A fast and robust constraint-based online re-optimization approach for automated online adaptive intensity modulated proton therapy in head and neck cancer

Oud, Michelle; Breedveld, Sebastiaan; Rojo-Santiago, Jesús; Giżyńska, Marta Krystyna ; Kroesen, Michiel; Habraken, S.J.M.; Perko, Z.; Heijmen, Ben; Hoogeman, M.S.

DOI
10.1088/1361-6560/ad2a98

Publication date
2024

Document Version
Final published version

Published in
Physics in Medicine and Biology

PAPER • OPEN ACCESS

A fast and robust constraint-based online re-optimization approach for automated online adaptive intensity modulated proton therapy in head and neck cancer

IOP Publishing Phys. Med. Biol. 69 (2024) 075007 https://doi.org/10.1088/1361-6560/ad2a98

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

OPEN ACCESS

RECEIVED
21 September 2023

REVISED
6 February 2024

ACCEPTED FOR PUBLICATION
19 February 2024

PUBLISHED
14 March 2024

PAPER

A fast and robust constraint-based online re-optimization approach for automated online adaptive intensity modulated proton therapy in head and neck cancer

Michelle Oud[1,2], **Sebastiaan Breedveld**[1], **Jesús Rojo-Santiago**[1,2], **Marta Krystyna Giżyńska**[2], **Michiel Kroesen**[1,3], **Steven Habraken**[1,2], **Zoltán Perkó**[4], **Ben Heijmen**[1] and **Mischa Hoogeman**[1,2]

[1] Erasmus MC Cancer Institute, University Medical Center Rotterdam, Department of Radiotherapy, Rotterdam, The Netherlands
[2] HollandPTC, Department of Medical Physics & Informatics, Delft, The Netherlands
[3] HollandPTC, Department of Radiation Oncology, Delft, The Netherlands
[4] Delft University of Technology, Faculty of Applied Sciences, Department of Radiation Science and Technology, The Netherlands

E-mail: m.oud@erasmusmc.nl

Keywords: intensity modulated proton therapy (IMPT), daily online adaptive radiotherapy using dose restoration, head-and-neck cancer, inter-fraction anatomy variation, automated treatment planning

Supplementary material for this article is available online

Abstract

Objective. In head-and-neck cancer intensity modulated proton therapy, adaptive radiotherapy is currently restricted to offline re-planning, mitigating the effect of slow changes in patient anatomies. Daily online adaptations can potentially improve dosimetry. Here, a new, fully automated online re-optimization strategy is presented. In a retrospective study, this online re-optimization approach was compared to our trigger-based offline re-planning (offline$_{TB}$ re-planning) schedule, including extensive robustness analyses. *Approach*. The online re-optimization method employs automated multi-criterial re-optimization, using robust optimization with 1 mm setup-robustness settings (in contrast to 3 mm for offline$_{TB}$ re-planning). Hard planning constraints and spot addition are used to enforce adequate target coverage, avoid prohibitively large maximum doses and minimize organ-at-risk doses. For 67 repeat-CTs from 15 patients, fraction doses of the two strategies were compared for the CTVs and organs-at-risk. Per repeat-CT, 10.000 fractions with different setup and range robustness settings were simulated using polynomial chaos expansion for fast and accurate dose calculations. *Main results*. For 14/67 repeat-CTs, offline$_{TB}$ re-planning resulted in $<$50% probability of $D_{98\%} \geqslant 95\%$ of the prescribed dose (D_{pres}) in one or both CTVs, which never happened with online re-optimization. With offline$_{TB}$ re-planning, eight repeat-CTs had zero probability of obtaining $D_{98\%} \geqslant 95\% D_{pres}$ for CTV$_{7000}$, while the minimum probability with online re-optimization was 81%. Risks of xerostomia and dysphagia grade $\geqslant$ II were reduced by 3.5 ± 1.7 and 3.9 ± 2.8 percentage point [mean $\pm$ SD] ($p < 10^{-5}$ for both). In online re-optimization, adjustment of spot configuration followed by spot-intensity re-optimization took 3.4 min on average. *Significance*. The fast online re-optimization strategy always prevented substantial losses of target coverage caused by day-to-day anatomical variations, as opposed to the clinical trigger-based offline re-planning schedule. On top of this, online re-optimization could be performed with smaller setup robustness settings, contributing to improved organs-at-risk sparing.

Introduction

Daily online plan adaptation has the potential to reduce dose degradation caused by inter-fraction anatomy and position variability (Bertholet *et al* 2020, Qiu *et al* 2023, Trnkova *et al* 2023). This is particularly relevant for intensity modulated proton therapy (IMPT), where the delivered dose is substantially more sensitive to these

variations. For head-and-neck (H&N) cancer patients, the large setup robustness settings required to achieve the desired target coverage in the absence of online plan adaptations can result in significantly increased toxicity risks (van de Water *et al* 2016, Oud *et al* 2022). Therefore, there have been continuous efforts to investigate the potential and feasibility of various online adaptation strategies.

Several studies have highlighted the feasibility of fast online adaptation for IMPT. Table 1 provides an overview of proposed online adaptive strategies, and dosimetric evaluations of such strategies, explicitly for H&N IMPT. Approaches can be divided in four categories, depending on the amount of information inherited from the original plan (Paganetti *et al* 2021): (1) online re-planning strategies entail the generation of a full new treatment plan with the original treatment planning pipeline, as introduced by Matter *et al* (2019), Nenoff *et al* (2019). (2) Online re-optimization strategies adapt the original plan to obtain similar or improved plan quality, without aiming at reproducing the original dose distribution (Botas *et al* 2018, Bobić *et al* 2021, 2023, Lalonde *et al* 2021). (3) Online dose restoration strategies adapt the original plan to obtain a similar dose distribution for the daily anatomy (Bernatowicz *et al* 2018, Borderías-Villarroel *et al* 2022, 2023, Miyazaki *et al* 2022). (4) Plan library strategies do not adapt the original plans. Instead, a pre-treatment generated library of plans is employed with daily selection of the library plan that best fits the geometry-of-the-day (Oud *et al* 2022).

Effective and efficient online adaptation requires a fully automated algorithm that guarantees adequate target coverage. However, existing methods can fall short, primarily related to two main factors. First, adapted plans may not have guaranteed robustness against residual errors such as intra-fraction motion, beam alignment to the isocenter of the CT, and range errors. Second, the in-room plan re-optimization workflow may be inefficient, especially in the presence of large anatomical variations. Manual tweaking of objective weights to balance multiple objectives (target coverage, maximum doses to the CTVs, and organ-at-risk (OAR) doses) may be required to achieve adequate target coverage. Constraining the minimum robust target coverage of the CTV during optimization could offer a solution. So far, none of the published online adaptation approaches for H&N cancer have incorporated a combined approach that integrates robust optimization with imposed hard constraints on target coverage.

