Probabilistic evaluation guided IMPT planning with realistic setup and range uncertainties improves the trade-off between OAR sparing and target coverage in neuro-oncological patients

Jong, J.I. de; Habraken, S.J.M.; Rojo-Santiago, J.; Lathouwers, D.; Perkó, Z.; Breedveld, S.; Hoogeman, M.S.

Citation

Jong, J. I. de, Habraken, S. J. M., Rojo-Santiago, J., Lathouwers, D., Perkó, Z., Breedveld, S., & Hoogeman, M. S. (2025). Probabilistic evaluation guided IMPT planning with realistic setup and range uncertainties improves the trade-off between OAR sparing and target coverage in neuro-oncological patients. *Radiotherapy &Amp; Oncology*, *213*. doi:10.1016/j.radonc.2025.111171

Radiotherapy and Oncology 213 (2025) 111171

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

Probabilistic evaluation guided IMPT planning with realistic setup and range uncertainties improves the trade-off between OAR sparing and target coverage in neuro-oncological patients

Jenneke I. de Jong [a,b,*], Steven J.M. Habraken [c,b], Jesús Rojo-Santiago [a,b], Danny Lathouwers [d], Zoltán Perkó [d], Sebastiaan Breedveld [a], Mischa S. Hoogeman [a,b]

[a] Radiotherapy, Erasmus MC Cancer Institute, Rotterdam, Netherlands
[b] Medical Physics & Informatics, HollandPTC, Delft, Netherlands
[c] Radiation Oncology, Leiden University Medical Center, Leiden, Netherlands
[d] Radiation Science and Technology, Delft University of Technology, Delft, Netherlands

ARTICLE INFO

Keywords:
Probabilistic planning
Robustness evaluation
Proton therapy
IMPT
Polynomial chaos expansion
Neuro-oncological

ABSTRACT

Objective: Scenario-based evaluation in proton therapy often relies on a small number of error scenarios, leading to limited insight into the DVH values under uncertainty and suboptimal trade-offs. In this study, we investigated if re-optimization based on probabilistic evaluation improves the trade-off between OAR sparing and target coverage in neuro-oncological patients.

Materials and methods: 22 neuro-oncological patients were included. 18 met their original target goals (group A), while in 4, target coverage was compromised to spare OARs (group B).

The probabilistic goal for the CTV was calibrated to be consistent with PTV-based photon plans, resulting in $D_{99.8\%,CTV} = 0.95D_{pres}$ with a 90 % confidence level. The probabilistic OAR constraints were set to meet the clinical constraints with a 95 % confidence level.

For both groups, the clinical plans were re-optimized, keeping the clinical objectives and constraints, but reducing robustness for the CTV objective (group A) to meet the probabilistic goal, or for the dose-limiting OAR objectives (group B) without exceeding the constraints.

For the original and re-optimized plans, polynomial chaos expansion was applied to simulate 10,000 fractionated treatments, deriving probability distributions for relevant DVH parameters.

Results: For group A, re-optimization resulted in a population median decrease of 8.2 (range: 0.4–20.8) Gy RBE in the total OAR-related clinical goal values.

For group B, re-optimization resulted in a population median increase of 2.7 (range: 1.3–6.8) Gy RBE in the $D_{99.8\%,CTV}$. The population median $V_{95\%,CTV}$ improved from 97.4 % to 99.1 %.

Conclusion: We demonstrated that probabilistic evaluation guided IMPT planning enables either OAR sparing or target coverage enhancement.

Introduction

In radiotherapy, mitigating the impact of uncertainties, such as anatomical changes, patient positioning and beam (setup) errors, is crucial. In photon therapy, this is achieved through the planning target volume (PTV), generated as an expansion margin around the clinical target volume (CTV). This approach works because the shape of photon dose distributions is largely unaffected by setup errors, a principle known as the static dose cloud approximation [1].

Margin-based approaches are not applicable to intensity-modulated proton therapy (IMPT) with pencil-beam scanning (PBS). This is caused by highly modulated beam doses with steep gradients, combined with the medium-dependence of the proton stopping power and the stopping-power prediction (range) error [2–4]. While range-adapted PTVs are applied for evaluation and optimization in proton therapy with passive scattering [5,6], more advanced scenario-based evaluation and optimization strategies [7,8] are used in IMPT.

International consensus on scenario-based evaluation is lacking,

---

* Corresponding author at: P.O. Box 2040, Doctor Molewaterplein 40, 3015GD Rotterdam, The Netherlands.
E-mail address: j.i.dejong@erasmusmc.nl (J.I. de Jong).

https://doi.org/10.1016/j.radonc.2025.111171
Received 23 May 2025; Received in revised form 18 August 2025; Accepted 15 September 2025

leading to variability across different centers [9]. Common practices include the simulation of a set of error scenarios based on systematic setup and range errors, with magnitudes calibrated according to the specific protocols of individual centers [8]. For visualization, the dose-volume histograms (DVH) and the DVH parameters belonging to the clinical goals are often presented for the different error scenarios [8,10].

In the Netherlands, scenario-based evaluation has been standardized among the three proton centers, involving the voxel-wise minimum (VWmin) and maximum (VWmax) dose distributions to assess CTV and OAR dose respectively [11]. This approach is based on 28 error scenarios and has been calibrated to maintain consistency with PTV-based evaluation in photon therapy [11].

However, such scenario-based evaluation methods involve only a small number of error scenarios and, therefore, do not provide a statistically robust framework to obtain insight into the DVH parameter values under uncertainty. This is exemplified by the VWmin/VWmax dose distributions, which are virtual and do not represent real dose distributions in any actual physical error scenario. This introduces two key limitations in clinical practice:

1. **High inter-patient variation**: While VWmin and VWmax DVH parameter values may be consistent across different patients, there is often significant variability in actual DVH parameter values under uncertainty [12]. In earlier work in a neuro-oncological patient group we found that despite all patients receiving the exact same VWmin $D_{98\%,CTV}$ value, the 10th percentile of the probability distribution of the $D_{98\%,CTV}$ ranged from $0.956D_{pres}$ to $0.973D_{pres}$ across patients [13].
2. **Suboptimal trade-offs:** We hypothesize that for patients with trade-offs between achieving tumor coverage and OAR sparing, the ability to effectively balance these goals may be compromised due to limited insight into the DVH parameter values under uncertainty.

