Automation and individualization of radiotherapy treatment planning in head and neck cancer patients
Kierkels, Roel Godefridus Josefina

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

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

Publication date:
2019

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 13-12-2019
CHAPTER 8

Summarizing discussion and future perspectives
8.1 Introduction

In this thesis, several approaches to improve treatment planning efficiency and quality in terms of automation and individualization for head and neck cancer (HNC) radiotherapy were explored. Semi-automated planning by means of multicriteria optimization (MCO) with navigation for intensity modulated radiotherapy (IMRT) has been investigated. MCO substantially reduced treatment planning time as compared to traditional dosimetrist-optimized IMRT planning. Moreover, biologically related objective functions based on multivariable normal tissue complication probability (NTCP) models have been introduced and evaluated in a fully automated Pareto-based planning framework. The NTCP-based objectives have the advantage of fewer unknown optimization parameters and an intrinsic mechanism of individualization (i.e. accounting for clinically relevant parameters instead of dose only).

NTCP models are also increasingly used to select patients for proton therapy, requiring a comparison between photon and proton treatment plans. To improve efficiency, a robust dose mimicking and a robust dose reduction algorithm have been introduced to automatically create robust intensity modulated proton therapy (IMPT) plans. Currently, robust IMPT plans are compared against ‘traditional’ planning target volume (PTV) based optimized photon plans using IMRT or volumetric modulated arc therapy (VMAT). In addition, the potential of robust VMAT optimization has been investigated and assessed in detail using a dosimetric evaluation of daily acquired cone-beam CTs (CBCTs) with deformable dose mapping and accumulation. Furthermore, the CBCT-based dose calculation has been evaluated against a dose calculation on weekly-acquired evaluation CTs. To assess the unknown errors related to deformable image registration (DIR) several automated evaluation metrics have been implemented and evaluated with radiotherapy planning CT and evaluation CT images.

Recently, the Dutch National Platform for Radiotherapy of Head and Neck Tumors (LPRHHT) initiated a plan comparison study among 15 radiotherapy institutes in the Netherlands aiming at improved consensus of HNC radiotherapy planning. The CT images and target and OARs delineations of a HNC case were distributed by the VU University Medical Center Amsterdam and planned by each institute according to their institutional protocol. The treatment planning optimization strategies described in this thesis, were also applied to this case (the representative case in this thesis equals the distributed HNC case from LPRHHT study). The NTCP estimates of xerostomia, swallowing dysfunction, and tube feeding dependence derived from the dose distributions of the plans of the 15 radiotherapy institutes are shown in figure 8.1 (grey lines and circles). Overall, the NTCP of swallowing dysfunction and tube feeding dependence improved for the plans derived from the optimization strategies as described in this thesis (blue – yellow markers, figure 8.1). Relevant dosimetric parameters
and NTCP estimates from the corresponding dose distributions are given in table 8.1. It should be noted that all photon and proton plans achieved a clinical target volume (CTV) coverage of $V_{95}\geq 99.9\%$. In addition, the planning target volume (PTV) fulfilled the criterion $D_{98}\geq 95.0\%$ for all photon plans.

**Figure 8.1** Normal tissue complication probability (NTCP) values for xerostomia (left panel), grade 2-4 swallowing problems (middle panel), and tube feeding dependence (right panel).