Furthermore, accurate and systematic evaluation of the dosimetric impact of online adaptation is crucial for the decision-making process regarding its introduction in clinical practice, given the considerable resources associated with online procedures. The potential dosimetric benefit is currently unclear because published assessments do not meet one or both of the following two requirements, (1) no evaluation of robustness of adapted plans against unavoidable residual errors. The omission of such analyzes bears the potential of bias in the conclusions regarding truly delivered doses in CTVs and OARs. Recently, studies employing polynomial chaos expansion (PCE) on the planning-CT have shown that large numbers of dose distributions under the influence of potential residual errors can be generated rapidly and used for statistically accurate plan robustness analysis (Perkó *et al* 2016, Rojo-Santiago *et al* 2021, 2023). This has not yet been employed to evaluate plan adaptation strategies, (2) no comparison with current state-of-the-art clinical treatment planning strategy, such as robust optimization with trigger-based offline adaptive re-planning. Bobić *et al* (2023) compared their online re-optimization strategy to their clinical offline adaptation strategy. However, robustness evaluation was not performed and they exclusively included patients that needed an offline adaptation, not providing a representative sample of the patient population. In Oud *et al* (2022), our plan library strategy was compared in compliance with the two requirements. However, online dose restoration, re-optimization, and re-planning strategies can potentially further improve dosimetry.

In this study, a fully automated online re-optimization strategy is proposed that guarantees CTV coverage by using hard planning constraints and by employing mini-max robust optimization (Fredriksson *et al* 2011). Setup robustness settings of 1 mm and automated multi-criterial optimization are used to maximally reduce OAR doses. Spots are added to the original spot distribution in poorly covered areas of the CTV, to ensure a good spot distribution while maintaining the original spot configuration as much as possible. Our novel online re-optimization strategy was validated for H&N cancer by dosimetric comparisons to our current clinical treatment strategy, which entails trigger-based offline robust re-planning. Dosimetric comparisons between the novel online re-optimization strategy and trigger-based offline re-planning included extensive robustness analyzes on repeat-CTs using PCE evaluations (Perkó *et al* 2016, Rojo-Santiago *et al* 2021).

Methods and materials

Patient data

In this retrospective study, CT-scans of fifteen primary H&N cancer patients treated with IMPT at Holland Proton Therapy Center in 2019 and 2020 were included. The following inclusion criteria had to be met: (1) availability of three or more repeat-CTs in treatment position, acquired during the fractionated treatment to verify the need for offline re-planning due to anatomical changes; and (2) Robust CTV coverage constraints

Table 1. Overview of publications on overview of proposed online adaptive strategies, and dosimetric evaluations of such strategies in head-and-neck IMPT. PCE = Polynomial chaos expansion (see text). RS = Robustness setting (setup robustness/range robustness). * Only patients with offline plan adaptations, ** only used when target coverage constraints were not met on repeat-CTs.

	Matter *et al* (2019)	Nenoff *et al* (2019)	Botas *et al* (2018)	Bobić *et al* (2021)	Lalonde *et al* (2021)	Bob *et al* (2023)	This study	Bernatowicz *et al* (2018)	Borderías-Villarroel *et al* (2022)	Miyazaki *et al* (2022)	Borderias Villarroel *et al* (2023)	Oud *et al* (2022)
Online adaptation approach	Re-planning	Re-planning	Re-optimization	Re-optimization	Re-optimization	Re-optimization	Re-optimization	- Dose restoration - Re-optimization	Dose restoration	Dose restoration	Combined approach: dose restoration and re-optimization	Plan library
PTV margins or RS for adaptation	PTV	- PTV 5 mm - PTV 1 mm + range-specific distal margin of 3%	None	None	None	PTV 1 mm	Robust 1 mm/3%	PTV 3–4 mm	Robust 3 mm/3%	Robust 3 mm/3.5% (inherited nominal dose on planning-CT)	Robust 4 mm/3%	Robust 1–5 mm/3%
Spot adjustment strategy	New spot configuration	New spot configuration	Position, energy	None	None	None	Energy, spot addition	Energy	Energy	New spot configuration	New spot configuration	n.a.
Contours used for online adaptation	Manual	Manual	Deformably propagated	Deformably propagated	Deformably propagated	Deformably propagated	Manual	Rigidly propagated	Rigidly propagated	Rigidly propagated and manual	Rigidly propagated and deformably propagated	Manual (CTV only)
Number of H&N patients	1	5	10	10	10	8*	15	2	10	2	10	15
Number of repeat-images/patient	1	1	5–7	31–35	31–35	26–35	3–6	1	4–6	1	35**	3–6
Type of repeat-images	n.a.	Simulated CTs	Scatter-corrected CBCT	Scatter-corrected CBCT	Scatter-corrected CBCT	Scatter- corrected CBCT	CT	CT	CT	CT	Corrected-CBCT	CT
Evaluation: planning strategy that online adaptation was compared to (adaptive strategy, RS or PTV)	Full planning, PTV	- Non adaptive, 5 mm PTV - Non adaptive 1 mm PTV + range-specific distal margin of 3%	Non adaptive, no margin or robustness	- Non adaptive, no margin or robustness - Weekly online re-optimization, no margin or robustness	- Non adaptive, robust 3 mm/0% - Non adaptive, anatomical robust 3 mm/0%	Offline adaptive, 4 mm PTV	Trigger-based offline adaptive, robust 3 mm/3%	Non adaptive, 3–4 mm PTV	- Non adaptive, robust 3 mm/3% - Full re-planning, robust 3 mm/3%	- Non adaptive, robust 3 mm/3.5% -Full re-planning, robust 3 mm/3.5%	Non adaptive, robust 4 mm/3%	- Non adaptive, robust 1–5 mm/3% - Trigger-based offline adaptive, robust 3 mm/3%

Table 1. (Continued.)

	Matter *et al* (2019)	Nenoff *et al* (2019)	Botas *et al* (2018)	Bobić *et al* (2021)	Lalonde *et al* (2021)	Bob *et al* (2023)	This study	Bernatowicz *et al* (2018)	Borderías-Villarroel *et al* (2022)	Miyazaki *et al* (2022)	Borderias Villarroel *et al* (2023)	Oud *et al* (2022)
Evaluation: contouring method on repeat-CTs	Manual	Manual	Deformably propagated	Deformably propagated	Deformably propagated	Deformably propagated	Manual	Rigidly propagated	Manual	Manual	Deformably propagated and manual	Manual
Evaluation: robustness analysis	PTV	Simulated treatments	Not performed (CTV)	Not performed (CTV)	Not performed (CTV)	Not performed (CTV)	PCE	Not performed (CTV)	1 mm/3% (worst-case), per-fraction	3 mm/3.5% (DVH-bandwidth)	Not performed (CTV)	Simulated treatments
Evaluation: per-fraction or accumulated dose	n.a.	Accumulated	Per-fraction and accumulated	Per-fraction and accumulated	Per-fraction and accumulated	Per-fraction and accumulated	Per-fraction	Per-fraction	Per-fraction	Per-fraction	Per-fraction and accumulated	Accumulated
Optimization times (excl. dose computation)	2.3 s	—	61.7 s average	720 s median (Paganetti *et al* 2021)	Approximately 360 s	—	206 s average	[20–80] s	672 s median	372 s average	[780–1020] s	—

could be met on the planning-CT without exceeding constraints on serial OARs due to proximity of the CTV. Prescribed doses (D_{pres}) were 70 GyRBE to the high-dose CTV, including the GTV and positive lymph nodes (CTV_{7000}) and 54.25 GyRBE to the elective areas (CTV_{5425}) in 35 fractions. A constant relative biological effectiveness (RBE) of 1.1 was used in planning. For each patient, 3–6 repeat-CTs were available. The acquisition of repeat-CTs was part of the standard protocol, and the frequency per patient was mostly based on the availability of personnel and CT scanner.