Probabilistic evaluation, based on a large number of error scenarios [14–17], may resolve these limitations. The fast simulation of numerous error scenarios has recently been enabled by advancements in algorithms, such as GPU-based Monte Carlo calculations [18,19] and Polynomial Chaos Expansion (PCE) [20]. Until now, probabilistic evaluation has primarily served as a gold standard in research for analyzing current clinical methods [12,21–24]. However, the potential clinical benefits of re-optimization based on probabilistic evaluation are yet unexplored.

In this paper, we aim to re-optimize clinical plans based on probabilistic evaluation and assess whether it leads to a reduction in OAR doses or improved target coverage in cases in which this was originally sacrificed. We investigate our approach in a cohort of recently treated neuro-oncological patients.

Materials and methods

Clinical plan creation and evaluation

The clinical plans of 22 neuro-oncological patients treated at HollandPTC in 2023 were included. Prescribed doses (D_{pres}) were 50.4 Gy RBE (16/22) or 59.4 Gy RBE (6/22), delivered in 1.8 Gy RBE fractions. All plans were created in RayStation 10b using the clinical beam model and the Monte Carlo (MC) dose engine with a MC uncertainty of 1 %.

All clinical plans were generated using a single-field optimization IMPT (SFO-IMPT) approach using two or three beams. Robustness against setup and range errors was aimed for through minimax optimization based on 21 error scenarios. The 21 error scenarios were formed by a combination of a ±3 mm setup error [25] in each of three principal directions and a stopping-power prediction (range) error of -3 %, 0 %, or $+3$ %. The CTV objectives and critical OAR constraints were set to ‘robust’, guiding the optimizer to achieve the desired outcomes across all error scenarios.

The clinical plans were evaluated based on 28 error scenarios [11]. These error scenarios included 14 setup error combinations with directions determined by the 6 faces and 8 vertices of a cube, all normalized to 3 mm. The setup errors were combined with range errors of -3 % or 3 %. CTV dose was evaluated using the VWmin dose distribution of the 28 error scenarios, with the clinical goal being VWmin $D_{98\%,CTV} = 0.95D_{pres}$. The doses to the brainstem and optic system were evaluated using the VWmax dose distribution of the 28 error scenarios.

Out of the 22 plans, 18 met the clinical robustness criteria for the CTV (group A). In the remaining 4 cases, CTV coverage was compromised to spare the brainstem or the optic system (group B).

Polynomial chaos expansion (PCE)

PCE was used to determine the dose distribution under varying error scenarios [20]. The dose D_i to each voxel is dependent on a setup error $\vec{\xi} = \left(\xi_x, \xi_y, \xi_z\right)$ and a relative range error ρ and was approximated by the analytical series expansion $D_i\left(\vec{\xi},\rho\right) = \sum_{k=0}^{p} a_{i,k}\Psi_k\left(\vec{\xi},\rho\right)$, with expansion coefficients $a_{i,k}$ and multi-dimensional Hermite polynomials $\Psi_k\left(\vec{\xi},\rho\right)$. These coefficients were determined through linear regression using 208 input–output combinations, where the input was an error scenario $\left(\vec{\xi},\rho\right)$ and the output was the corresponding dose $D_i\left(\vec{\xi},\rho\right)$ for each voxel i calculated within the treatment planning system (TPS) with a MC uncertainty of 1 % [20], assuming Gaussian distributions. PCE was previously validated for this application in [20].

Probabilistic evaluation

For every clinical plan, PCE was used to simulate 10,000 fractionated treatments. For every treatment, a systematic setup error Σ (mean: 0.00 mm, SD: 0.92 mm) and a range error ρ (mean: 1.2 %, SD: 1 %) were randomly sampled from their Gaussian distribution. The 1.2 % range error mean accounts for the CT-to-proton stopping-power prediction (SPP) underestimation based on the Gammex 467 phantom calibration [26]. Additionally, a randomly sampled random setup error σ (mean: 0.00 mm, SD: 1.00 mm, also Gaussian) was applied to each fraction. The means and standard deviations for the uncertainties were chosen to align with the 3 mm margin M according to Van Herk’s margin recipe: $M = 2.5\Sigma + 0.7\sigma$[25], leading to values close to the ones measured in a head & neck cohort of our center [21] Fig. 1 illustrates both the current clinical evaluation method and the probabilistic evaluation method.

Determining the probabilistic clinical goals

The probabilistic goal for the CTV was calibrated based on the prescription: $D_{98\%,CTV} = 0.94D_{pres}$, ensuring consistency with photon plans according to [11]. Since the clinical plans were originally based on the slightly more conservative prescription of VWmin $D_{98\%,CTV} = 0.95D_{pres}$, all dose distributions were scaled to exactly meet VWmin $D_{98\%,CTV} = 0.94D_{pres}$ before calibration. After scaling, the probability distribution of the volume that receives 95 % of D_{pres} ($V_{95\%,CTV}$) was generated for each patient using PCE. For each patient, the 10th percentile of this distribution was identified, representing the volume receiving at least $0.95D_{pres}$ with a 90 % probability [Supplementary Material 1]. To finalize the calibration, the median of these 10th percentile values across all patients was calculated to define the probabilistic $V_{95\%,CTV}$ goal for this group. We found a median 10th percentile $V_{95\%,CTV}$ value of 99.8 % across all patients. This results in the probabilistic goal: $D_{99.8\%,CTV} = 0.95D_{pres}$ with a 90 % probability. The probabilistic OAR constraints were based on meeting the constraints of the EPTN guidelines [27] with a 95 % probability, considered reasonable from a clinical perspective.