<table>
<thead>
<tr>
<th>Target dose (Gy)</th>
<th>VMAT*</th>
<th>MCO VMAT</th>
<th>NTCP-driven VMAT</th>
<th>Robust VMAT</th>
<th>IMPT*</th>
<th>DMR IMPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{98}$ CTV$_{primary}$</td>
<td>69.7</td>
<td>68.0</td>
<td>68.7</td>
<td>69.0</td>
<td>69.1</td>
<td>69.2</td>
</tr>
<tr>
<td>$D_{05}$ CTV$_{primary}$</td>
<td>70.6</td>
<td>70.3</td>
<td>70.7</td>
<td>70.4</td>
<td>70.1</td>
<td>71.0</td>
</tr>
<tr>
<td>$D_{03}$ CTV$_{primary}$</td>
<td>71.5</td>
<td>71.2</td>
<td>73.4</td>
<td>73.0</td>
<td>72.4</td>
<td>73.4</td>
</tr>
<tr>
<td>$D_{98}$ CTV$_{elective}$</td>
<td>53.3</td>
<td>53.2</td>
<td>52.1</td>
<td>52.8</td>
<td>53.4</td>
<td>54.6</td>
</tr>
<tr>
<td>$D_{05}$ CTV$_{elective}$</td>
<td>69.9</td>
<td>68.6</td>
<td>69.0</td>
<td>69.0</td>
<td>69.4</td>
<td>69.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>OAR mean dose (Gy)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Contra Parotid</td>
<td>12.7</td>
<td>12.2</td>
<td>13.0</td>
<td>11.9</td>
<td>7.3</td>
<td>8.5</td>
</tr>
<tr>
<td>Ipsi Parotid</td>
<td>28.4</td>
<td>26.1</td>
<td>32.0</td>
<td>29.4</td>
<td>27.5</td>
<td>26.8</td>
</tr>
<tr>
<td>PCM superior</td>
<td>64.5</td>
<td>63.5</td>
<td>61.3</td>
<td>65.0</td>
<td>64.6</td>
<td>65.6</td>
</tr>
<tr>
<td>PCM medial</td>
<td>48.9</td>
<td>44.9</td>
<td>46.7</td>
<td>47.5</td>
<td>44.2</td>
<td>38.1</td>
</tr>
<tr>
<td>PCM inferior</td>
<td>37.5</td>
<td>27.5</td>
<td>26.8</td>
<td>33.7</td>
<td>28.8</td>
<td>24.4</td>
</tr>
<tr>
<td>Oral Cavity</td>
<td>40.7</td>
<td>31.7</td>
<td>33.5</td>
<td>33.2</td>
<td>26.0</td>
<td>25.3</td>
</tr>
<tr>
<td>Cricopharyngeal muscle</td>
<td>39.3</td>
<td>29.8</td>
<td>29.4</td>
<td>31.6</td>
<td>33.2</td>
<td>27.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NTCPs* (%)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Xerostomia</td>
<td>30.0</td>
<td>29.5</td>
<td>30.3</td>
<td>29.1</td>
<td>24.5</td>
<td>25.6</td>
</tr>
<tr>
<td>Swallowing problems</td>
<td>31.5</td>
<td>26.5</td>
<td>26.3</td>
<td>28.0</td>
<td>24.4</td>
<td>24.6</td>
</tr>
<tr>
<td>Tube feeding dependence</td>
<td>29.4</td>
<td>24.6</td>
<td>23.5</td>
<td>27.2</td>
<td>24.0</td>
<td>23.1</td>
</tr>
<tr>
<td>$\Sigma$NTCP</td>
<td>90.9</td>
<td>80.6</td>
<td>80.1</td>
<td>84.3</td>
<td>72.9</td>
<td>73.3</td>
</tr>
</tbody>
</table>

**Abbreviations:** CTV = clinical target volume; PCM = pharyngeal constrictor muscle; $D_x$ indicates the dose in the fractional volume $x$. *Dosimetrist-optimized plans from chapter 1 in italic. **NTCP models as in chapter 1. The OAR in bold are also considered in the NTCP models.
The following paragraphs provide a summarizing discussion of the presented work as well as future perspectives. Throughout these paragraphs, the treatment optimization strategies as applied to the HNC case from the LPRHHT study (same case as in chapter 1) are presented and discussed.

### 8.2 Automated treatment planning optimization in head and neck cancer

Multicriteria optimization is a semi-automated planning solution requiring manual Pareto surface navigation. The main advantage of MCO is that the Pareto optimal plans are obtained fully automatically. The resulting library plans are Pareto optimal in fluence space, however not necessarily clinically optimal. Navigation of the Pareto surface, with the objective to find a clinically acceptable plan, is therefore required. This was investigated in 20 HNC patients and described in chapter 2. In the presented study, the plan quality of the conventional IMRT plans and the MCO-plans were scored using an in-house-developed assessment form. Each plan was then independently scored by two radiation-oncologists specialized in HNC. The plan quality score indicated no difference between the conventional IMRT and the MCO-plans and 87% of all ratings were within one-point difference between the radiation-oncologists. In 1% of all ratings there was less consensus, which was mainly caused by different scoring of target coverage close to an organ at risk such as the parotid glands. Moreover, MCO-plans showed slightly lower parotid gland doses at the cost of target coverage, indicating that clear instructions to Pareto navigation are required. We further found that navigation of a Pareto front requires less planning time than conventional IMRT planning (approximately 40 vs. 200 minutes, respectively) but remains challenging due to the relative large number of trade-off objectives.