CTV contours were propagated from the planning-CT to the corresponding repeat-CTs. The CTV_{7000} and the part of the CTV_{5425} that was within a 5 mm margin to the CTV_{7000} were rigidly propagated from the planning-CT to the repeat-CTs and were manually adjusted if contours were outside the external patient contour or inside bone or if large discrepancies occurred (e.g. the hyoid bone moved resulting in a different position of the GTV). The remainder of the CTV_{5425} was deformably propagated to the repeat-CTs and manually adjusted in case of noticeable mismatches with the repeat-CTs. The contours were checked for consistency by expert clinicians. Contours of the OARs were deformably propagated to the repeat-CTs, and manually adjusted.

Online re-optimization

The novel online re-optimization strategy consisted of daily adaptation of the initial treatment plan using contoured repeat-CTs. The initial treatment plan was obtained through full multi-criterial optimization (see below). The re-optimization algorithm was embedded in our in-house system for automated treatment planning (Erasmus-iCycle), and used a weighted-sum cost-function and hard constraints for plan generation. Weights in the applied weighted-sum cost-function were Lagrange parameters obtained from the initial treatment plan (Breedveld *et al* 2009). Such weight extractions could be done automatically, prior to the first treatment fraction. Online re-optimization for repeat-CTs then consisted of five steps: (1) restoration of spot Bragg peak positions to their intended positions by adjusting their energies to account for changes in water-equivalent-path-lengths (WEPL) (Jagt *et al* 2017). Restored energies were interpolated to the same energy grid as used for initial treatment planning, (2) the online re-optimization strategy employs a novel method to improve the established original spot distribution, in which new spots were added specifically to target areas that were not covered by spots after step 1 due to anatomical changes. Hereto, a dense candidate spot distribution with spots originating from all beam directions was first placed over the CTV + 5 mm expansion, followed by an iterative selection of 2000 of these spots with a Bragg peak location at 8 mm or more from restored original spots, (3) computation of dose-deposition matrices for the repeat-CT anatomy for all spots (restored original and added), (4) constrained robust optimization of the intensities of all spots using the weighted-sum cost-function. The same robustness settings as the initial plan were used (1 mm/3% setup/range robustness settings in 29 scenarios, see below). The constraints ensured appropriate coverage of the two CTVs, (5) spots that had an intensity below the minimum required monitor units were removed to ensure deliverability of the treatment plan.

Trigger-based offline re-planning schedule

Our current clinical adaptation strategy consists of trigger-based offline re-planning ($offline_{TB}$ re-planning). For the simulation of this strategy, treatment plans were generated with full multi-criterial optimization (see below), and the clinical offline re-planning schedule was followed. In our clinical workflow, triggering offline adaptations is guided by dose assessments on repeat-CTs and the evaluation of sequential daily CBCTs. In this dataset, nine plan offline adaptations were performed for seven patients in total. Four of these plan adaptations were performed on the last repeat-CT. In the simulations, offline adapted plans were used from the next repeat-CT onwards. Adaptations based on the last repeat-CT were therefore not taken into account in the evaluations. This resulted in a total of 5 plan adaptations for 5 patients that were taken into account. In supplementary data A, a schematic representation of the treatment planning schedule in the online re-optimization and $offline_{TB}$ re-planning strategies is shown.

Full multi-criterial optimization of treatment plans

For the online re-optimization strategy, initial treatment plans for the planning-CTs were generated with full multi-objective optimization. For the $offline_{TB}$ re-planning strategy, all treatment plans on the planning-CT and the 5 plan adaptations were generated with this approach. Wish-list driven fully-automated software was used for generation of the treatment plans: Erasmus-iCycle (Breedveld *et al* 2012, van de Water *et al* 2013). To simultaneously also ensure plan robustness, these optimizations were mini-max scenario-based (Fredriksson *et al* 2011, Liu *et al* 2013). Smaller setup robustness settings were applied for treatment plans used in the online re-optimization strategy. A resampling approach was used for spot selection (van de Water *et al* 2013), where candidate spots are iteratively selected from a dense grid and added to the optimization problem. After each

iteration of optimization, non-contributing spots are removed from the optimization problem. Details on the applied wish-list, sequential minimization of prioritized objectives subject to hard constraints, and treatment-site specific configuration of the employed treatment planning system can be found in Breedveld *et al* (2012), van de Water *et al* (2013), Heijmen *et al* (2018). Details on the specific configuration of automated plan generation in this study are described in the supplementary data of Oud *et al* (2022).

During optimization, the nominal scenario and 28 uncertainty scenarios were used to account for variations in patient setup and uncertainties in proton ranges (Korevaar *et al* 2019). For offline$_{TB}$ re-planning, 3 mm/3% robustness settings (setup robustness/range robustness) were used, while 1 mm/3% was used for online re-optimization. The selected setup robustness setting of 3 mm for offline$_{TB}$ re-planning is the same as currently used in our clinic. The 1 mm setup robustness setting for online re-optimization was based on literature (Nenoff *et al* 2021, Bobić *et al* 2023). Both for CTV$_{7000}$ and for CTV$_{5425}$, generated treatment plans had to meet a coverage constraint: $V_{95\%} > 98\%$ in the voxelwise minimum dose distribution (Korevaar *et al* 2019). Note that voxelwise dose distributions were only used in the planning phase, not for plan evaluations and comparisons (below).

Evaluation and comparison of online re-optimization and offline$_{TB}$ re-planning

Dosimetric evaluations and comparisons of the two investigated adaptive strategies were performed for the available repeat-CTs. The impact of inter-fraction anatomical changes was incorporated through repeat-CTs. While adaptive approaches can mitigate anatomical changes, they cannot compensate for residual errors: errors in matching the gantry to the isocenter of the CT, uncertainties in couch positioning, registration errors with the MR that was used for target delineation, registration errors with the CT, intra-fraction motion and proton range uncertainties (supplementary data B). On the other hand, dosimetric variations caused by these uncertainties will in reality occur and need to be accounted for in evaluations and comparisons of optimized/predicted doses.

PCE (Perkó *et al* 2016, Rojo-Santiago *et al* 2021) was used for extensive evaluation of robustness of generated dose distributions against residual uncertainties. The rationale to use PCE comes from its ability to accurately approximate 3D dose distributions for all (i.e. thousands of) uncertainty scenarios of a treatment plan in a matter of seconds, allowing statistical robustness evaluation instead of using the common nominal or 29 scenarios. Instead of executing forward dose computations to obtain the 3D dose distribution for all the scenarios, PCE constructs and employs a computational model to predict dose distributions. This model is a multi-dimensional polynomial function of the stochastic input variables (geometrical errors and proton range-errors in this study). Expansion coefficients of the function are obtained by linear regression, see Perkó *et al* (2016), Rojo-Santiago *et al* (2021) for details. Our PCE approach is implemented in Matlab (version 2021b) and was previously validated for the employed ASTROID dose engine (Kooy *et al* 2010, Perkó *et al* 2016, van der Voort *et al* 2016). In this study, the expansion coefficients were obtained based on the input of 208 dose distributions computed in fixed uncertainty scenarios. Once these computations are completed, computing the expansion coefficients took around 2 s, and the generation of 10.000 scenario dose distributions for a structure took around 2 s. Time required for assessment of the DVHs highly depends on the employed settings and organ size. In our case, DVH computation for the two CTVs took around 30 min for an average patient (analysis performed on an Intel® Xeon® Gold 6248). PCE models were constructed for each repeat-CT, for both online re-optimization and offline$_{TB}$ re-planning strategy.

