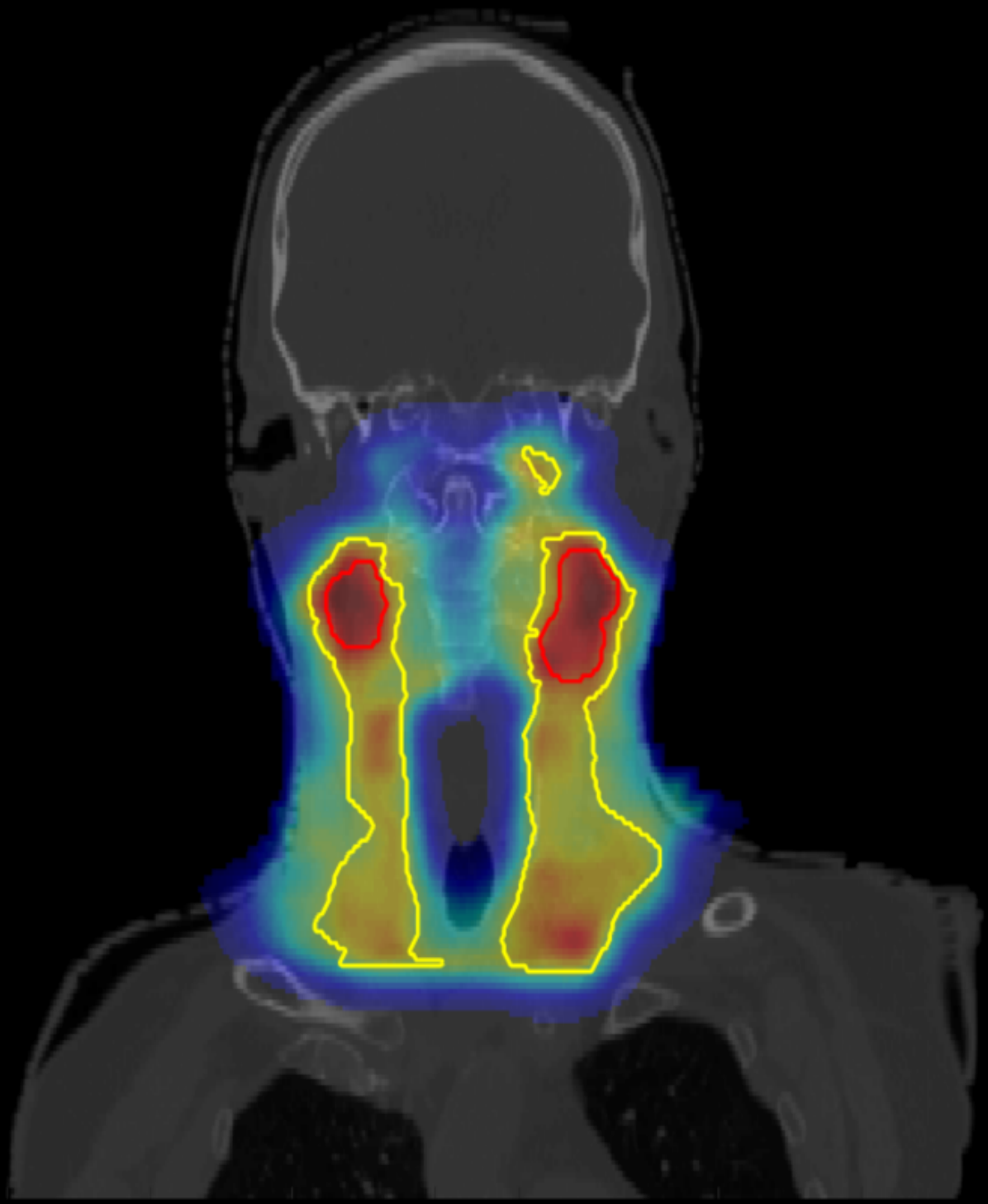
Robust optimization in intensity modulated proton therapy treatment planning



Robust optimization in intensity modulated proton therapy treatment planning

by

Manon van Zon

in partial fulfilment of the requirements for the degrees of

Master of Science
in Applied Mathematics

and

Master of Science
in Biomedical Engineering

at the Delft University of Technology
to be defended publicly on July 14, 2023 at 10:00

Thesis committee Applied Mathematics:

Prof. dr. ir. Martin van Gijzen	TU Delft
Dr. ir. Danny Lathouwers	TU Delft
Dr. Johan Dubbeldam	TU Delft

Thesis committee Biomedical Engineering:

Dr. ir. Danny Lathouwers	TU Delft
Prof. dr. ir. Martin van Gijzen	TU Delft
Dr. ir. Sebastiaan Breedveld	Erasmus MC
Ir. Michelle Oud	Erasmus MC
Dr. Johan Dubbeldam	TU Delft

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

Abstract

Intensity modulated proton therapy is an advanced radiotherapy technique that is used to treat cancer patients. In order to successfully treat a patient, sufficient dose to the tumor is required. However, during the fractionated treatment, multiple errors can cause a difference between the planned and actual dose delivery. To ensure adequate dose delivery in potential error scenarios, robust treatment plans are acquired as these are less sensitive to uncertainties inherent to proton therapy. However, robust optimization is challenging.