For illustration purpose a VMAT MCO plan is shown in figure 8.2a. The treatment plan showed lower mean dose to the OARs related to swallowing than the dosimetrist-optimized plan (chapter 1, italic numbers table 8.1). This resulted in a drop of 10% in ΣNTCP. To simplify the navigation process in this case, two composite volumes were constructed: for the salivary-related OARs and the swallowing-related OARs. Navigation of the Pareto surface was then performed in approximately 10 minutes. This is shorter than reported in chapter 2, in which the navigation process took approximately 20 minutes using RayStation v2.4. We believe that the application of composite structures is an intermediate approach to a more biologically meaningful navigation, including OAR specific weighting. Therefore, we are developing new methods to combine single criteria into NTCP-based navigation tools such that the overall complication risk can directly be minimized. A first step to directly optimize on NTCP was introduced in chapter 3 and further evaluated in chapter 4.
Other researchers have improved the MCO algorithms and assessed its performance on different treatment sites. These improvements include MCO optimization to aim for a plan close to the segmented deliverable plan for IMRT and VMAT [1,2]. In addition, MCO considering plan robustness (for intensity modulated photon and proton planning) was investigated and implemented [3]. Treatment sites that have been studied with MCO include prostate, brain and stereotactic lung radiotherapy [4–6]. The general conclusion of these studies and our study is that treatment planning times can be considerably reduced with similar or improved plan quality compared to the conventionally optimized plans.

**Figure 8.2** The dose distribution of a VMAT plan optimized with multicriteria optimization (A) and a dose distribution optimized with objective functions based on multivariable NTCP models (B).

The low to high dose is indicated by the blue to red colorwash. The grey and white contours indicate the elective and primary PTVs, respectively. The oral cavity is indicated by the dark green contour and the parotids are indicated by the flashy green and purple contours. The swallowing muscles are indicated by the pink to red contours.

Due to improvements in both computational performance and radiotherapy planning optimization algorithms it is expected that the role of MCO will change such that fully automated optimization algorithms will steer the optimizer towards the clinically interesting region on the Pareto surface. MCO will then be used by dosimetrists or radiation oncologists to select the preferred plan per patient.

### 8.3 Plan optimization using multivariable NTCP-based objective functions

The objective list of inverse treatment planning in HNC patients can be rather long. Finding the optimal thresholds and weights of the objectives that result in the lowest estimated NTCPs (without compromising target coverage) can therefore be challenging and time-consuming. The purpose of our study (chapter 3) was to introduce objective functions based on multivariable NTCP models, combining multiple dosimetric parameters into a single objective, simplifying the optimization process while aiming for the lowest NTCP.
We demonstrated that objective functions based on multivariable NTCP models produce clinically acceptable treatment plans with slightly lower NTCP estimates compared to conventional optimized plans. We anticipate that the NTCP-based objectives were combined with physical objectives to account for both adequate target coverage and the volumes that were not described by the NTCP models. Figure 8.2b shows the dose distribution of a VMAT plan that was optimized with the hybrid objective function using log-transformed multivariable NTCP models for xerostomia, grade 2-4 swallowing problems (including the mean dose of the superior pharyngeal constrictor muscle (PCM) and the mean dose of the oral cavity), and tube feeding dependence [7]. Similar to the MCO plan, the NTCP-driven plan shows slightly lower OAR doses and NTCP values than the dosimetrist-optimized VMAT plan (table 8.1). Since the ipsilateral parotid gland and medial PCM were not included in any of the NTCP models, the mean dose to these structures were not fully minimized.