Both for online re-optimization and offline$_{TB}$ re-planning, PCE was used to calculate for each repeat-CT 10.000 dose distributions, each for a randomly selected total residual setup error for each of the three principal directions, and a randomly selected range error for offline$_{TB}$ re-planning and online re-optimization. These errors were sampled from Gaussian distributions describing total residual geometric uncertainties. The standard deviations (SD) of these Gaussian distributions were derived from quality assurance (QA) and treatment data at Holland Proton Therapy Center by quadratically adding SDs of the various residual errors involved. This resulted in total SDs of 1.18, 1.16, and 1.22 mm for the setup errors in lateral, longitudinal, and vertical directions, respectively. A description of the residual errors and the employed SD can be found in supplementary data B. The Gaussian distribution of range errors was assumed to have a SD of 1.5% in correspondence to Taasti *et al* (2018).

For each repeat-CT, the obtained dosimetric values were multiplied by 35 to arrive at values for full treatments to improve interpretability. The probabilities for adequate CTV coverage and for exceeding preferred maximum doses in a fraction were established with the 10.000 PCE dose distributions. The $D_{98\%}$ was evaluated as it is an ICRU-recommended metric that is numerically more robust than point minimum doses. Reported OAR mean doses were obtained by averaging achieved mean doses in the 10.000 PCE dose distributions. For each repeat-CT, NTCPs were calculated using these OAR doses. The risks of xerostomia and dysphagia grade $\geqslant$ II and $\geqslant$ III were evaluated. We used NTCP models described in the Dutch National Indication protocol

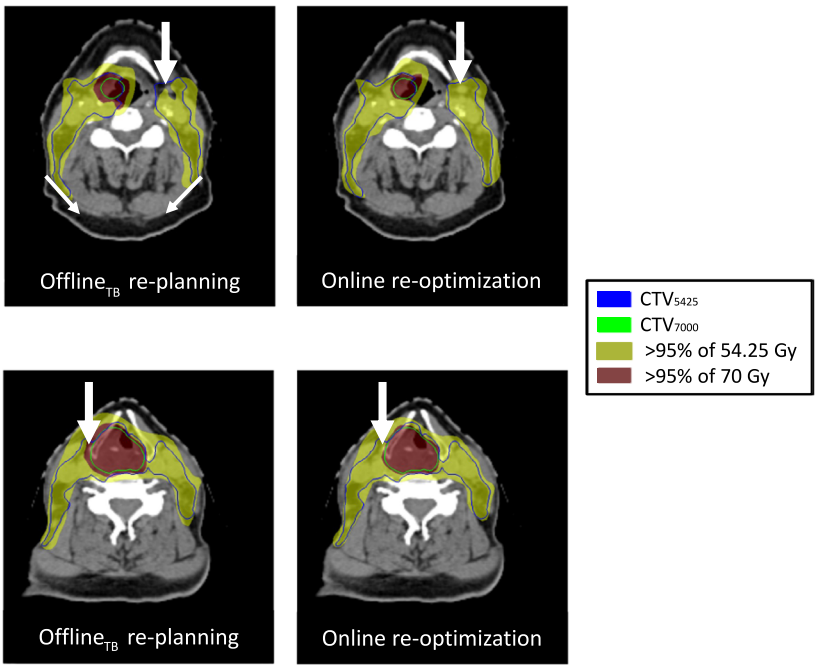

Figure 1. Slices of an example repeat-CT (#38 in figure 3), showing the dose distribution corresponding to the scenario having the 90%-worst-case $D_{98\%}$ among the 10 000 simulated fractions for the CTV_{5425} in the trigger-based offline re-planning ($offline_{TB}$ re-planning) strategy. The arrow in the upper left panel points at an underdose in CTV_{5425} with the $offline_{TB}$ re-planning strategy, while proper dose is delivered with online re-optimization (upper right). Arrows in the lower panels point at too wide high dose delivery around CTV_{7000} with $offline_{TB}$ re-planning, which does not happen with online re-optimization.

(National Association for Radiotherapy in the Netherlands 2019), which were constructed using the data of 750 patients using multivariable regression analysis and were validated on an independent dataset.

Online re-optimization times were evaluated on an Intel® Xeon® Gold 6248, ignoring dose computation and contouring times.

Statistical significance of differences between online re-optimization and $offline_{TB}$ re-planning was tested using paired Wilcoxon signed-rank tests for paired data, and Wilcoxon rank sum tests for unpaired data ($\alpha < 0.05$).

Results

CTV coverage for example repeat-CT

Figures 1 and 2 illustrate results for an example repeat-CT. Figure 1 presents the 90%-worst-case obtained with online re-optimization for CTV_{5425}. This particular repeat-CT was selected because it shows the efficacy of online re-optimization in areas of large anatomical variations (top panels): underdosage in the CTV_{5425} with $offline_{TB}$ re-planning, while coverage was maintained with online re-optimization. The example also shows the ability of online re-optimization to maintain good conformity, which is reduced with $offline_{TB}$ re-planning (bottom panels).

Figures 2(a) and (b) show the $D_{98\%}$ histogram and the corresponding cumulative $D_{98\%}$ histogram for the example repeat-CT in figure 1, derived from the 10.000 fractions obtained with the PCE simulations. With $offline_{TB}$ re-planning, the probabilities of obtaining at least 95% D_{pres} in CTV_{5425} and CTV_{7000} were 50% and 98% (points A1 and A2 in figure 2(b)) while for online re-optimization they were 88% and 83% (points A3, A4 in figure 2(b)).

CTV coverage in population of repeat-CTs

Figure 3 shows for each repeat-CT for the two adaptive strategies separately, the percentage of the 10.000 PCE dose distributions with a $D_{98\%}$ of at least 95%D_{pres}, both for CTV_{5425} and CTV_{7000}. While for the majority of repeat-CTs the chances of reaching $D_{98\%} \geqslant 95\%D_{pres}$ were higher with $offline_{TB}$ re-planning, for a large minority of repeat-CTs, the probability for adequate coverage was zero or close to zero with $offline_{TB}$ re-planning because of large changes in patient geometry, while online re-optimization was able to guarantee high coverage probabilities. For 14/67 repeat-CTs, $offline_{TB}$ re-planning resulted in <50% probability of $D_{98\%} \geqslant 95\%D_{pres}$ in one or both CTVs, which never happened with online re-optimization. With $offline_{TB}$ re-planning,

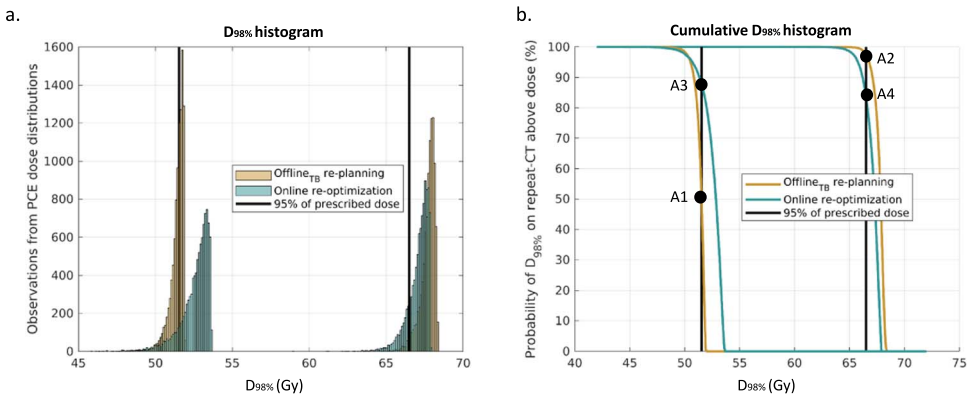

Figure 2. For the example repeat-CT in figure 1, histograms showing distributions of $D_{98\%}$ values in the 10.000 polynomial chaos expansion (PCE) dose distributions (a) and the corresponding $D_{98\%}$-histogram. (b) Points A1–4 are discussed in the text.