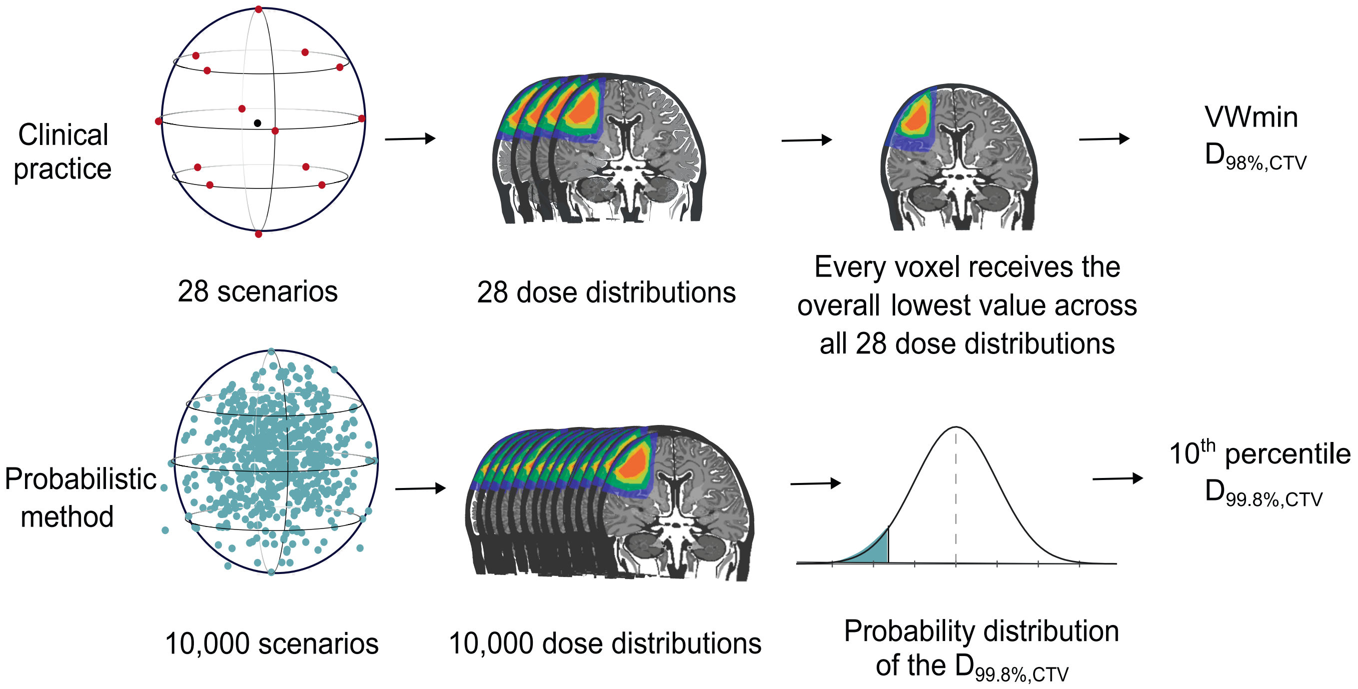

Fig. 1. Visualization of the difference between the clinical evaluation method and the probabilistic evaluation method. The clinical method is based on 28 scenarios and the corresponding VWmin dose distribution that is formed by taking the overall lowest values across all distributions for every voxel. The probabilistic method is based on 10,000 scenarios that allow us to create the full probability distributions of dosimetric outcomes. The turquoise shaded lower tail of the probability distribution corresponds to 10 % of the sampled scenarios with the lowest $D_{99.8\%,CTV}$ values. The choice of the $D_{99.8\%,CTV}$ as the DVH parameter for probabilistic evaluation is explained in the Methods section under 'Determining the probabilistic clinical goals'.

Re-optimization of clinical plans based on probabilistic evaluation

Re-optimization based on probabilistic evaluation consisted of two steps: (i) re-optimization in RayStation and (ii) scaling. During re-optimization in RayStation (i), robustness was deliberately reduced for either the CTV (group A) or the dose-limiting OARs (group B) to approach the newly set probabilistic CTV goal and OAR constraints (Fig. 2, step 1). Normally, robust optimization requires that the CTV and critical OAR objectives are satisfied across all 21 error scenarios. However, for the re-optimization step, these error scenarios were excluded from the optimization process. As a result, the optimizer focused solely on achieving the optimization objectives under nominal conditions for the CTV (group A) or the dose-limiting OARs (group B), without accounting for setup or range uncertainties. The original pencil-beam energies and positions were retained, preserving some degree of robustness. Re-optimization was done through 180 iterations, found to be sufficient to reach convergence.

After re-optimization in RayStation, dose scaling (ii) ensured that the probabilistic goal for the CTV was achieved exactly. For group B, if the target dose could still be increased without exceeding the probabilistic OAR constraints, upwards scaling of the dose was applied until either the probabilistic CTV goal was met or one of the probabilistic OAR constraints was reached (Fig. 2, step 2). If probabilistic OAR constraints were exceeded after re-optimization, downwards scaling of the dose was applied to ensure that none of the OAR constraints were surpassed.

To evaluate the value of re-optimization, we also created plans where scaling alone was used to meet the newly set probabilistic CTV goal and OAR constraints.

Comparison of clinical and re-optimized plans

For every patient, a probabilistic evaluation was conducted for both the clinical plan and the re-optimized plan. The comparison was based on the DVH parameters belonging to the clinical goals used in clinical practice [Supplementary Material 2]. For DVH parameters belonging to the CTV, the 10th percentile values were selected, while for the OAR-related DVH parameters, the 95th percentile values were selected. The difference between the clinical plan value and the re-optimized plan value was determined for every DVH parameter. Additionally, the total difference was evaluated by summing the differences across all OAR-related DVH parameters.

Results

For group A, the median decrease for the sum of all OAR-related DVH parameters listed in Fig. 3 was 8.2 (range: 0.4–20.8) Gy RBE, with individual patient reductions up to 20.8 Gy RBE (Table 1). Scale factors applied after re-optimization had a median of 1.00 (range: 0.97–1.03). Scaling alone resulted in a median decrease of 2.8 Gy RBE for the sum of all OAR-related DVH parameters, which is 5.4 Gy RBE lower than the reduction found after re-optimization and scaling. This can be explained by the fact that re-optimization not only shifts the probability density functions (PDF) but also alters the shape, offering more optimization space [Supplementary Material 3].

Dose reductions were particularly achieved for the brainstem and optic system (Fig. 4). Specifically, for the $D_{0.03cc,Brainstem\_Core}$ and the $D_{0.03cc,Brainstem\_Surf}$ reductions up to 3.3 and 3.5 Gy RBE were achieved. The $D_{0.03cc,OpticNrv\_L}$, the $D_{0.03cc,OpticNrv\_R}$ and the $D_{0.03cc,OpticChiasm}$ showed reductions up to 4.6, 3.6 and 3.6 Gy RBE, respectively. A few cases had slight increases in individual DVH parameters, but none exceeded clinical constraints. The $D_{99.8\%,CTV}$ decreased with a median of 1.1 (range: 0.3–1.9) Gy RBE without compromising clinical goal achievement.