Biologically driven objective functions are not new. TCP, NTCP, and gEUD-based objectives have been proposed by others [8–11]. To find the global minimum of the search space (i.e. the optimal plan) a convex objective function is recommended (e.g. [12]). Therefore, typical objective functions optimize the weighted mean-squared difference of multiple criteria. These objective functions are convex in fluence space but require a threshold value. Hoffmann et al. investigated the convexity of the biologically driven objectives and used a normal tissue threshold value [13]. However, this threshold is not always known prior to optimization. Moreover, NTCP models are generally sigmoidal shaped and only convex up to the point of inflection (see figure 1.2, NTCP curve of tube feeding dependence). To maintain convexity during optimization the proposed NTCP-based objectives were log-transformed. With the direct use of the log-transformed NTCP objective (i.e. no mean-squared difference) at each iteration, the optimizer is rewarded with a lower NTCP value, always searching for a lower OAR doses.

We acknowledge that dose volume histogram points (such as V60, the volume receiving 60 Gy) incorporated in the NTCP-based objectives cause nonconvex functions. Therefore, the preferred NTCP-based objectives include mean dose parameters. In terms of deliverable plans, the dose distribution is also a nonconvex function of the MLC leaf positions. Therefore, there is no guarantee that the objective value converges to the global optimum in finite time. Furthermore, to increase the speed of the optimization process, in future work, step size calculation per iteration will be included by adding the Hessian function to the algorithm.

Traditional NTCP models (such as the Lyman-Kutcher-Burman NTCP model) contain one dose parameter only. It is expected that the application of an NTCP-based objective with only one dose-volume parameter is limited as compared to conventional physical objectives. However, NTCP models increasingly contain multiple dose parameters and clinical factors,
such as the NTCP model to predict tube feeding dependence 6 months after therapy [14]. This NTCP model comprises four dose parameters and therefore translates four conventional dose-volume or gEUD objectives into one. We evaluated plan optimization using the NTCP model for tube feeding dependence within hundred HNC patients and demonstrated improved plan quality for a few patients only (chapter 4). One explanation is that within patients in which the risk of tube feeding dependence was already low the dose could not be further minimized (given that approximately 5 - 10% of patients are tube feeding dependent 6 months after therapy). Another explanation is inherent to the relationship between xerostomia, dysphagia, and tube feeding dependence models. The NTCP models for these clinical end-points included similar OAR dose parameters (e.g. the mean dose to the contralateral parotid gland and pharyngeal constrictor muscles). Therefore, the dose distribution to these volumes were not only optimized by the tube feeding dependence NTCP-based part of the objective function, but also by objective functions related to xerostomia and dysphagia. Furthermore, the reference plans were already optimized on xerostomia and dysphagia, albeit with dose-volume objectives.

The current and future research agenda contains projects to improve NTCP models considering multiple OARs. Recent work (unpublished data) has identified both parotid glands as well as both submandibular glands to play a significant role in a comprehensive NTCP model for patient-rated xerostomia. It is expected that the use of these NTCP models in optimization will increase plan quality. Moreover, promising results of radiomics (i.e. imaging biomarkers) incorporated into NTCP models have been presented recently [15]. If the radiomics-based NTCP models will be used within the objective functions, the image features will play a role in the optimization of the dose distribution. However, the additional predictive performance of CT-based image features of the parotid glands in the existing NTCP models for xerostomia was limited [15]. MRI features seem to be more promising for this purpose [16].