First, as robust optimization accounts for multiple error scenarios, the time needed to generate optimal treatment plans increases significantly. Therefore, it is investigated in this thesis if the optimization time can be reduced while preserving treatment plan quality. This is investigated through two different methods. In the first method, the number of error scenarios accounted for during optimization is reduced. We found that this can significantly reduce optimization time, while improved target coverage and lower risk on side effects are obtained. However, the near-maximum dose to the tumour was found to be less favourable. The second method investigated is variance optimization. This method significantly reduces optimization time. However, for similar target coverage, the risk of side effects increases.

Another challenge related to robust optimization is the increase in delivered dose to healthy tissues surrounding the tumor, which increases the risk of side effects. Therefore, it is investigated in this thesis if the risk of side effects can be lowered by allowing higher maximum dose to the tumor. It is found that this method indeed reduces the risk of side effects. However, the increased maximum dose to the tumor may not be clinically desired as the increase may lead to higher risks of other side effects: edema and fibrosis. The clinically desired trade-off between near-maximum dose and normal tissue sparing should be established.

Preface

This thesis marks the end of my combined master in Applied Mathematics and Biomedical Engineering at Delft University of Technology. My thesis, worth 42 ECTS, separately highlights the mathematical and biomedical aspects of the research but also shows the connection between the two fields. First, Section 2 provides a general introduction into the thesis topic, which combines mathematical and biomedical aspects. Sections 3 and 4 focus on the biomedical aspects. Whereas Section 5 is centered around the mathematical aspects of this thesis.

Acknowledgements

I would like to thank my supervisors, Michelle Oud, Sebastiaan Breedveld, Martin van Gijzen and Danny Lathouwers for their time, guidance and invaluable input. I also would like to thank Johan Dubbeldam for his time and willingness to be a committee member. Furthermore, I would like to thank my fellow master students at Erasmus MC, I really enjoyed working together with them. Last but not least, I want to thank my family and friends for their unwavering support.

Manon van Zon,
Delft, June 2023

Contents

List of Abbreviations

In this section, all abbreviations are listed including the page number that indicates where the corresponding definition is given.

Abbreviation	**Meaning**	**Page**
CT	computed tomography	4
CTV	clinical target volume	4
CTV_{5425}	part of CTV that is prescribed 54.25 Gy	12
CTV_{7000}	part of CTV that is prescribed 70.00 Gy	12
D2%	the maximum dose to 2% of the voxels	12
D98%	the minimum dose to 98% of the voxels	9
IMPT	intensity modulated proton therapy	4
MCO	multi-criterial optimization	7
NTCP	normal tissue complication probability	12
OAR	organ at risk	4
PCE	polynomial chaos expansion	10
pCT	planning CT	4
PDF	probability density function	10
PTV	planning target volume	6
rCT	repeat CT	5

1 Introduction

In 2020, over 900.000 people worldwide were diagnosed with head and neck cancer and over 450.000 people died from this type of cancer [1]. Head and neck cancer includes all types of cancer in the head and neck region excluding the brain, eyes, thyroid and spinal cord.

One of the advanced modalities to treat cancer is intensity modulated proton therapy (IMPT), where protons are used to eradicate the tumor. As treatment is delivered in multiple fractions, divided over a number of weeks, it is key to generate a treatment plan that ensures sufficient dose to the tumor in each of these fractions. Many things such as tumor shrinkage, as result of successful treatment, and weight loss of the patient can impact the dose delivery in the different treatment fractions. This, together with other uncertainties inherent to IMPT, makes it challenging to generate an optimal treatment plan. To deal with these challenges, robust treatment plans are generated as these are less sensitive to uncertainties. However, with robust optimization it takes many hours to generate an optimal treatment plan. To improve the efficiency of robust treatment plan generation, it is crucial to minimize optimization times.

