dose restoration approach, we recommend maintaining the default setting. This is based on the observation that changing the margins and robustness settings in the prior plans can yield insufficient target coverage or increased OAR doses. Full plan adaptation appeared to be more robust against variations of the parameter settings, except for the choice of optimization method. To minimize the impact of the choice optimization method, a more advanced tuning of the RPM-parameters is recommended. Apart from this we recommend using the default settings.

Declaration of Competing Interest

Varian Medical Systems has in part financed this research. Erasmus MC Cancer Institute also has research collaborations with Elekta AB, Stockholm, Sweden and Accuray Inc., Sunnyvale, USA.

Acknowledgments

The CT-data with contours were collected at Haukeland University Hospital, Bergen, Norway and were provided to us by responsible oncologist Svein Inge Helle and physicist Liv Bolstad Hysing. This study was financially supported by ZonMw, the Netherlands Organization for Health Research and Development, grant number 104003012 and by Varian Medical Systems.

References

- [1] Thörnqvist S, Muren LP, Bentzen L, Hysing LB, Høyer M, Grau C, et al. Degradation of target coverage due to inter-fraction motion during intensity-modulated proton therapy of prostate and elective targets. *Acta Oncol (Madr)* 2013;52(3):521–7. <https://doi.org/10.3109/0284186X.2012.752860>.
- [2] Lomax AJ. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: The potential effects of calculational uncertainties. *Phys Med Biol* 2008;53(4):1027–42. <https://doi.org/10.1088/0031-9155/53/4/014>.
- [3] 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(4):1043–56. <https://doi.org/10.1088/0031-9155/53/4/015>.
- [4] Nguyen TTC, Nguyen BT, Mai NV. Robustness evaluation of Intensity Modulated Proton Therapy plans using Dose Volume Population Histogram. *Phys Medica* 2019;65:219–26. <https://doi.org/10.1016/j.ejmp.2019.09.070>.
- [5] Jagt T, Breedveld S, van de Water S, Heijmen B, Hoogeman M. Near real-time automated dose restoration in IMPT to compensate for daily tissue density variations in prostate cancer. *Phys Med Biol* 2017;62(11):4254–72. <https://doi.org/10.1088/1361-6560/aa5c12>.
- [6] Jagt T, Breedveld S, van Haveren R, Heijmen B, Hoogeman M. An automated planning strategy for near real-time adaptive proton therapy in prostate cancer. *Phys Med Biol* 2018;63(13):135017. <https://doi.org/10.1088/1361-6560/aacaa7>.
- [7] Jagt TZ, Breedveld S, van Haveren R, Heijmen BJM, Hoogeman MS. Online-adaptive versus robust IMPT for prostate cancer: How much can we gain?