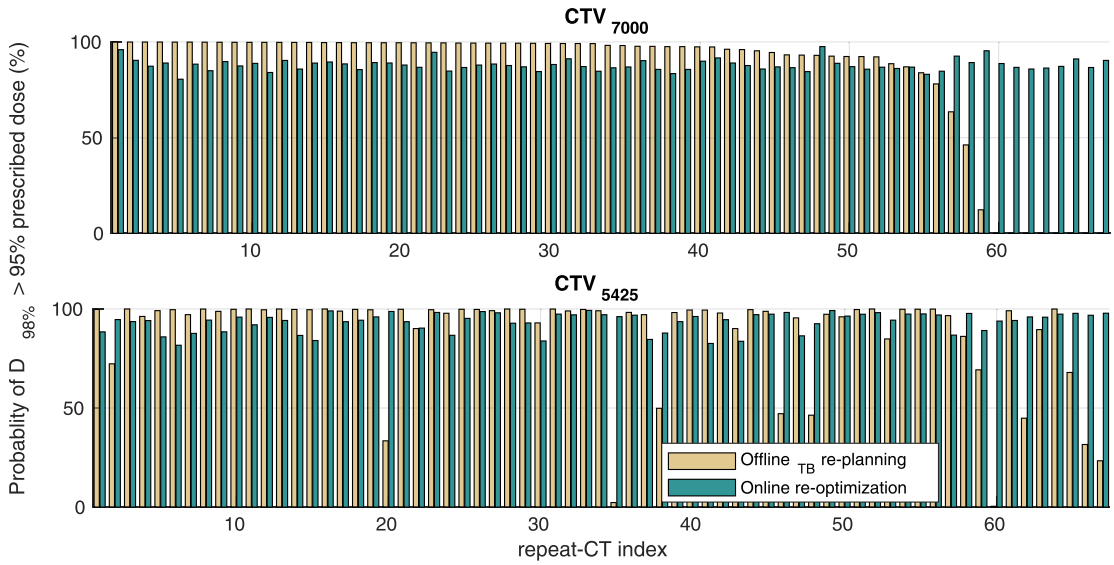

Figure 3. For each of the 67 repeat-CT, the percentage of the 10.000 PCE dose distributions with $D_{98\%} > 95\%$ of the prescribed dose for the two investigated adaptive strategies, for CTV_{7000} (top) and CTV_{5425} (bottom). Repeat-CTs were sorted based on CTV_{7000} $D_{98\%}$ in the trigger-based offline adaptive ($offline_{TB}$ re-planning) strategy.

eight repeat-CTs had zero coverage probability for CTV_{7000}, while the minimum repeat-CT coverage probability with online re-optimization was 81%. For online re-optimization, the mean percentages with [10th, 90th percentile] were 93.6% [84.7%, 91.1%] for CTV_{5425} and 87.9% [86.0%, 98.2%] for CTV_{7000}. For $offline_{TB}$ re-planning, this was 87.7% [46.6%, 100%] and 83.0% [0, 99.9%], for CTV_{5425} and CTV_{7000}. Comparing the 10th and 90th percentiles of the distributions for $offline_{TB}$ re-planning and online re-optimization, target coverage is more consistent across the repeat-CTs for online re-optimization. With online re-optimization, probabilities for adequate target coverage in repeat-CTs were highly similar to those in the corresponding planning-CT plans (supplementary data C), meaning that the intended target coverage from the planning-CTs was maintained for repeat-CTs. This was certainly not achieved with $offline_{TB}$ re-planning, as in a large minority the probability was extremely low.

Figure 4(a) shows cumulative $D_{98\%}$ histograms for CTV_{5425} and CTV_{7000} for the 95th, 90th, 85th, and 5th percentiles of the population of 67 repeat-CTs. For both CTVs, comparison of the 95th percentile curves points to enhanced rates of adequate coverage with online re-optimization compared to $offline_{TB}$ re-planning. Comparisons of the 5th percentile curves' intersections with 90% probability of $D_{98\%}$ above dose (B1 with B3 and B2 with B4) in figure 4(a) further point at clear coverage advantages for online re-optimization, as 95% of the repeat-CTs had $\geqslant$90% chance of receiving a $D_{98\%}$ of $\geqslant$49.4 Gy in the CTV_{5425} and $\geqslant$64.0 Gy in the CTV_{7000} for $offline_{TB}$ re-planning, while this was $\geqslant$51.1 Gy and $\geqslant$66.1 Gy for online re-optimization. The 95%D_{pres} intersections with the 5th percentile also highlight the advantage of online re-optimization, showing that for 95% of the patient population online re-optimization ensures a minimum CTV_{5425} coverage probability as high

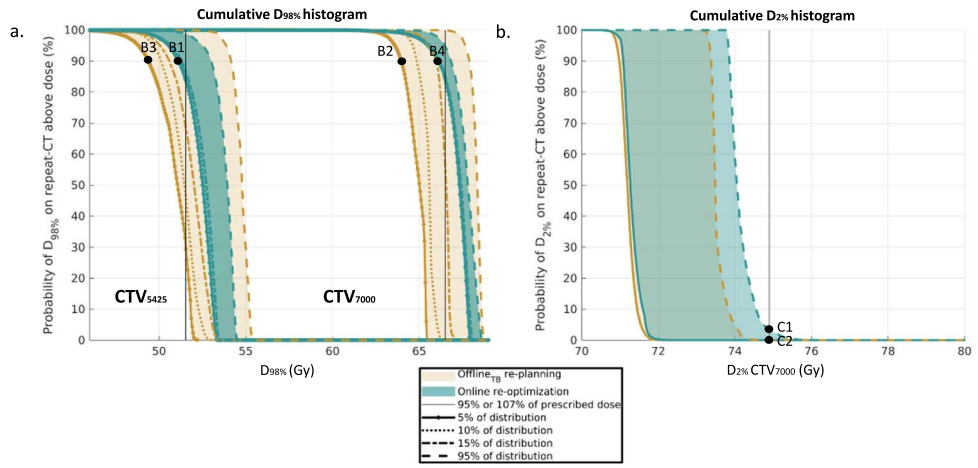

Figure 4. For the population of 67 repeat-CTs, $D_{98\%}$- (a) and $D_{2\%}$- (b) cumulative histograms obtained from the 10.000 PCE dose distributions for each of the repeat-CT scans. For each of the CTVs, the construction of the presented cumulative histograms started by establishing the cumulative histogram for each repeat-CT separately. Next, the 95%, 90%, 85%, and 5% of these 67 cumulative histograms were taken and displayed. Points B1–4, C1–4 and D1–2 are discussed in the text.

as 84%, while this drops to 31% with offline$_{TB}$ re-planning (points C1 and C3), and for CTV$_{7000}$ online re-optimization ensures a minimum coverage probability of 84% versus 0% with offline$_{TB}$ re-planning (points C2 and C4). Furthermore, comparisons of the difference between the 5th percentile curves and the 95th percentile curves between the two adaptive strategies in figure 4(a) show enhanced consistency in target coverage across the patient population with online re-optimization.