Another issue in IMPT treatment planning is that by irradiating the tumor, unavoidable doses are also delivered to normal tissues. Therefore, one of the challenges is to reduce the dose to normal tissues as damage to those can cause side effects. One of the main side effects for head and neck patients is xerostomia, also known as dry mouth syndrome. In the worst case scenario, patients with xerostomia have to drink water every 30 minutes to keep their mouth hydrated. This means both during the day and at night. The other main side effect is dysphagia, which means difficulty in swallowing. Patients with severe cases of dysphagia can not eat and have to rely on a feeding tube. As these side effects have a substantial impact on the patient's quality of life, it is key to optimize treatment plans such that the tumor is eradicated but normal tissues are spared.

Hence, the focus of this thesis is two-fold, which is summarized in the following research questions:

- Can we reduce the optimization time while maintaining treatment plan quality?
- Can we reduce the risk of side effects?

To answer these questions, the remainder of this thesis is structured as follows. In Section 2 the basic principles of proton therapy and treatment plan generation are covered. Then, in Section 3, we present a method that can be used to reduce the optimization time. By employing this method, we hypothesize that the risk of side effects can be reduced by increasing the maximum allowable dose to the tumor, which is further investigated in Section 4. In Section 5, we explore if variance optimization can be used to again reduce the optimization time. Finally, in Section 6 we formulate answers to the research questions.

2 Technical background

The aim of this section is to provide some practical background that is needed to understand this thesis. Therefore, in Section 2.1 the basic principles of conventional radiotherapy and proton therapy are explained. In Sections 2.2 and 2.3, we explain what radiotherapy treatment planning and fractionation is, respectively. Then, in Sections 2.4 and 2.5, we explain the uncertainties that are related to radiotherapy treatment planning and two methods that deal with these uncertainties. Once a treatment plan is generated, it is important to assess its quality. To this end, Section 2.6 presents two methods that are used to evaluate treatment plans. In Section 2.7, we provide medical information about head and neck patients as they are the focus group for this research. Finally, in Section 2.8 we elaborate on some criteria that are used for treatment plan assessment in this study.

2.1 Radiotherapy

Cancer is frequently treated using radiotherapy, where radiation is used to eradicate cancerous tissue [2]. Various types of radiotherapy are applied, which differ in the radiation dose as a function of the penetration depth. In photon therapy, which is most commonly used, patients are treated with X-rays. For X-rays, the peak radiation dose is deposited close to the skin, after which the dose deposition decreases exponentially. Proton therapy is another, rapidly growing, radiotherapy treatment modality [3]. In proton therapy, patients are treated with protons rather than X-rays. The energy of a proton determines how deep it can penetrate, also called its range. However, its penetration depth also depends on the density of the material it encounters. This behaviour is described by the stopping power. Hence, the depth of the peak radiation dose can be changed by altering the proton energy. For proton therapy, this peak is known as the Bragg peak. Moreover, the radiation dose decreases instantly after depositing the peak radiation dose. For both types of therapy, the relation between dose deposition and penetration depth is shown in Figure 2.1.1.

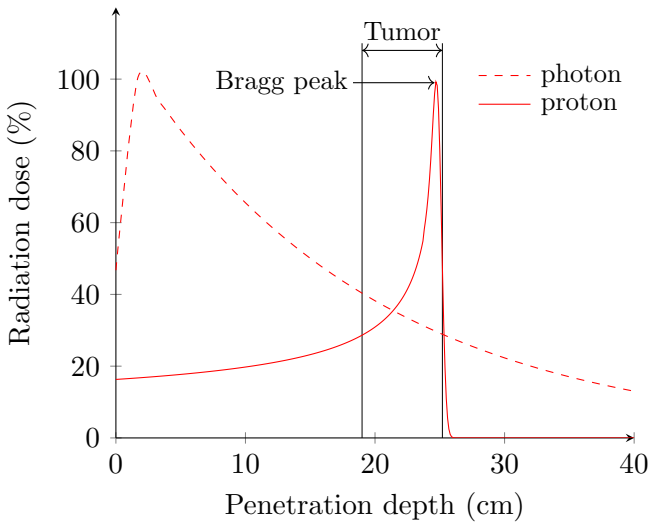


Figure 2.1.1: The relation between dose deposition and penetration depth of photons and protons [4].

Here, the proton energy is adjusted such that the Bragg peak coincides with the tumor. This figure illustrates the potential dosimetric advantages of proton therapy compared to photon therapy. Because of the differences in the dose-depth curves, the radiation dose deposited to healthy tissue surrounding the tumor can be reduced in proton therapy compared to photon therapy.

The type of proton therapy considered throughout this research is called intensity modulated

proton therapy (IMPT). Here, the proton beam is divided into smaller pencil beams, where each pencil beam can be steered with magnets and their energy and intensity can be tuned.