In the original plan, the $D_{99.8\%,CTV}$ ranged from $0.956D_{pres}$ to $0.988D_{pres}$,whereas re-optimization eliminated all interpatient variation in the $D_{99.8\%,CTV}$ [Supplementary material 4].

For group B, re-optimization resulted in a mean increase of 2.7 (range: 1.3–6.8) Gy RBE in the $D_{99.8\%,CTV}$ (Fig. 4). The population

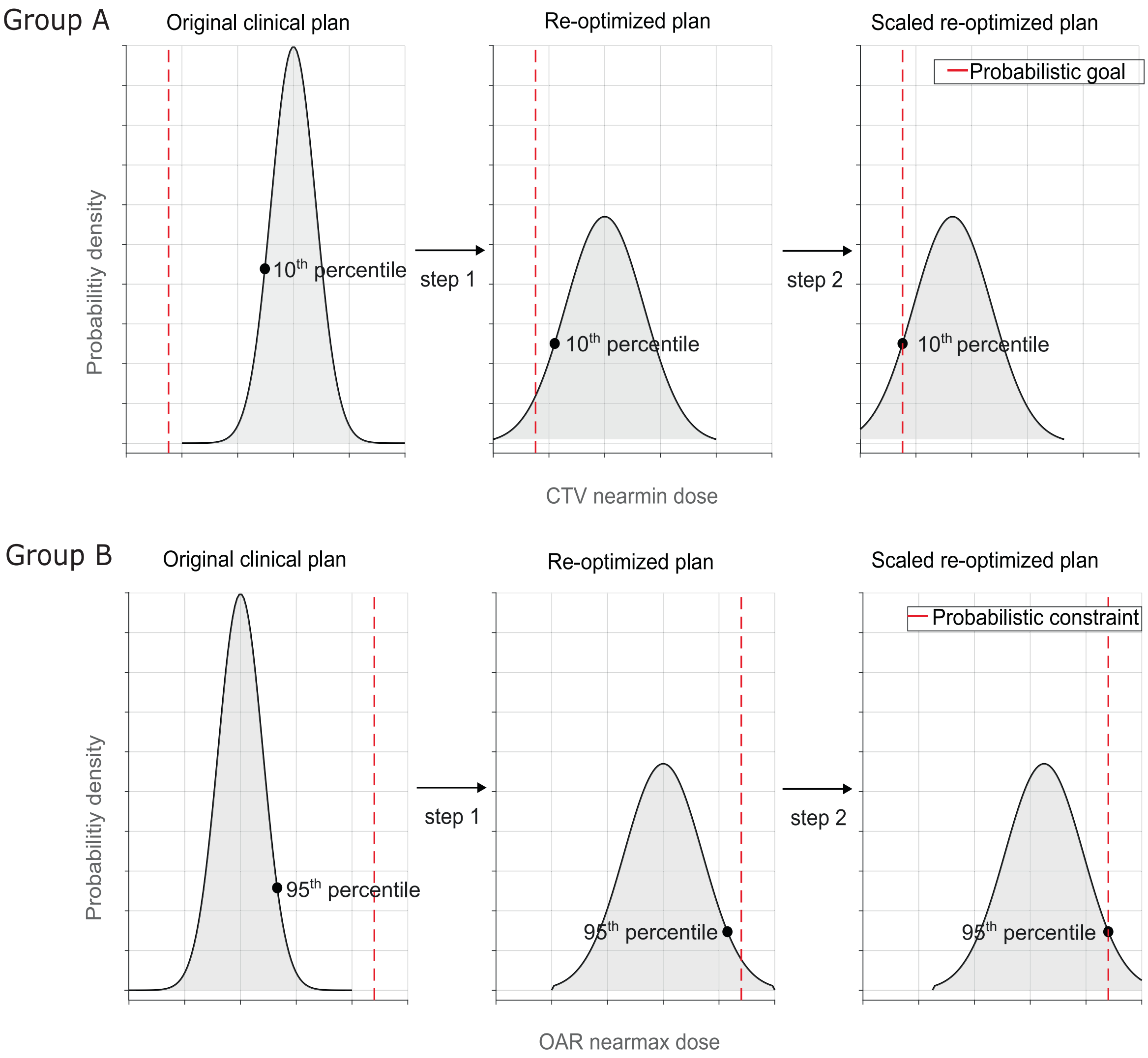

Fig. 2. Visualization of the re-optimization method for the clinically robust patients (group A) and the patients in which target coverage was originally sacrificed (group B). For group A, the first plot represents the original clinical plan, where the probability distribution of the CTV dose shows a clear gap between its 10th percentile value and the probabilistic goal, indicating that there is space to decrease this value. The second plot presents the results after re-optimization. The final plot presents the results after fine-tuning with scaling. For group B, the plots illustrate the change of the probability distribution for the DVH parameter associated with the OAR that prevents the CTV from receiving the prescribed dose. The first plot represents the original clinical plan, where the probability distribution shows a clear gap between the 95th percentile of the curve and the probabilistic constraint, indicating that there is space to increase the OAR (and CTV) dose value. The second plot presents the results after re-optimization. The final plot presents the results after fine-tuning with scaling to ensure that the 95th percentile of the DVH parameter is precisely aligned with the probabilistic constraint.

median of the $V_{95\%,CTV}$ improved from 97.4 % to 99.1 %. For patient 21, the probabilistic clinical goal for the CTV was achieved after re-optimization. Additionally, the dose limiting OARs experienced an average increase of 1.4 (range: –0.2–6.9) Gy RBE without exceeding the clinical constraints.

The DVH parameter values for the OARs that were not dose-limiting showed a change of –4.4 to 1.4 (median: 0.0) Gy RBE.

Fig. 5 presents examples of the nominal dose distribution and DVH curves for a representative patient from both groups.

Discussion

In this paper, we aimed to evaluate the impact of re-optimization based on probabilistic evaluation to determine whether it improves the trade-off between OAR sparing and target coverage in neuro-oncological patients.