Figure 4(b) displays cumulative population $D_{2\%}$ histograms for CTV$_{7000}$ for the two strategies with corresponding 5th and 95th percentile curves. The 95th percentile curves (right dashed curves) show a slight disadvantage for online re-optimization: for $\geqslant$95% of the repeat-CTs the probability that $D_{2\%}$ exceeded 74.9 Gy (107%D_{pres}) was only limited at $\leqslant$3%, while this probability was zero for offline$_{TB}$ re-planning (see points D1 and D2 in figure 4(b)).

OARs

Differences between online re-optimization and offline$_{TB}$ re-planning in OAR mean doses are shown in figure 5(a). Online re-optimization was superior for all dose differences ($p < 10^5$ for all). The highest reduction was observed for the middle constrictor muscle (−5.0 Gy on average, ranging from −23.8 to 3.8 Gy). Figure 5(b) shows to what extent the superiority of online re-optimization in OAR doses impacts NTCPs. The risk of xerostomia and dysphagia grade $\geqslant$ II could be reduced significantly by 3.5 ± 1.7 percentage point and 3.9 ± 2.8 percentage point [mean ± SD] ($p < 10^5$ for both). The risk of xerostomia and dysphagia grade $\geqslant$ III could be reduced by 1.1 ± 0.6 percentage point and 1.0 ± 1.0 percentage point ($p < 10^5$ for both).

Optimization times

In the online re-optimization strategy, restoration of the planned WEPL took 2.5 ± 0.6 s [mean ± SD] per repeat-CT. Spot addition took 14.8 ± 2.8 s. Re-optimization times were 189 ± 31 s.

Discussion

In this study, we have proposed a novel approach for online adaptive dose re-optimization in IMPT and evaluated this for patients with H&N cancer. To enforce adequate target coverage and minimize OAR dose, the online re-optimization method employs automated constraint-based multi-criterial re-optimization, mini-max robust optimization with a 1 mm setup-robustness setting, spot restoration and spot addition. Dose distributions obtained with online re-optimization were benchmarked against our current offline adaptive re-planning protocol. To obtain an accurate and relevant comparison of the two strategies, a representative patient population and realistic robustness settings during optimization were used, and a comprehensive statistical analysis of the robustness against residual errors was performed.

Three important advantages of the online re-optimization strategy compared to offline$_{TB}$ re-planning were identified: first, with the novel online re-optimization strategy, the intended planning-CT target coverage was maintained in the repeat-CTs. Second, online re-optimization resulted in full avoidance of very low probabilities for adequate target coverage, while with offline$_{TB}$ re-planning in 21% of the analyzed repeat-CTs the chance of

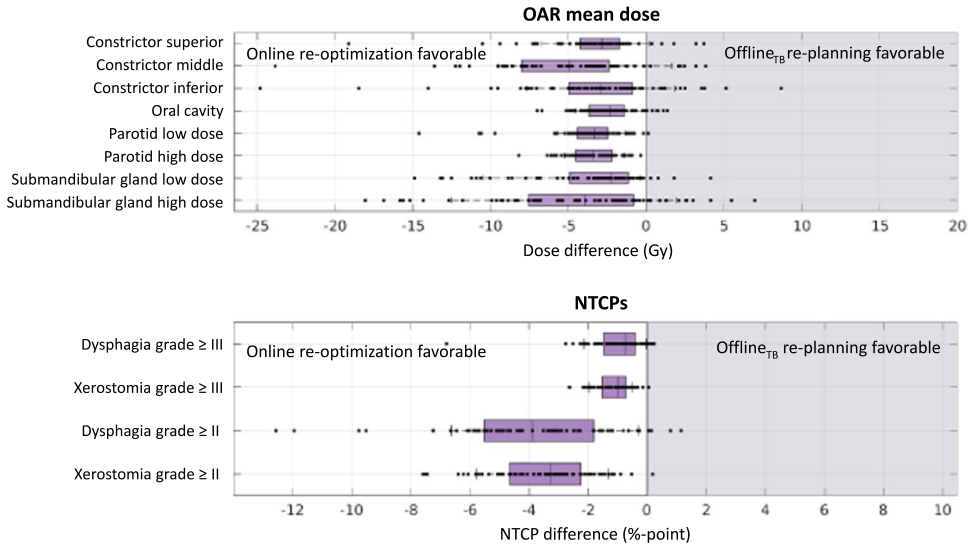

Figure 5. Differences between online re-optimization and trigger-based offline adaptive planning (offline$_{TB}$ re-planning) in (top) OAR mean doses and (bottom) resulting NTCPs. Box plots show distributions for the 67 evaluated repeat-CTs, every dot indicates a repeat-CT. Whiskers extend to the 10th and 90th percentiles of the distributions. All differences in OAR mean doses and in NTCPs were statistically significant with $p < 10^5$.

adequate target coverage was <50% in one or both CTVs. Third, online re-optimization significantly reduced OAR doses, which resulted in reduced NTCPs. On the other hand, two dosimetric advantages of offline$_{TB}$ re-planning were found: first, near-maximum doses in the CTV$_{7000}$ were slightly advantageous with offline$_{TB}$ re-planning. Second, apart from very low coverage probabilities with offline$_{TB}$ re-planning for some repeat-CTs (above), there is also a substantial fraction of repeat-CTs with coverage probabilities approaching 100% (48% of repeat-CTs with 100% CTV$_{7000}$ coverage probability, see figure 3). However, very high coverage probabilities in the latter are a result of the population-based setup robustness settings determined by the near-worst performing patients. The overall increased NTCPs with offline$_{TB}$ re-planning are likely (partly) related to these high coverages. With online re-optimization such higher-than-requested coverage probabilities were avoided: they were more in line with intended coverage probabilities in the corresponding planning-CT plans.

For online re-optimization, obtained probabilities for adequate target coverage ($D_{98\%} > 95\%D_{pres}$) on the repeat-CTs were 87.9% [86.0%, 98.2%] (mean [10th, 90th percentile]) for CTV$_{7000}$ and 93.6% [84.7%,91.1%] for CTV$_{5425}$. Further research, including dose accumulation studies, is needed to establish the optimal choice for the metrics used to assess adequate target coverage, as well as for the optimal probabilities of obtaining adequate target coverage. With the proposed online re-optimization approach, steering the $D_{98\%}$ is possible with the applied setup robustness settings in planning-CT plan generations (1 mm in the current study, as proposed in literature).