During radiotherapy treatment delivery, the patient is positioned on a couch inside the treatment room. A schematic overview of a treatment device is shown in Figure 2.1.2. The couch can be translated and rotated horizontally inside the treatment room. Connected to the couch is a machine referred to as the gantry, which can rotate around the patient. At the head of the gantry, called the snout, proton beams are released and directed towards the patient. As the snout can be positioned 360° around the patient, its position is known as the beam angle.

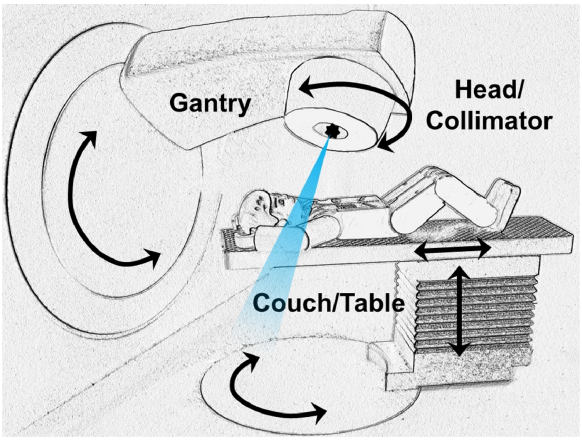


Figure 2.1.2: Schematic overview of a treatment device [5].

2.2 Treatment planning

In proton therapy treatment planning, a patient-specific optimal dose distribution is generated. To this end, numerous machine parameters have to be determined. These parameters include beam and couch angles as well as the energy, direction and intensity of the pencil beams. To determine these parameters, an entire workflow is developed [6]. First, a planning computed tomography (CT) scan of the patient is required. Using this CT scan, the proton stopping power is calculated as well as the dose that will be delivered to the patient by a pencil beam. Furthermore, on this planning CT (pCT) scan both the tumor and the surrounding healthy tissues in the vicinity of the tumor, known as organs at risk (OARs), are delineated. Delineation allows for fast and easy distinction of structures. The volume that contains the tumor as visible on the CT scan and parts where the microscopic parts of the tumor are expected to be, is the clinical target volume (CTV). The CTV is the volume that must be irradiated during therapy and thereby requires a minimum amount of dose to eradicate the cancerous tissue. However, it is crucial to limit OAR dose to minimize side effects. Therefore, structure-wise dose constraints and objectives are defined. Subsequently, a computer program is used to find an optimal dose distribution within given constraints. To compute the dose distribution inside a patient, the planning CT scan is discretized into voxels. The relation between the dose in a voxel due to a pencil beam is given by the dose deposition matrix $D \in \mathbb{R}^{n \times m}$ as follows:

$$\mathbf{d} = D\mathbf{x}. \tag{2.1}$$

Here, $\mathbf{d} \in \mathbb{R}^n$ represents the vector containing the dose in each voxel for all n voxels and $\mathbf{x} \in \mathbb{R}^m$ represents the pencil beam intensities for the m pencil beams [5]. In this research, the in-house developed program Erasmus-iCycle is used to generate and optimize the dose distribution. Hence, the output of the treatment planning process contains the values for each of the parameters that are required to deliver the desired dose distribution.

2.3 Fractionation

Radiotherapy is typically delivered in multiple fractions. As healthy tissue generally recovers faster than malignant cells, fractionation increases the therapeutic window [7]. Prior to each treatment fraction it is important to accurately position the patient to limit errors in dose deposition. To reduce geometrical variability of head and neck patients between and during treatment fractions, patients wear a mask that significantly limits freedom of movement. Also, to verify the patients' position, a cone-beam CT scan (a low dose CT-scan) is acquired to move the patient in the planned position. Still, over the course of a treatment, the patients' anatomy can change due to various reasons. These reasons include weight loss and tumor shrinkage as result of successful treatment. Therefore, during the course of the fractionated treatment, repeat CT (rCT) scans are made. These scans are used to verify whether the original treatment plan is still sufficient for the current anatomy. If not, a new treatment plan is generated and used for the remainder of treatment fractions.

In the current treatment schedule we use, head and neck cancer patients are treated with 35 fractions, spread over 7 weeks. Due to resources, it is impossible to generate a new planning CT scan every fraction. However, in every fraction, sufficient dose should be delivered to the CTV and minimal dose to the OARs. This highlights the importance of finding an optimal treatment plan that optimizes dose delivery to the CTV and spares healthy tissues in each treatment fraction.