Our main findings are that integrating re-optimization based on probabilistic evaluation increases OAR sparing or improves target coverage when originally sacrificed. Additionally, we observed a large reduction in interpatient variation in the probabilistic $D_{99.8\%,CTV}$, highlighting the potential for more consistent dosimetric outcomes across the patient population.

We found a population-based median $V_{95\%,CTV}$ of 99.8 % for our neuro-oncological cohort, in agreement with the value found in a previous cross-calibration of IMPT and VMAT plans for head-and-neck cancer [21]. The fact that dose should be prescribed to a larger

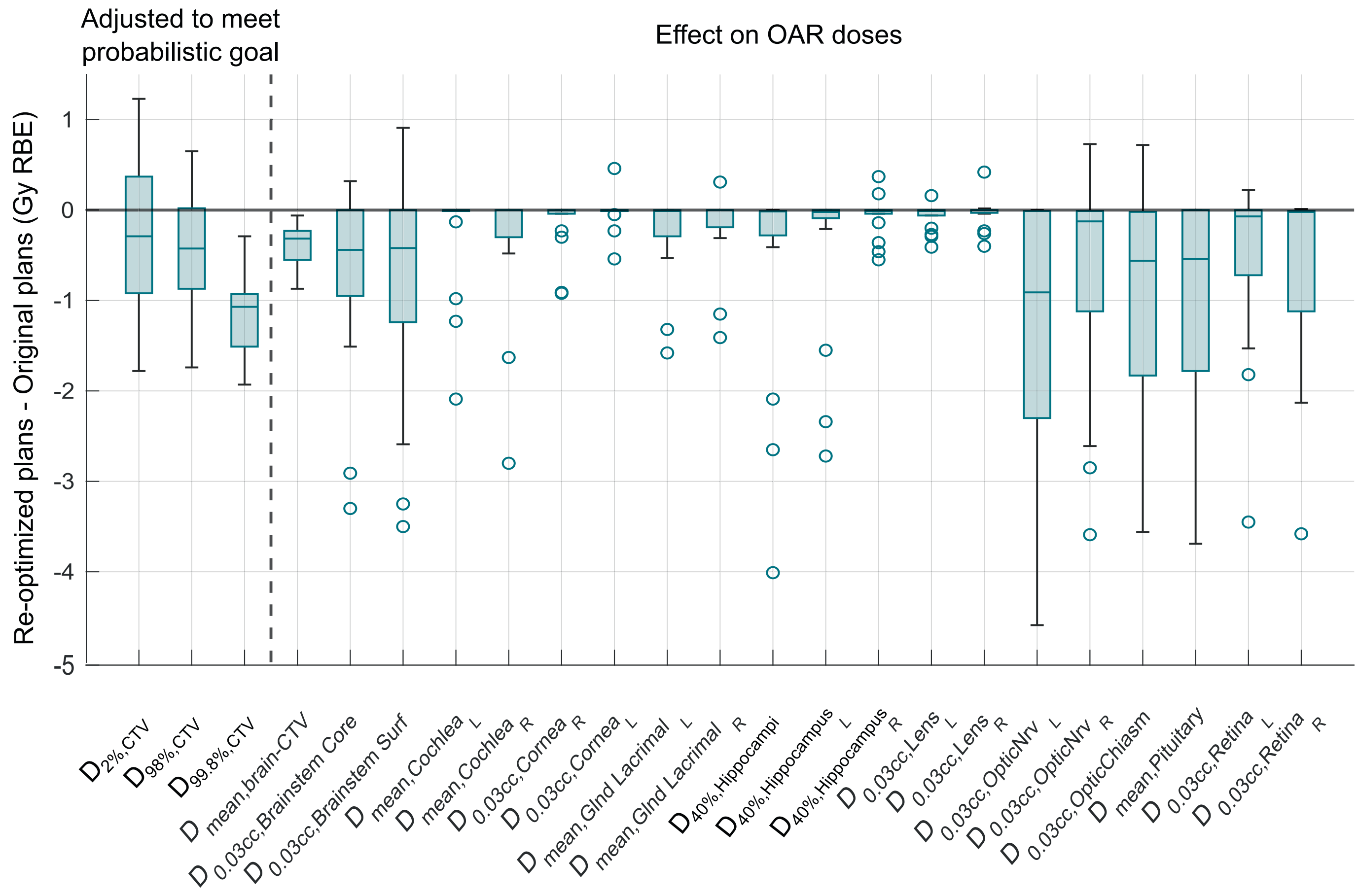

Fig. 3. Results for group A. The boxplots illustrate the differences between the original clinical values and the values after re-optimization for all DVH parameters associated with the clinical goals used in treatment plan evaluation. For DVH parameters belonging to the CTV, the 10th percentile values of their probability distributions were selected, while for the OAR-related DVH parameters, the 95th percentile values of their probability distributions were selected. The upper and lower edges of the box represent the 0.75 and 0.25 quantiles. The whiskers correspond to the most extreme data points that are still within 1.5 times the interquartile distance from the upper and lower edges of the box. All values beyond this range are considered outliers and are shown as individual data points.

Table 1
Total reduction in summed values of all DVH parameters associated with clinical goals used in treatment plan evaluation after re-optimization based on D99.8%, CTV = 0.95Dpres with a 90 % probability. For every DVH parameter, the 95th percentile value of it probability distribution was selected.

Patient	Total OAR dose reduction (Gy RBE)	Scale factor
1	−5.3	0.997
2	−2.8	1.008
3	−7.4	0.986
4	−19.3	1.009
5	−8.9	1.018
6	−12.2	1.019
7	−17.7	0.981
8	−0.9	0.970
9	−10.5	0.996
10	−5.8	1.001
11	−3.4	1.018
12	−20.3	0.986
13	−20.8	1.033
14	−0.4	0.990
15	−19.0	0.994
16	−0.6	0.985
17	−0.8	0.988
18	−12.5	1.003
Median	−8.2	0.997

percentage of the CTV volume rather than 98 % for consistency with photons can be expected, reflecting that the CTV is a subvolume of the PTV.

The results of this study demonstrate that it was possible to reduce OAR dose in all patients of group A while still achieving the probabilistic CTV goal with a 90 % probability. This finding can be explained by our current planning approach based on VWmin $D_{98\%,\mathrm{CTV}} = 0.95D_{\mathrm{pres}}$, which is conservative compared to [11]. However, the magnitude of OAR dose reduction varied among patients. One important contributor to this variability was the variation in the $D_{99.8\%,\mathrm{CTV}}$ values of the original clinical plans, which affected the potential for dose reduction.