Guaranteeing adequate target coverage without prohibitively high maximum doses requires an appropriate spot distribution. An improved spot distribution in the daily situation can be obtained by complete replacement of the spot configuration (Borderias Villarroel *et al* 2023), but generating this distribution can be time-consuming and large changes in spot configurations are undesirable for online QA purposes. Previous studies found that restoration could be performed by keeping the original spot distribution (Bobić *et al* 2021, Lalonde *et al* 2021) or only changing the energies of the spots (Botas *et al* 2018). However, our findings were that spot addition was necessary to achieve dose distributions with acceptable target coverages and acceptable maximum doses in targets and surroundings. This is related to the imposed dosimetric constraints on target coverage in our strategy. Contrary to our findings, imposed constraints on target coverage in Lalonde *et al* (2021) without the use of spot addition resulted in acceptable near-maximum doses in the high-dose CTV. Possibly, this is related to their larger number of spots in original treatment plans, the omission of robust optimization and evaluation, or a different degree of variability in patient anatomies. Although the near-maximum doses with our online re-optimization strategy were still slightly higher compared to offline$_{TB}$ re-planning, it may not be clinically relevant to further improve these.

In our comparisons, the clinical trigger for offline re-planning was based on plans generated with a 5 mm setup robustness setting while we used 3 mm setup robustness settings as this complies with current standard clinical care of H&N proton treatment in the Netherlands. This may have contributed to a relatively lower coverage in the trigger-based offline adaptive approach.

For both strategies, all treatment plans were generated with Erasmus-iCycle, intentionally omitting the use of the clinical treatment plan. The rationale behind this decision was to allow for fully-automated plan generation for both adaptive strategies and to keep optimization differences between the strategies minimal, e.g. using the same dose calculation algorithm, the same implementation of cost functions, etc. Employing treatment plan generation with identical optimization settings (except the setup robustness setting) ensured a systematic and consistent dosimetric evaluation. Erasmus-iCycle generates treatment plans with comparable-to-better quality compared to clinical treatment plans, as shown by Huiskes *et al* (2023).

Our comparison of online re-optimization and offline$_{TB}$ re-planning was based on 3–6 repeat-CTs per patient, while other studies (table 1) performed their analyzes of online re-optimization using corrected CBCTs. A limitation of our study was that evaluating with a limited number of repeat-CTs may have led to the possibility that offline adaptations were initiated earlier in clinical practice, potentially resulting in an overrepresentation of the number of fractions with inadequate target coverage in this study. On the other hand, corrected CBCTs provide daily information on the anatomy. However, HU in corrected CBCTs are less accurate compared to repeat-CTs (Park *et al* 2015, Lalonde *et al* 2020, Thummerer *et al* 2020). Therefore, using corrected CBCT also introduces potential bias because online re-optimization will compensate for inaccuracies in HU in contrast to offline re-planning.

This study only compared single-fraction delivered doses. For interpretability, fraction doses were multiplied by the number of fractions during the complete treatment. Analyzes of accumulated dose would provide insight in the total delivered dose during the treatment. Currently, single-fraction dosimetric constraints in adaptive treatments are not available. Accumulated doses, in combination with comparisons to current clinically applied strategies, could potentially aid in determining whether the obtained coverage with online re-optimization is adequate. Investigation of accumulated dose was hampered in this study by manual contour adjustments that invalidated deformable vector fields obtained from HU matching.

The contours that were used on the repeat-CTs were obtained by automatic propagation of planning-CT contours, followed by manual correction in case of observed inaccuracies. Editing of the automatically propagated contours was performed to enhance the accuracy of the performed dosimetric analyzes for the comparison of online re-optimization with offline$_{TB}$ re-planning. In an online setting, available time for editing may be limited. Further enhancement of image quality and contour propagation algorithms may limit the time needed for editing. Possibly, the dosimetric impact of using non-edited or minimally-edited contours for online adaptive IMPT obtained from propagation or AI-based solutions is small and only CTV contours may require manual adjustments (Guo *et al* 2021, Smolders *et al* 2023).

In this study, the impact of inter and intra-observer variations in contouring on planning and repeat-CTs was ignored during simulations. Inter and intra-observer variations in target volume delineations are one of the largest contributors to variations in dose delivery (Barbara Segedin 2016, Apolle *et al* 2019, van der Veen *et al* 2019, Nash *et al* 2022). Furthermore, contouring on repeat-CTs without MRI also introduces extra uncertainty as the target is not always well visible.

Implementation of this online re-optimization strategy requires not only fast re-optimization, high quality imaging, and appropriate contours, but also fast online QA procedures, which are currently unavailable in clinical settings. However, recent advancements in independent dose computations and log-file-based QA strategies have been made (Li *et al* 2013, Meier *et al* 2015, Meijers *et al* 2020, Burlacu *et al* 2023). Clinical application of online adaptive approaches also requires fast dose computation algorithms. Fortunately, algorithms that provide fast dose computations are becoming available (Pastor-Serrano and Perkó 2022).

Conclusion

In this study, we introduced a fast and fully automated online robust re-optimization strategy for daily adaptation of initial treatment plans to cope with day-to-day anatomical variations for H&N IMPT. Hard constraints and spot addition were used to maintain adequate CTV dose. Based on a comprehensive robustness evaluation, we conclude that substantial loss of target coverage, as observed with our current clinical trigger-based offline adaptive (offline$_{TB}$ re-planning) strategy, could be fully avoided with the online re-optimization strategy, while using small setup robustness settings. This resulted in improved OAR doses and reduced NTCPs compared to offline$_{TB}$ re-planning.

Data availability statement

The data cannot be made publicly available upon publication due to legal restrictions preventing unrestricted public distribution. The data that support the findings of this study are available upon reasonable request from the authors.

Funding statement

This work was partly funded by a research grant of Varian, a Siemens Healthineers Company. The Erasmus MC Cancer Institute also has research collaborations with Elekta AB, Stockholm, Sweden, and Accuray Inc., Sunnyvale, USA.

Ethical statement

The research was conducted in accordance with the principles embodied in the declaration of Helsinki and in accordance with local statutory requirements. The data used in this study originates from the research database of Holland Proton Therapy Center. This database consists of data from all consenting patients treated at HollandPTC. These patients gave written informed consent. The local Institutional Review Board of LUMC waived the need to assess the protocol of the research database. The METC approval number was P18 053.

ORCID iDs

Michelle Oud https://orcid.org/0000-0002-0988-0345
Sebastiaan Breedveld https://orcid.org/0000-0001-8954-4554
Jesús Rojo-Santiago https://orcid.org/0000-0002-4598-7430
Marta Krystyna Giżyńska https://orcid.org/0000-0001-5880-6814
Michiel Kroesen https://orcid.org/0000-0002-1346-2277
Zoltán Perkó https://orcid.org/0000-0002-0975-4226
Ben Heijmen https://orcid.org/0000-0003-1647-0528
Mischa Hoogeman https://orcid.org/0000-0002-4264-9903