2.4 Uncertainties

In both photon and proton therapy an adequate combination of machine parameters and patient setup are crucial to obtain optimal treatment delivery as slight deviations in these parameters can lead to undesired dose distributions [8]. In fact, some uncertainties might interfere with treatment quality. The main setup uncertainties and their corresponding magnitude, obtained from measurements at HollandPTC, are listed in Table 2.4.1. These uncertainties include errors in the CT isocenter, which is the error during positioning of the patient while acquiring the planning CT scan. Furthermore, MR-imaging is used for delineation of the CTV. Next, the CTV is propagated onto the planning CT scan. As the match between MRI and CT is likely to be imperfect, this results in an MR-registration error. Right before the treatment, the patients' position relative to the gantry is verified by the acquisition of a cone-beam CT scan. The cone beam CT scan is aligned with the planning CT scan. After alignment of the CT scans, a residual setup error remains [9]. The match between the cone-beam and planning CT scans is used to move the couch in the required position, which is not guaranteed to be exact. Other uncertainties are related to intra-fraction movement of the patient [6]. The total set-up error can be divided into a systematic and random part. The systematic error is defined as the mean error. Hence, it is the same for every treatment fraction. The random errors are variations around the mean error during a fractionated treatment.

Table 2.4.1: Main setup uncertainties in proton therapy with their corresponding magnitudes as measured in HollandPTC [8, 10].

Uncertainty	**Systematic error (mm)**	**Random error (mm)**
Isocenter CT	0.5	0
Isocenter gantry	0	0.5
Couch	0	0.5
MR registration	0.5	0
Residual setup (x,y,z)	0.21, 0.32, 0.38	0.52, 0.48, 0.61
Intra-fraction motion (x,y,z)	0.35, 0.31, 0.55	0.62, 0.49, 0.84

In proton therapy, there is an additional type of uncertainty compared to photon therapy. This

uncertainty is related to the proton range as the stopping power is predicted with uncertainty [11,12]. Underestimating and overestimating tissue density can lead to overshoot and undershoot of the proton beam, respectively. In head and neck cancer, the range error is estimated at 1.5% [13].

In photon therapy, uncertainties are typically dealt with by putting a margin around the CTV, resulting in the so called planning target volume (PTV). Subsequently, the PTV is irradiated with the prescribed dose to ensure sufficient CTV coverage. This approach is suitable for photon therapy, since we may assume that small deviations do not influence the shape of the dose distribution [14]. This assumption is known as the static dose cloud approximation. However, the static dose cloud approximation is not valid for protons. This can be observed in Figure 2.1.1. Here, the curve for protons is much steeper and therefore much more sensitive to small shifts of the curve. Hence, another approach has to be incorporated to obtain appropriate CTV coverage in proton therapy.

2.5 Robust optimization

One of the approaches that deals with uncertainties in IMPT treatment planning is robust optimization [15, 16]. In robust optimization, the treatment plan is optimized while assuming specific errors might occur. To this end, multiple treatment scenarios are considered. In the nominal treatment scenario, it is assumed that no errors occur. The other treatment scenarios are included to account for errors in proton range and setup. For errors in range, undershoot and overshoot are considered. For errors in setup, it is typically assumed that a patient is shifted in a specific direction. Any combination of treatment scenarios can be included in the optimization problem. With robust optimization, the aim is to obtain the planning goals in each treatment scenario accounted for. For each treatment scenario k, the dose distribution as presented in Equation (2.1) changes into the following discrete dose form:

$$\mathbf{d}_k = D_k \mathbf{x}. \tag{2.2}$$

Note that range and setup uncertainties correspond to uncertainties in the dose deposition matrix [15].

2.5.1 Minimax optimization

There are multiple robust optimization methods. One of the most used methods, which is currently used in Erasmus-iCycle, is minimax optimization [15–17]. In minimax optimization, the aim is to optimize the dose distribution for the worst treatment scenario that is accounted for. Therefore, it is also known as worst-case optimization. As a consequence, the quality of the treatment plan is determined by the best treatment plan that can be obtained in the worst possible treatment scenario. Mathematically, the minimax approach is represented as

$$\underset{\mathbf{x}}{\text{minimize}} \max_k [f(\mathbf{d}_k(\mathbf{x}))] \tag{2.3}$$

$$\text{subject to } \mathbf{d}_k = D_k \mathbf{x} \tag{2.4}$$

$$\mathbf{x} \geq 0. \tag{2.5}$$

Here, the objective function f is considered for each treatment scenario k. This objective function can be, for example, the function that computes the mean or maximum value of the dose. The goal is then to minimize the maximum objective function value that can be found for any of the treatment scenarios accounted for. Hence, we are looking for those pencil beam intensities that minimize the objective function value.