8.4 Automated robust proton planning optimization

The availability of proton therapy is currently limited and adequate patient selection is therefore needed. In the Netherlands, the so-called “model-based approach” [17] has been approved by the National Healthcare Institute (Zorginstituut Nederland) to select patients for proton therapy using NTCP models. Since HNC potentially benefits from proton therapy in terms of toxicity reductions, the research presented in chapter 5 focused on HNC radiotherapy in the framework of the model-based approach.
To increase treatment planning efficiency, we developed a dose mimicking and dose reduction (DMR) algorithm to automatically create a robust proton plan given a reference photon dose (chapter 5). The algorithm was developed on the assumptions that the photon dose distribution and contours are available, that the target DVH(s) remains similar between both treatment modalities (dose mimicking optimization) and that the dose to OARs can be further reduced (dose reduction optimization) due to the physical properties of the proton beam. Our results demonstrated that the dose distributions of the DMR plans of 95% (38/40) of the cases was very similar to the “dosimetrist-optimized” plans. For two cases, the discrepancy was introduced by an underdosage of the photon plan, e.g. for targets near the skin, which consequently lead to IMPT plans with an underdosage as well, albeit on a different location, e.g. in an OAR. This can be explained by the fact that DVH-based mimicking was used (instead of voxel-based mimicking of the dose). To account for this underdosage, in a future step, a clinical goal-based normalization method for the underdosed photon plans can be added prior to the dose mimicking and reduce optimization.

To be able to efficiently process all requests for a model-based plan comparison from the UMCG and other referring institutes, the DMR algorithm will be implemented in the near future. Moreover, to create a clinically realistic treatment plan, beam angle selection and avoidance structures (e.g. to avoid metallic implants) are required. These settings are not incorporated in the presented DMR algorithm and need further investigation to make the workflow fully automated. To further reduce the departmental workload, the performance of the DMR algorithm on other treatment sites will be the subject of ongoing research. The DMR proton dose distribution of the representative case is shown in figure 8.3a. Although similar target coverage was reached, the mean dose to the OARs varied a few Gy between the DMR plan and the dosimetrist-optimized IMPT plan (table 8.1). This led to a ΣNTCP difference of 0.4% for the presented case.

The DMR algorithm has great potential for multiple applications. The DMR algorithm can be used to derive a deliverable IMPT plan from a predicted dose distribution derived from e.g. a machine learning (ML) based auto-planning approach [18]. Automated photon and proton treatment planning using ML techniques is currently part of ongoing research. Our preliminary results on ML-based automated photon planning shows promising results and creates a dual arc VMAT plan for a HNC patient in less than 10 minutes. To turn a predicted proton dose distribution into a robust IMPT plan, however, robust DMR optimization is required (chapter 5).

For adaptive radiotherapy, a modified DMR algorithm can be used to mimic the actual photon or proton plan to the new geometry of the same patient. Therefore, the algorithm should account for a different frame of reference (i.e. reference CT vs. evaluation CT).
Furthermore, the speed of the algorithm should be improved, especially if the algorithm will be applied for online adaptive protocols.

![Figure 8.3 The dose distribution of the IMPT plan derived from the automated DMR optimization (A) and the dose distribution of the robustly optimized VMAT plan (B) (same patient as in figure 8.1). Contour colors and colorwash as in figure 8.2.](image)

### 8.5 Assessment of deformable image registration errors

In an adaptive workflow, accurate deformable image registration is required to find the transformation from the evaluation CT (or cone-beam CT) scan to the reference CT scan (i.e. the planning CT). The transformation can then be used to automatically propagate the contours and map the accumulated dose to the evaluation CT. Hence, the accumulated dose can be used as a reference background dose to calculate an adapted plan that can compensate for underdosage of the tumor or overdosage of the healthy tissues in the accumulated dose. A prerequisite of the clinical introduction of such a workflow is a clear understanding of the uncertainties related to DIR.

DIR is a key element in adaptive radiotherapy. However, the determination of the transformation between two image sets is an ill-posed optimization problem. As a consequence, the transformation derived from a DIR comprises unknown errors. When used for contour propagation, the quality of the contours can be assessed by visual inspection, whereas this is not possible for the mapped dose distribution. In literature, a number of publications report on tools to quantify the DIR error [19,20]. Frequently reported tools use the target registration error (based on anatomical landmarks) or contour comparison metrics to assess the DIR error [21]. These tools are, however, time consuming, subject to user variability and mainly suitable for the commissioning and evaluation of DIR algorithms. More details are recently reported by the AAPM Task Group 132 [21].
In order to use DIR for dose mapping, the errors related to DIR should be identified for each deformable registration separately. Therefore, we reviewed the current literature and combined the most promising uncertainty metrics for DIR in a fully automatic, quantitative, and a case-specific evaluation method of DIR (chapter 6). Metrics related to the physiological feasibility and the numerical robustness were investigated within a deformable registration platform with a B-spline transformation model. The results of the different uncertainty metrics were compared against a ground-truth transformation derived from a synthetic CT image with a known deformation with respect to the reference CT.