Additionally, our study showed that CTV coverage could be improved when it was initially compromised. This can be attributed to the VWmax method predicting OAR dose values close to the clinical constraints, while in reality, with a 95 % probability, these values were lower, consistent with findings from previous research [12].

Another observation is that re-optimization based on probabilistic evaluation removed interpatient variation in the $D_{99.8\%,\mathrm{CTV}}$, reflecting its ability to directly evaluate any DVH parameter. In the original clinical plans, interpatient variation in the $D_{99.8\%,\mathrm{CTV}}$ was substantial. This finding is not surprising, as the VWmin $D_{98\%,\mathrm{CTV}}$ does not directly measure the dose delivered to the ‘worst’ two percent of the CTV. Consequently, this leads to large variability in $D_{99.8\%,\mathrm{CTV}}$ values despite similar VWmin $D_{98\%,\mathrm{CTV}}$ values.

For this study, we limited ourselves to re-optimization of clinical plans, retaining the spot selection and optimization objectives and constraints. Further gains may be realized by optimizing from scratch, which would allow for the selection of different spots and increase the optimization space.

Furthermore, the values for the standard deviations of the systematic

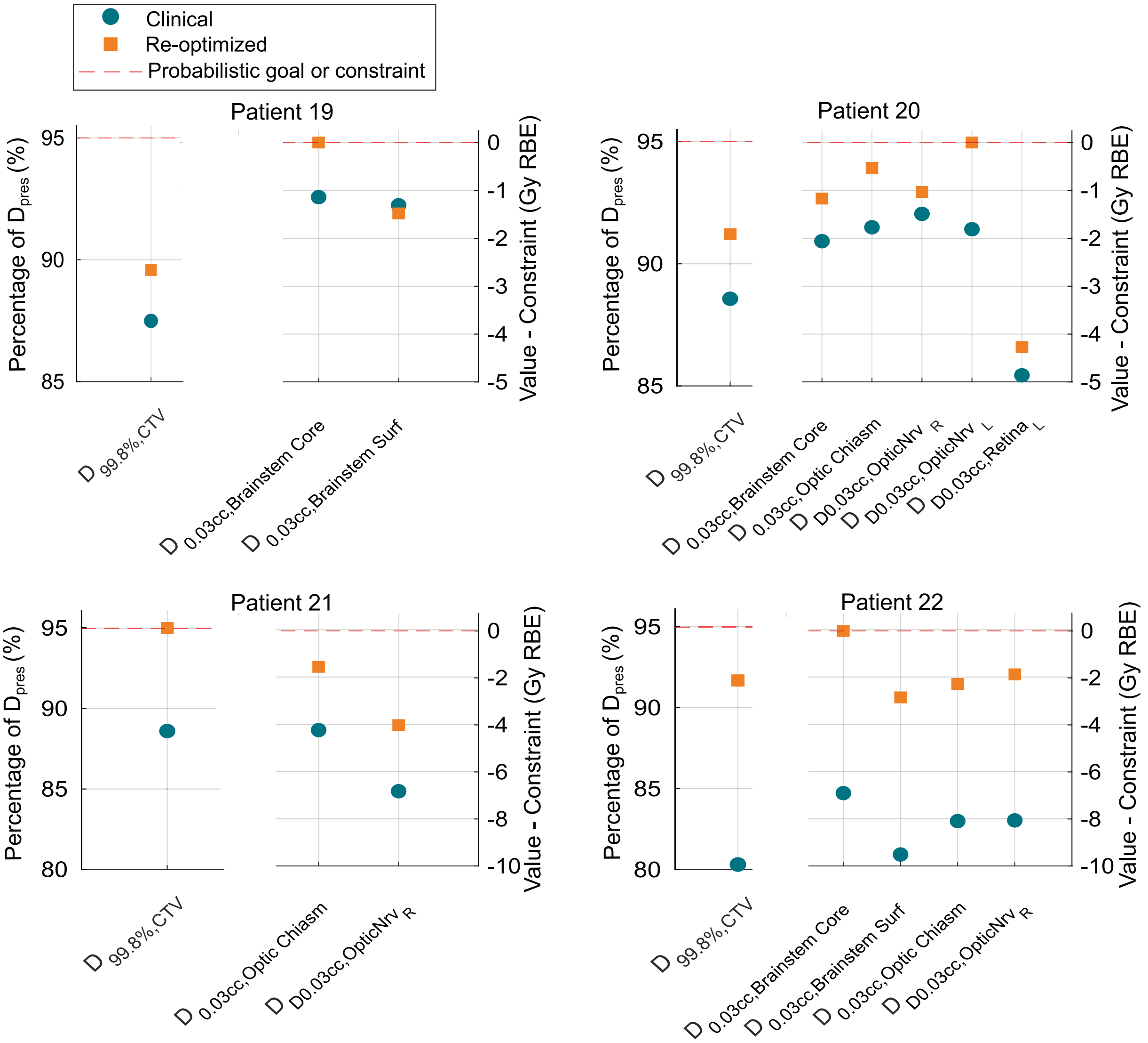

Fig. 4. Results for group B. The plots illustrate the differences between the original values and the values after re-optimization for the $D_{99.8\%,CTV}$ and the OARs that were the reason that target coverage was sacrificed in the original clinical plan. For the DVH parameter belonging to the CTV, the 10th percentile value of its probability distribution was selected, while for the OAR-related DVH parameters, the 95th percentile values of their probability distributions were selected. For patient 21 we were able to meet the probabilistic goal for the CTV with a 90 % probability. For all other patients, one of the probabilistic OAR constraints was the limiting factor.

and random errors in the probabilistic evaluation were chosen to align with a 3 mm margin based on van Herk's margin recipe, to ensure consistency between the clinical plans and the re-optimized plans. However, we know that the standard deviations at our center are slightly smaller in reality [22], which would make achieving plan robustness easier and provide additional optimization space. Nevertheless, there are also factors that could decrease the optimization space. For the probabilistic OAR constraints, we ensured that DVH parameters did not exceed the clinical constraints with 95 % probability, but stricter adherence may be necessary for critical OARs. Individualized trade-offs for each patient could further enhance the approach.