2.5.2 Stochastic programming

Another commonly used robust optimization method is stochastic programming, also referred to as the probabilistic approach [15,16]. Like minimax optimization, stochastic programming aims

to find the optimal solution while taking multiple treatment scenarios into account. However, in contrast to minimax optimization, the quality of the expected treatment scenario is optimized rather than the worst case treatment scenario. There exist multiple ways to formulate the corresponding optimization problem. In this research, it is assumed that every included treatment scenario k has an assigned weight w_k. These weights can be used to indicate the probability that treatment scenario k occurs. Hence, higher weights are assigned to the most likely treatment scenarios. During optimization, the expected value of the weighted sum of objective values is minimized. Therefore, the optimization problem is mathematically formulated as

$$\begin{aligned} &\underset{\mathbf{x}}{\text{minimize}} \sum_k w_k f(\mathbf{d}_k(\mathbf{x})) \\ &\text{subject to } \mathbf{d}_k = D_k \mathbf{x} \\ &\mathbf{x} \geq 0. \end{aligned}$$

Unkelbach et al. show the potential of stochastic programming in IMPT treatment planning [14, 18]. In [18] they tested the method on a circular phantom, representing a slice of a patient, containing a tumor that surrounds an OAR. The geometry of the phantom is presented in Figure 2.5.3.

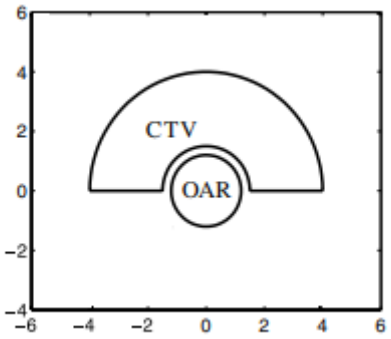


Figure 2.5.3: Geometry of the phantom where the OAR is surrounded by the CTV [18].

To test the method, two optimization cases are considered. In the first case, the treatment plan for the nominal treatment scenario is optimized without taking uncertainties into account. In the second case, a range uncertainty of 5 mm is considered and stochastic programming is used to find an optimal treatment plan. It was found that in the second case, improved target coverage and homogeneity were obtained compared to optimizing for the nominal treatment scenario. Furthermore, excessive OAR dose and underdosage to the tumor were avoided in case of range errors. However, in general, the OAR dose increases when accounting for range uncertainty.

In [14], the stochastic programming approach is tested for a patient with a paraspinal tumor. Again, two cases are considered. In the first case, the treatment plan is optimized accounting only for the nominal treatment scenario. In the second case, the systematic setup (2.5 mm) and range uncertainties (5 mm) are accounted for while optimizing the treatment plan using stochastic programming. They found that by accounting for uncertainties, the spinal cord is protected against overdoses. Furthermore, target coverage and homogeneity were obtained, respectively.

2.5.3 Multi-criterial optimization

Optimization of IMPT treatment plans is a multi-criterial optimization (MCO) problem [19]. In MCO, multiple objective functions are defined. Objectives include for example minimum dose in the target, maximum dose in the target and mean dose to OARs.

Erasmus-iCycle is a system for automatic MCO treatment planning and sequentially optimizes objectives while satisfying constraints. The objectives and constraints are specified in a wish-list, where all objectives have assigned priorities. In Table 2.5.2 a simplified version of a wish-list is shown. The top row indicates a constraint and in the bottom rows, the prioritized objectives are listed. The aim is to find a plan that satisfies all objectives within the given constraints, where it is most important to satisfy the objective with highest priority. The general formula of the optimization problem becomes:

$$\begin{aligned} &\underset{\mathbf{x}}{\text{minimize}}[f_1(\mathbf{x}), ..., f_n(\mathbf{x})] \\ &\text{subject to } g_j(\mathbf{x}) \leq b_j, \\ &\mathbf{x} \geq 0. \end{aligned}$$

Here, n is the number of objectives, g_j with $1 \leq j \leq m$ is the number of constraints and b_j is the goal value of a particular constraint [5]. In each step of the MCO process one of the objective functions f_i is minimized while satisfying all constraints and higher prioritized objectives. After the MCO process, a Pareto optimal treatment plan is obtained. Hence, an objective can only be improved at the cost of another objective.

The optimization process of Erasmus-iCycle consists of two steps. In the first step, if an objective can attain its desired value, the objective is assigned its highest acceptable value (for a minimization problem). Hence, the objective is not further minimized even though this might be possible. Thereby offering room for the other objectives to be satisfied as well. In the second step, all objectives are assigned their optimal value, in order of their priority. To find the optimum of an objective function, a primal-dual interior-point approach is used. This approach requires both the gradient and the Hessian of the objective functions to find the step direction for each iteration [20].