We found DIR errors up to 10 mm in the studied head and neck cases (n = 26). Approximately 20% of the voxels showed an error > 2.0 mm. Since DIR errors in the volumes with high dose gradients will result in the largest errors in accumulated dose, information about the errors in these volumes is crucial. In the head and neck cases, the largest errors were found in the vertebrae and oral cavity. Overall, the log-transformed distance discordance metric (DDM) and the log-transformed harmonic energy (HE) showed the highest correlation with the ground-truth DIR error. To identify DIR errors with a negative predictive value of 0.90 at a ground-truth tolerance level of 2.0 mm (i.e. the ratio of voxels correctly identified with a DIR error ≤2.0 mm), we proposed tolerance thresholds for the DDM and HE of 0.49 and 0.014, respectively. Similar to the work of others, it was found that the DDM outperformed the inverse consistency error and transitivity error metric [20,22,23]. Vickress et al. also demonstrated that the DDM was most predictive with a magnitude of dose uncertainty of 2.5 Gy (which was defined as the range of dose uncertainties at an anatomic landmark)[23]. Until now, the correlation between the DIR error and dose using the DDM was only investigated in lung cancer patients.

We observed that the correlation between the DIR error metrics and the ground-truth DIR error remains rather low (r < 0.6). This was similar to the finding of Saleh et al. and further improvements of DIR error metrics is needed [22]. It should however be mentioned that DIR algorithms have substantially improved in the last decade, in terms of speed and quality, e.g. by adding contour-based registration. In future work, the uncertainty metrics (i.e. the DDM and HE) should be used to assess the quality of other DIR algorithms, e.g. contour-driven DIR algorithms.

Currently, DIR and automated planning are emerging as important plan adaptation tools in radiotherapy. Two other prerequisites for an online adaptive workflow are fast approval of contours and the adapted dose distribution as well as fast quality assurance of the adapted treatment plan. Two potential solutions for the latter include independent Monte Carlo based dose calculations and the use of log-file based dose calculations in which log-files of the actual machine settings are used to reconstruct the dose distribution (see e.g. [24]). Both are also part of ongoing research.
8.6 Robust plan optimization

On one side of the spectrum there is online adaptive radiotherapy, which has the potential to deliver the ‘optimal’ plan to a given geometry. At time of writing, however, adaptive radiotherapy, and especially online adaptive approaches are too resource intensive and clinical implementation remains limited. On the other side of the spectrum there is conventional radiotherapy that provides a plan that is fairly insensitive to errors like patient setup errors and anatomical variations. The traditional approach to account for random and systematic positioning errors is the use of a CTV to PTV margin. This margin, however assumes that the dose distribution is invariant to setup uncertainties. Setup errors may, however, affect the shape of the dose distribution, especially near high-density gradients such as bone and air cavities. This effect may be small for photon therapy but more profound for proton therapy, due to the sensitivity of the individual Bragg peaks to positioning errors. Therefore, robust optimization has been proposed as an alternative approach to conventional margin-based planning.

Probabilistic plan optimization and minimax optimization are two well-described robust optimization strategies. The former has been implemented and extensively evaluated in 20 HNC IMRT by Fontanaroza et al. [25]. They demonstrated that probabilistic optimization led to OAR dose reductions of up to more than 6 Gy, whereas a similar probability of effective CTV coverage was achieved as compared to PTV-based plans. The potential of worst-case minimax robust plan optimization for VMAT in HNC has however not been fully explored. Therefore, we extensively investigated minimax robust optimization in detail and evaluated the robustness of the plans with CBCT-based dose calculations to account for random setup errors and geometrical variations. The resulting dose distributions were mapped to the reference planning CT and accumulated. Similarly, the robustness of the PTV-based plans was evaluated and compared against the accumulated dose of the robustly optimized plan. Similar to the work of Fontanaroza et al. [25] the target objectives of the PTV-plan were adapted to robust objective functions. Other settings were not changed and differences in OAR dose were primarily due to the robust target objective functions.