For the calibration of the CTV goal, we scaled all plans to VWmin $D_{98\%,CTV} = 0.94D_{pres}$, which is consistent with photon plans as outlined in [11]. However, since our calibration was based on this previous calibration rather than directly incorporating photon plans in our research, this increases the potential for inaccuracies.

In addition, there is a discrepancy between the ±3 % range errors used in standard clinical robust optimization and the values suggested in literature (mean = 1.2 %, SD = 1 %) [26]. The ±3 % values were originally chosen due to limited knowledge of actual range uncertainties at the time. These values should be reconsidered to better align clinical robust optimization with our improved understanding of range uncertainty in neuro-oncological patients.

While our study focused exclusively on setup and range uncertainties, considered most important for the neuro-oncological treatment site, other sources of uncertainty, such as anatomical changes, rotations and LET, also affect IMPT treatment planning [28]. Future research is needed on how to incorporate these additional sources of uncertainty.

Finally, in this study, we evaluated the probabilities of meeting CTV and OAR constraints independently. This approach does not take into account the joint probability of meeting a combination of constraints, which may be as low as the product of the individual probabilities.

Before probabilistic evaluation can be implemented into clinical practice, several steps are needed. First, it is crucial to research what probabilistic evaluation metrics clinicians need to effectively use the large number of dose distributions for clinical decision-making.

Second, the trade-off between statistical robustness and evaluation

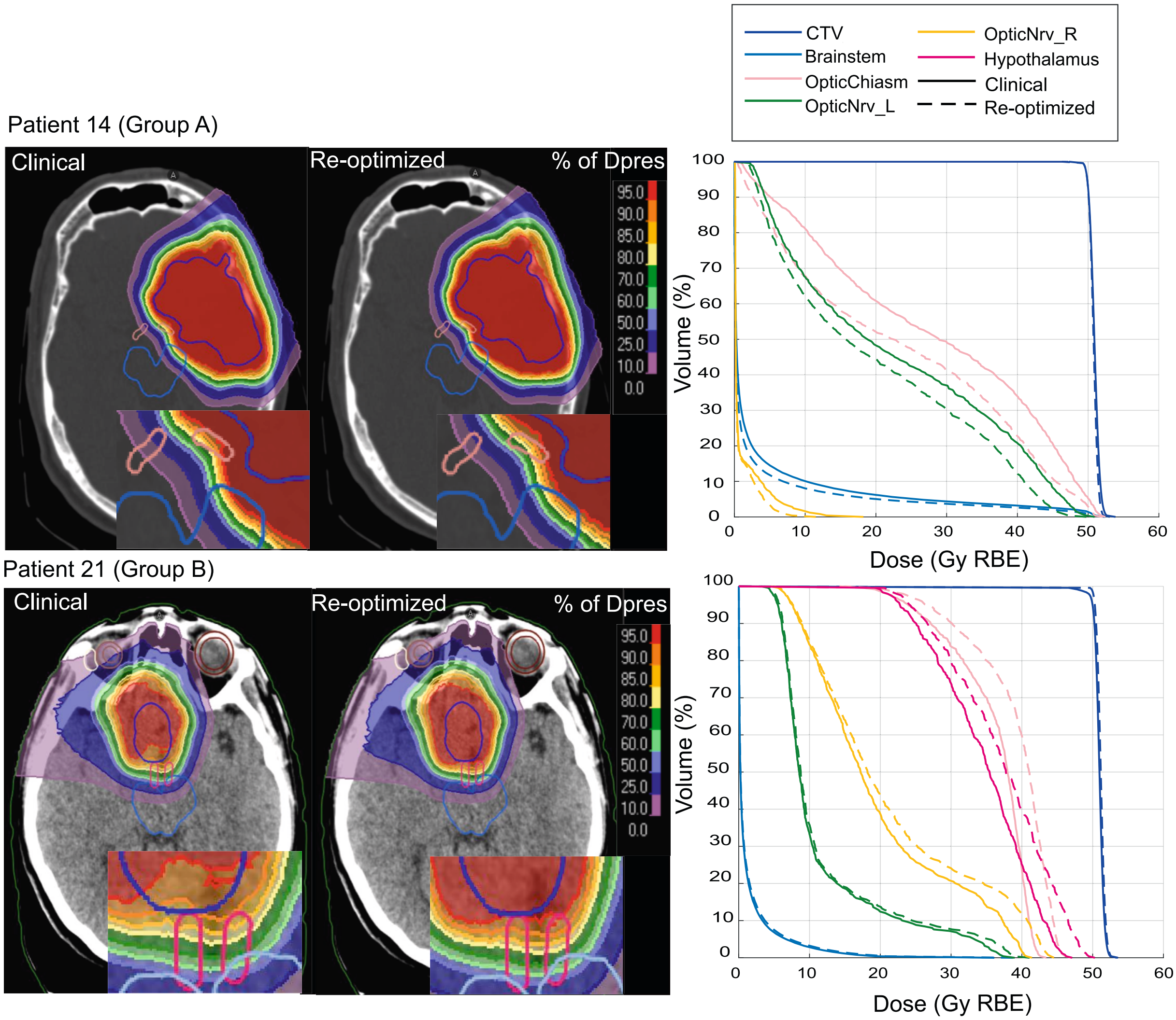

Fig. 5. The nominal dose distributions and DVH curves for both the clinical and re-optimized plans of patient 14 (Group A) and 21 (Group B), as representative examples. For Patient 14, re-optimization resulted in improved dose conformity to the tumor, with a reduction in the dose to the brainstem and optic system. For Patient 21, re-optimization improved target coverage at the expense of an increased dose to the hypothalamus and the optic system. The full probability distributions of the $D_{99.8\%,CTV}$ for patient 14 and the dose-limiting OAR-related DVH parameters for patient 21 before and after re-optimization can be found in Supplementary Material 3.

time must be explored. The method used in this study had an average computation time of 236 min, in contrast to the average computation time of 13 min for the method currently employed in clinical practice. Research is needed to determine if faster methods can provide statistically reliable results for clinical use.

Third, further insight is needed into how optimization settings impact probabilistic evaluation outcomes. The re-optimization approach used in this study worked well for our patient cohort due to the inherent robustness of the initial plans. However, this method may not be universally applicable, as planning techniques vary between centers. A study on re-optimization based on probabilistic evaluation using the optimization techniques employed in clinical practice could provide valuable insights. Future research should also focus on the feasibility of probabilistic optimization methods which directly incorporate probabilistic objectives.