Table 2.5.2: An example of a wish-list.

	Structure	**Type**	**Goal**	**Sufficient**
constraint	CTV	min	67.9 Gy	67.9 Gy
Priority				
1	CTV	minimize max	74.2 Gy	74.2 Gy
2	Brain stem	minimize max	30 Gy	1 Gy
2	Spinal cord	minimize max	30 Gy	1 Gy
3	Brain	minimize max	30 Gy	1 Gy
4	Left parotid	minimize mean	1 Gy	1 Gy
4	Right parotid	minimize mean	1 Gy	1 Gy
5	Oral cavity	minimize mean	1 Gy	1 Gy

2.6 Evaluation methods

In treatment planning it is crucial to obtain sufficient CTV coverage while satisfying OAR constraints [15]. In robust evaluation it is investigated whether developed treatment plans obtain sufficient CTV coverage in the treatment scenarios accounted for. In conventional photon therapy there exist common ways to evaluate robustness. These evaluation methods for photons are PTV based. Here, PTV V95% $>$ 98% is typically used, which means the volume of the PTV that receives more than 95% of the prescribed dose is over 98%. However, as discussed in Section 2.4, the PTV approach is not valid for proton therapy. Therefore, other evaluation methods are needed to evaluate robustness.

To develop similar standards for robustness evaluation of IMPT treatment plans, Korevaar et al. [11] investigated multiple methods. These are explained in Section 2.6.1.

2.6.1 Voxelwise evaluation

To develop similar rules for robustness evaluation of IMPT treatment plans compared to photon treatment plans, Korevaar et al. [11] used a two-step approach.

In the first step, the goal is to find summary evaluation dose distributions to allow for easy plan assessment in daily clinical practice. A summary evaluation dose distribution summarizes the location of under-dosage in treatment scenarios. Three summary evaluation dose distributions that have been investigated are voxelwise minimum dose, voxelwise mean dose and worst case scenario dose. An overview of these methods is presented in Figure 2.6.4. The voxelwise minimum dose is the minimum dose to each voxel over all treatment scenarios. The voxelwise mean dose, is the mean dose for each voxel over all treatment scenarios. In worst-case treatment scenario dose, an entire treatment scenario which shows lowest CTV D98% is stored. Here, CTV D98% is the minimum dose to 98% of the CTV.

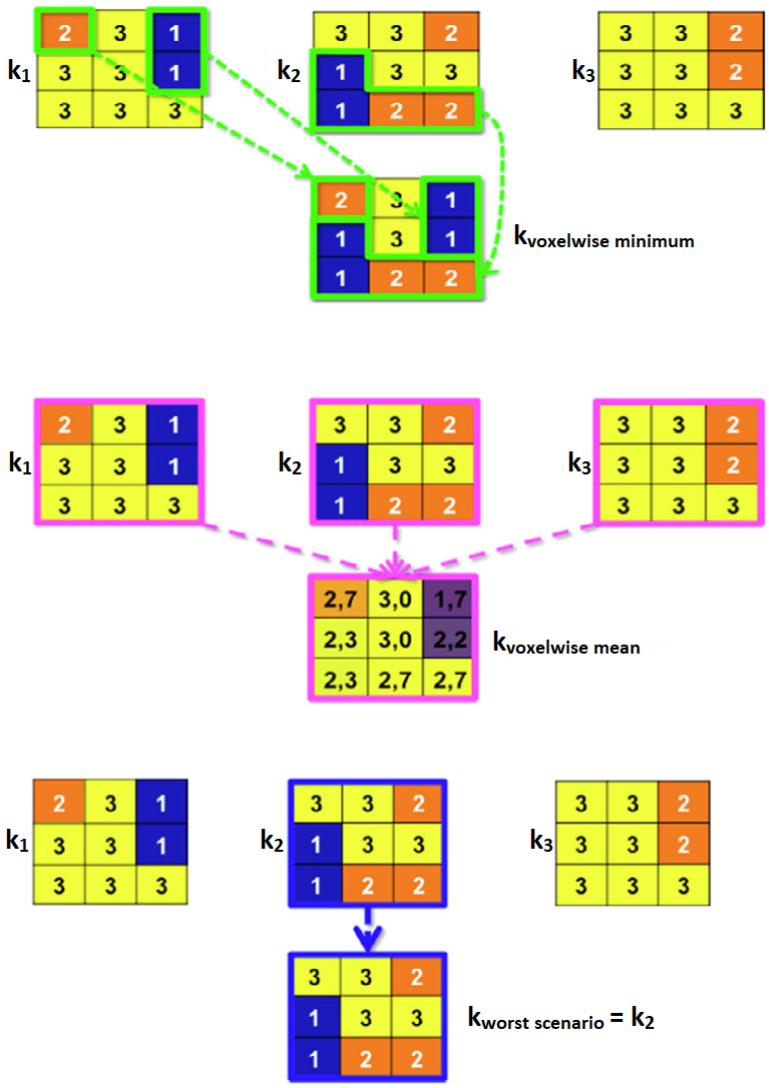


Figure 2.6.4: An illustration of voxelwise min, voxelwise mean dose and worst scenario summary evaluation dose distributions [11].