We demonstrated that robust VMAT optimization produces clinically acceptable treatment plans with improved target coverage and lower OAR dose (on average up to approximately 3 Gy) as compared to PTV-based plans. This translated into approximately 1 - 3% lower NTCP values for xerostomia, grade 2-4 dysphagia, and tube feeding dependence. This is also demonstrated for the representative case of figure 8.3b (see also table 8.1, column 4).

Since adaptive radiotherapy has become increasingly available, the effect of a mid-treatment course plan adaptation was investigated for PTV-plans and robustly optimized
Summarizing discussion and future perspectives

VMAT plans. Therefore, the nominal plan was adapted in patients that had an increase of >2.5% in NTCP after three weeks of treatment (arbitrarily chosen threshold). This was the case for 4/10 patients in our study. A mid-treatment course plan adaptation combined with robust optimization further decreased the dose with approximately 1 Gy in the OARs, whereas target coverage was assured. This translated in a further decreased of the NTCPs of approximately 1%. This improvement can be achieved, albeit for a limited group of patients. Adequate patient selection for adaptive radiotherapy is therefore needed (see e.g. [26]).

The studied plans were optimized to account for 5.0 mm setup uncertainty, which was the standard clinical value at time of research. Patient setup errors remain below this value and no CTV underdosage was observed. The number of cases that need re-planning will however increase with smaller uncertainty values (e.g. 3.0 mm) if the plan is optimized on the PTV. Robust optimization can, however, contribute to smaller robustness settings and/or fewer plan adaptations due to a higher inherent robustness.

Recently, Liu et al. investigated the probabilistic optimization approach in an adaptive setting for IMRT in HNC patients. Data from cone-beam CT images acquired during the first five fractions were fed back into the optimizer to further optimize setup uncertainties and geometrical variations in the adapted plan. Similar to van Kranen et al. they found a dose reduction of 1 Gy/mm margin reduction [27,28]. Using a margin-less 4D adaptive protocol they also estimated a potential NTCP reduction of ≥5 % or even ≥10% for a substantial portion of the studied patients. The reduction of ΔNTCP depended of course on the slope of the NTCP curve.

Robust treatment planning is characterized by PTV-less planning. A challenging task to the clinical introduction of robust optimization remains the clinical evaluation of the plans (i.e. the absence of the PTV). The traditional method relies on the dose coverage of the PTV (e.g. V95%≥98%, the relative volume receiving 95% of the prescribed dose should be ≥ 98%). Other evaluation methods to assess the plan quality in terms of robustness are therefore needed. In literature, different methods have been proposed, including the analysis of DVH-bands [29], error bar dose distributions [30], the worst-case dose distribution and the voxel-wise minimum dose distributions. All metrics require an ensemble of dose distributions derived from dose calculations with different setup error scenarios. This is currently very time consuming and fast robustness evaluation metrics are needed and part of ongoing research. Perkó and colleagues proposed a promising method based on the Polynomial Chaos Expansion principle in which a meta-model with multidimensional polynomials is used to approximate the voxel dose very fast [31]. Verification with other metrics is however needed.
8.7 Future perspectives

Major technological advancements in radiotherapy treatment planning strive for fully personalized treatments and comprise inverse planning optimization and automation amongst other tools to achieve this. In combination with improvements in computational performance, this opens doors to automatically simulate scenarios of new treatment schemes. Future treatment methods can then be more intensively simulated upfront to guarantee safe introduction of these new treatments. Moreover, to easily assess the quality of new treatment techniques, not only in HNC radiotherapy, an automated framework for batch planning analysis should be developed. This, in turn, will lead to efficient commissioning and implementation of new class solutions and emerging radiotherapy planning tools. Furthermore, biologically driven indices as well as volumetric and functional imaging during the course of treatment will play an increasingly important role to perform fractionated analysis within such a simulation framework. Future NTCP models together with TCP models and models for secondary cancer induction should be incorporated into such a framework.
8.8 References