- Radiother Oncol 2020;151:228–33. <https://doi.org/10.1016/j.radonc.2020.07.054>.
- [8] 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(22):R419–72. <https://doi.org/10.1088/0031-9155/59/22/R419>.
- [9] Breedveld S, Storchi PRM, Voet PWJ, Heijmen BJM. ICycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. *Med Phys* 2012;39(2):951–63. <https://doi.org/10.1118/1.3676689>.
- [10] van de Water S, Kraan AC, Breedveld S, Schillemans W, Teguh DN, Kooy HM, et al. Improved efficiency of multi-criteria IMPT treatment planning using iterative resampling of randomly placed pencil beams. *Phys Med Biol* 2013;58(19):6969–83. <https://doi.org/10.1088/0031-9155/58/19/6969>.
- [11] Kooy HM, Clasié BM, Lu H-M, Madden TM, Bentefour H, Depauw N, et al. A Case Study in Proton Pencil-Beam Scanning Delivery. *Int J Radiat Oncol Biol Phys* 2010;76(2):624–30. <https://doi.org/10.1016/j.ijrobp.2009.06.065>.
- [12] Voet PWJ, Dirx MLP, Breedveld S, Franssen D, Levendag PC, Heijmen BJM. Toward fully automated multicriterial plan generation: A prospective clinical study. *Int J Radiat Oncol Biol Phys* 2013;85(3):866–72. <https://doi.org/10.1016/j.ijrobp.2012.04.015>.
- [13] van de Water S, Kooy HM, Heijmen BJM, Hoogeman MS. Shortening delivery times of intensity modulated proton therapy by reducing proton energy layers during treatment plan optimization. *Int J Radiat Oncol Biol Phys* 2015;92(2):460–8. <https://doi.org/10.1016/j.ijrobp.2015.01.031>.
- [14] Breedveld S, Storchi PRM, Heijmen BJM. The equivalence of multi-criteria methods for radiotherapy plan optimization. *Phys Med Biol* 2009;54(23):7199–209. <https://doi.org/10.1088/0031-9155/54/23/011>.
- [15] Voglis C, Lagaris IE. Boxcqp : an Algorithm for Bound Constrained Convex Quadratic Problems. *Computing* 2004;8–10.
- [16] Breedveld S, Storchi PRM, Keijzer M, Heijmen BJM. Fast, multiple optimizations of quadratic dose objective functions in IMRT. *Phys Med Biol* 2006;51(14):3569–79. <https://doi.org/10.1088/0031-9155/51/14/019>.
- [17] Ogryczak W. Preemptive Reference Point Method. *Multicriteria Anal* 1997:156–67. https://doi.org/10.1007/978-3-642-60667-0_16.
- [18] Ogryczak W, Kozłowski B. Reference point method with importance weighted ordered partial achievements. *Top* 2011;19(2):380–401. <https://doi.org/10.1007/s11750-009-0121-4>.
- [19] van Haveren R, Breedveld S, Keijzer M, Voet P, Heijmen B, Ogryczak W. Lexicographic extension of the reference point method applied in radiation therapy treatment planning. *Eur J Oper Res* 2017;263(1):247–57. <https://doi.org/10.1016/j.ejor.2017.04.062>.
- [20] van Haveren R, Ogryczak W, Verduijn GM, Keijzer M, Heijmen BJM, Breedveld S. Fast and fuzzy multi-objective radiotherapy treatment plan generation for head and neck cancer patients with the lexicographic reference point method (LRPM). *Phys Med Biol* 2017;62(11):4318–32. <https://doi.org/10.1088/1361-6560/62/11/4318>.
- [21] van Haveren R, Heijmen BJM, Breedveld S. Automatically configuring the reference point method for automated multi-objective treatment planning. *Phys Med Biol* 2019;64:aaf9fe.;64(3):035002. <https://doi.org/10.1088/1361-6560/aaf9fe>.
- [22] Fredriksson A, Forsgren A, Hårdemark B. Minimax optimization for handling range and setup uncertainties in proton therapy. *Med Phys* 2011;38(3):1672–84. <https://doi.org/10.1118/1.3556559>.
- [23] Chen W, Unkelbach J, Trofimov A, Madden T, Kooy H, Bortfeld T, et al. Including robustness in multi-criteria optimization for intensity-modulated proton therapy. *Phys Med Biol* 2012;57(3):591–608. <https://doi.org/10.1088/0031-9155/57/3/591>.
- [24] Fredriksson A, Bokrantz R. A critical evaluation of worst case optimization methods for robust intensity-modulated proton therapy planning. *Med Phys* 2014;41(8Part1):081701. <https://doi.org/10.1118/1.4883837>.
- [25] Bernatowicz K, Geets X, Barragan A, Janssens G, Souris K, Sterpin E. Feasibility of online IMPT adaptation using fast, automatic and robust dose restoration. *Phys Med Biol* 2018;63(8):085018. <https://doi.org/10.1088/1361-6560/aaba8c>.
- [26] 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>.
- [27] Matter M, Nenoff L, Meier G, Weber DC, Lomax AJ, Albertini F. Intensity modulated proton therapy plan generation in under ten seconds. *Acta Oncol (Madr)* 2019;58(10):1435–9. <https://doi.org/10.1080/0284186X.2019.1630753>.
- [28] Lalonde A, Bobić M, Winey B, Verburg J, Sharp GC, Paganetti H. Anatomic changes in head and neck intensity-modulated proton therapy: Comparison between robust optimization and online adaptation. *Radiother Oncol* 2021;159:39–47. <https://doi.org/10.1016/j.radonc.2021.03.008>.
- [29] Bobić M, Lalonde A, Sharp GC, Grassberger C, Verburg JM, Winey BA, et al. Comparison of weekly and daily online adaptation for head and neck intensity-modulated proton therapy. *Phys Med Biol* 2021;66(5):055023. <https://doi.org/10.1088/1361-6560/abe050>.
- [30] Botas P, Kim J, Winey B, Paganetti H. Online adaption approaches for intensity modulated proton therapy for head and neck patients based on cone beam CTs and Monte Carlo simulations. *Phys Med Biol* 2019;64:aaf30b.;64(1):015004. <https://doi.org/10.1088/1361-6560/aaf30b>.