References

Apolle R *et al* 2019 Inter-observer variability in target delineation increases during adaptive treatment of head-and-neck and lung cancer *Acta Oncol.* **58** 1378–85
Bernatowicz K *et al* 2018 Feasibility of online IMPT adaptation using fast, automatic and robust dose restoration *Phys. Med. Biol.* **63** 085018
Bertholet J *et al* 2020 Patterns of practice for adaptive and real-time radiation therapy (POP-ART RT): II. Offline and online plan adaption for interfractional changes *Radiother. Oncol.* **153** 88–96
Bobić M *et al* 2021 Comparison of weekly and daily online adaptation for head and neck *intensity-modulated proton therapy Phys. Med. Biol.* **66** 055023
Bobić M *et al* 2023 Large anatomical changes in head-and-neck cancers – A dosimetric comparison of online and offline adaptive proton therapy *Clin. Trans. Radiat. Oncol.* **40** 100625
Borderias Villarroel E L *et al* 2023 Dose mimicking based strategies for online adaptive proton therapy of head and neck cancer *Phys. Med. Biol.* **68** 105002
Borderías-Villarroel E *et al* 2022 Evaluation of the clinical value of automatic online dose restoration for adaptive proton therapy of head and neck cancer *Radiother. Oncol.* **170** 190–7
Botas P *et al* 2018 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.* **64** 015004
Breedveld S *et al* 2012 iCycle: integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans *Med. Phys.* **39** 951–63
Breedveld S, Storchi P R M and Heijmen B J M 2009 The equivalence of multi-criteria methods for radiotherapy plan optimization *Phys. Med. Biol.* **54** 7199
Burlacu T, Lathouwers D and Perkó Z 2023 A deterministic adjoint-based semi-analytical algorithm for fast response change computations in proton therapy *J. Comput. Theor. Transp.* **51** 1–41
Fredriksson A, Forsgren A and Hårdemark B 2011 Minimax optimization for handling range and setup uncertainties in proton therapy *Med. Phys.* **38** 1672–84
Guo H *et al* 2021 The dosimetric impact of deep learning-based auto-segmentation of organs at risk on nasopharyngeal and rectal cancer *Radiat. Oncol.* **16** 1–14
Heijmen B *et al* 2018 Fully automated, multi-criterial planning for volumetric modulated arc therapy–an international multi-center validation for prostate cancer *Radiother. Oncol.* **128** 343–8
Huiskes M *et al* 2024 Validation of Fully Automated Robust Multicriterial Treatment Planning for Head and Neck Cancer IMPT *Int J Radiat Oncol Biol Phys.*
Jagt T *et al* 2017 Near real-time automated dose restoration in IMPT to compensate for daily tissue density variations in prostate cancer *Phys. Med. Biol.* **62** 4254
Kooy H M *et al* 2010 A case study in proton pencil-beam scanning delivery *Int. J. Radiat. Oncol.* Biol.* Phys.* **76** 624–30
Korevaar E W *et al* 2019 Practical robustness evaluation in radiotherapy—a photon and proton-proof alternative to PTV-based plan evaluation *Radiother. Oncol.* **141** 267–74
Lalonde A *et al* 2020 Evaluation of CBCT scatter correction using deep convolutional neural networks for head and neck adaptive proton therapy *Phys. Med. Biol.* **65** 245022
Lalonde A *et al* 2021 Anatomic changes in head and neck intensity-modulated proton therapy: comparison between robust optimization and online adaptation *Radiother. Oncol.* **159** 39–47

Li H et al 2013 Use of treatment log files in spot scanning proton therapy as part of patient-specific quality assurance Med. Phys. 40 021703
Liu W et al 2013 Effectiveness of robust optimization in intensity-modulated proton therapy planning for head and neck cancers Med. Phys. 40 051711
Matter M et al 2019 Intensity modulated proton therapy plan generation in under ten seconds Acta Oncol. 58 1435–9
Meier G et al 2015 Independent dose calculations for commissioning, quality assurance and dose reconstruction of PBS proton therapy Phys. Med. Biol. 60 2819
Meijers A et al 2020 Feasibility of patient specific quality assurance for proton therapy based on independent dose calculation and predicted outcomes Radiother. Oncol. 150 136–41
Miyazaki K et al 2022 Deformed dose restoration to account for tumor deformation and position changes for adaptive proton therapy Med. Phys. 50 675–87
Nash D et al 2022 Suitability of propagated contours for adaptive replanning for head and neck radiotherapy Phys. Med. 102 66–72
Nenoff L et al 2019 Daily adaptive proton therapy–the key to innovative planning approaches for paranasal cancer treatments Acta Oncol. 58 1423–8
Nenoff L et al 2021 Experimental validation of daily adaptive proton therapy Phys. Med. Biol. 66 205010
National Association for Radiotherapy in the Netherlands 2019 Landelijk Indicatieprotocol Protonentherapie Hoofdhals v2.2. Retrieved from https://nvro.nl/images/documenten/rapporten/2019-08-15__Landelijk_Indicatieprotocol_Protonentherapie_Hoofdhals_v2.2.pdf
Oud M et al 2022 An online adaptive plan library approach for intensity modulated proton therapy for head and neck cancer Radiother. Oncol. 176 68–75
Paganetti H et al 2021 Adaptive proton therapy Phys. Med. Biol. 66 22TR01
Park Y K et al 2015 Proton dose calculation on scatter-corrected CBCT image: feasibility study for adaptive proton therapy Med. Phys. 42 4449–59
Pastor-Serrano O and Perkó Z 2022 Millisecond speed deep learning based proton dose calculation with Monte Carlo accuracy Phys. Med. Biol. 67 105006
Perkó Z et al 2016 Fast and accurate sensitivity analysis of IMPT treatment plans using polynomial chaos expansion Phys. Med. Biol. 61 4646
Qiu Z et al 2023 Online adaptive planning methods for intensity-modulated radiotherapy Phys. Med. Biol. 68 10TR01
Rojo-Santiago J et al 2021 Accurate assessment of a dutch practical robustness evaluation protocol in clinical PT with pencil beam scanning for neurological tumors Radiother. Oncol. 163 121–7
Rojo-Santiago J et al 2023 PTV-based VMAT versus robust IMPT for Head-and-Neck Cancer: a probabilistic uncertainty analysis of clinical plan evaluation with the dutch model-based selection Radiother. Oncol. 186 109729
Segedin B and Petric P 2016 Uncertainties in target volume delineation in radiotherapy—are they relevant and what can we do about them? Radiol. Oncol. 50 254–62
Smolders A J et al 2023 Dosimetric comparison of autocontouring techniques for online adaptive proton therapy Phys. Med. Biol. 68 175006
Taasti V T et al 2018 Comparison of single and dual energy CT for stopping power determination in proton therapy of head and neck cancer Phys. Imaging Radiat. Oncol. 6 14–9
Thummerer A et al 2020 Comparison of CBCT based synthetic CT methods suitable for proton dose calculations in adaptive proton therapy Phys. Med. Biol. 65 095002
Trnkova P et al 2023 A survey of practice patterns for adaptive particle therapy for interfractional changes Phys. Imaging Radiat. Oncol. 26 100442
van der Veen J, Gulyban A and Nuyts S 2019 Interobserver variability in delineation of target volumes in head and neck cancer Radiother. Oncol. 137 9–15
van der Voort S et al 2016 Robustness recipes for minimax robust optimization in intensity modulated proton therapy for oropharyngeal cancer patients Int. J. Radiat. Oncol.* Biol.* Phys. 95 163–70
van de Water S et al 2013 Improved efficiency of multi-criteria IMPT treatment planning using iterative resampling of randomly placed pencil beams Phys. Med. Biol. 58 696919
van de Water S et al 2016 The price of robustness; impact of worst-case optimization on organ-at-risk dose and complication probability in intensity-modulated proton therapy for oropharyngeal cancer patients Radiother. Oncol. 120 56–62