Finally, future work could include validation of the probabilistic approach by analyzing the achievement of clinical goals on control CTs acquired during treatment, to assess how well the re-optimized plans perform under actual uncertainties.

In conclusion, to the best of our knowledge, this study is the first to quantify the gains of probabilistic evaluation guided IMPT planning. Re-optimization based on probabilistic evaluation enables either an increase in OAR sparing or enhancement of target coverage by effectively advancing personalized treatment planning.

CRediT authorship contribution statement

Jenneke I. de Jong: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Steven J.M. Habraken:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Methodology, Formal analysis, Conceptualization. **Jesús Rojo-Santiago:** Writing – review & editing, Software, Methodology, Conceptualization. **Danny Lathouwers:** Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Zoltán Perkó:** Writing – review & editing, Software, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Sebastiaan Breedveld:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Software, Methodology, Formal analysis, Conceptualization. **Mischa S. Hoogeman:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Funding

The research for this work was partly funded by RaySearch Laboratories and partly funded by the Surcharge for Top Consortia for Knowledge and Innovation (TKIs) from the Dutch Ministry of Economic Affairs and Climate.

Declaration of competing interest

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2025.111171.

References

[1] Karlsson K, et al. Accuracy of the dose-shift approximation in estimating the delivered dose in SBRT of lung tumors considering setup errors and breathing motions. Acta Oncol 2017;56(9):1189–96.
[2] Park PC, et al. Fast range-corrected proton dose approximation method using prior dose distribution. Phys Med Biol 2012;57(11):3555–69.
[3] Mohan R. A review of proton therapy – Current status and future directions. Precis Radiat Oncol 2022;6(2):164–76.
[4] 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.
[5] Park PC, et al. A beam-specific planning target volume (PTV) design for proton therapy to account for setup and range uncertainties. Int J Radiat Oncol Biol Phys 2012;82(2):e329–36.
[6] Knopf AC, et al. Adequate margin definition for scanned particle therapy in the incidence of intrafractional motion. Phys Med Biol 2013;58(17):6079–94.
[7] Unkelbach J, et al. Robust radiotherapy planning. Phys Med Biol 2018;63(22):22TR02.
[8] Sterpin E, et al. Robustness evaluation of pencil beam scanning proton therapy treatment planning: A systematic review. Radiother Oncol 2024;197:110365.
[9] Kaplan LP, et al. Plan quality assessment in clinical practice: results of the 2020 ESTRO survey on plan complexity and robustness. Radiother Oncol 2022;173:254–61.
[10] Trofimov A, et al. Visualization of a variety of possible dosimetric outcomes in radiation therapy using dose-volume histogram bands. Pract Radiat Oncol 2012;2(3):164–71.
[11] Korevaar EW, et al. Practical robustness evaluation in radiotherapy – A photon and proton-proof alternative to PTV-based plan evaluation. Radiother Oncol 2019;141:267–74.
[12] Rojo-Santiago J, et al. Robustness analysis of CTV and OAR dose in clinical PBS-PT of neuro-oncological tumors: prescription-dose calibration and inter-patient variation with the Dutch proton robustness evaluation protocol. Phys Med Biol 2023;68(17).
[13] de Jong JI, et al. Probabilistic consistency of voxel-wise robustness evaluation in IMPT for neuro-oncological patients. Radiother Oncol 2024;194:S4587–90.
[14] Sterpin E, et al. Development of robustness evaluation strategies for enabling statistically consistent reporting. Phys Med Biol 2021;66(4):045002.
[15] Tilly D, Ahnesjo A. Fast dose algorithm for generation of dose coverage probability for robustness analysis of fractionated radiotherapy. Phys Med Biol 2015;60(14):5439–54.
[16] Wahl N, et al. Analytical probabilistic modeling of dose-volume histograms. Med Phys 2020;47(10):5260–73.
[17] Yang Z, et al. Statistical evaluation of worst-case robust optimization intensity-modulated proton therapy plans using an exhaustive sampling approach. Radiat Oncol 2019;14(1):129.
[18] Souris K, Lee JA, Sterpin E. Fast multipurpose Monte Carlo simulation for proton therapy using multi- and many-core CPU architectures. Med Phys 2016;43(4):1700.
[19] Liu F, et al. Accelerating radiation therapy dose calculation with Nvidia GPUs. Ieee International Parallel and Distributed Processing Symposium Workshops (Ipdpsw) 2021;2021:449–58.
[20] Perko Z, et al. Fast and accurate sensitivity analysis of IMPT treatment plans using Polynomial Chaos Expansion. Phys Med Biol 2016;61(12):4646–64.
[21] Rojo-Santiago J, et al. PTV-based VMAT vs. robust IMPT for head-and-neck cancer: A probabilistic uncertainty analysis of clinical plan evaluation with the Dutch model-based selection. Radiother Oncol 2023;186:109729.
[22] Rojo-Santiago J, et al. Accurate assessment of a Dutch practical robustness evaluation protocol in clinical PT with pencil beam scanning for neurological tumors. Radiother Oncol 2021;163:121–7.
[23] Rojo-Santiago J, et al. A probabilistic evaluation of the Dutch robustness and model-based selection protocols for Head-and-Neck IMPT: A multi-institutional study. Radiother Oncol 2024;199:110441.
[24] van der Voort S, et al. Robustness recipes for minimax robust optimization in intensity modulated proton therapy for oropharyngeal cancer patients. Int J Radiat Oncol Biol Phys 2016;95(1):163–70.
[25] van Herk M, et al. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. Int J Radiat Oncol Biol Phys 2000;47(4):1121–35.
[26] Wohlfahrt P, et al. Dual-energy computed tomography to assess intra- and inter-patient tissue variability for proton treatment planning of patients with brain tumor. Int J Radiat Oncol Biol Phys 2019;105(3):504–13.
[27] Lambrecht M, et al. Radiation dose constraints for organs at risk in neuro-oncology; the European Particle Therapy Network consensus. Radiotherapy and Oncology 2018;128(1):26–36.
[28] Unkelbach J, Paganetti H. Robust proton treatment planning: physical and biological optimization. Sem Radiat Oncol 2018;28(2):88–96.