To quantify robustness in a similar manner as in the PTV approach for photons, Korevaar et al. determine the correlation between CTV D98%, CTV V95% loss and the average CTV D98% and CTV V95% loss for the summary evaluation dose distributions of all considered treatment scenarios, respectively for photon and proton plans. Here, V95% loss is the percentage of the CTV receiving less than 95% of the prescribed dose.

The proton and photon plans using voxelwise minimum, voxelwise mean and worst scenario dose all showed strong correlation with the CTV D98% metric. The V95 loss metric showed much weaker correlations. Even though voxelwise minimum, voxelwise mean and worst scenario all show similar correlation for the CTV D98% metric, voxelwise minimum is expected to be most convenient. Voxelwise minimum allows clinicians to judge under-dosage in a similar way to the conventional PTV-approach. In the other methods, locations of under-dosage can be missed.

In the second step, the acceptance criteria CTV D98% $>$ 95% was compared against the PTV D98% $>$ 95% for photons. Very strong correlation was found between PTV D98% and CTV D98% using voxelwise minimum. Therefore, according to [11], a treatment plan is called robust if CTV D98% $>$ 95% using the voxelwise minimum approach.

2.6.2 PCE evaluation

The approach used in Section 2.6.1 considers only a limited and fixed number of treatment scenarios to evaluate robustness. Whereas in reality, the error changes each treatment fraction. Therefore, another approach called polynomial chaos expansion (PCE) evaluation is used to simulate many fractionated treatments to evaluate the robustness while modeling the uncertainties more precisely. In PCE, the relation between a random variable (here, the dose distribution) and uncertain input values (here, the setup and range uncertainties) is represented as a polynomial function [8, 21]. To show how the PCE of the dose is computed, we follow the reasoning of [21]. To compute the PCE of the dose, we first define the vector $\boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\delta} \\ \rho \end{pmatrix}$ with $\boldsymbol{\xi} \in \mathbb{R}^N$. Here, N indicates the number of uncertainties accounted for. Furthermore, $\boldsymbol{\delta} \in \mathbb{R}^{N-1}$ contains setup uncertainties and ρ represents the uncertainty in range. Then, the dose as depending on $\boldsymbol{\delta}$ and ρ, using PCE, is represented as [8]

$$\mathbf{d}(\boldsymbol{\xi}) = \sum_{l=0}^{P} c_l \boldsymbol{\psi}_l(\boldsymbol{\xi}). \tag{2.6}$$

Here, c_l are expansion coefficients and $\boldsymbol{\psi}_l(\boldsymbol{\xi})$ are the polynomial basis vectors. P indicates the number of basis vectors that are used for the PCE computation. In general, the accuracy of the PCE increases if P increases. However, including more basis vectors also increases the computation time [22]. To determine the basis vectors, the distribution of the input values must be determined. The distribution of $\boldsymbol{\xi}$ is given by the joint probability density function (PDF) of all input variables. Assuming that the uncertainties accounted for are independent, we have that the joint PDF of the input variables is the product of the individual PDF of the input variables. Hence, in the remainder of this thesis we assume that the joint PDF $p_{\boldsymbol{\xi}}(\boldsymbol{\xi})$ is given by

$$p_{\boldsymbol{\xi}}(\boldsymbol{\xi}) = \prod_{j=1}^{N} p_{\xi_j}(\xi_j).$$

Next, orthogonal polynomial basis functions must be specified. Typically, the probabilist's Hermite polynomials are used for normally distributed parameters. The probabilist's Hermite polynomials are defined as:

$$He_n(\xi) = (-1)^n e^{\frac{\xi^2}{2}} \frac{\mathrm{d}^n}{\mathrm{d}\xi^n} e^{-\frac{\xi^2}{2}},$$

with n the degree of the Hermite polynomial. Then, the polynomial chaos basis vectors in Equation (2.6) can be written as

$$\boldsymbol{\psi}_l(\boldsymbol{\xi}) = \prod_{j=1}^{N} He_{j,\, g_{lj}}(\xi_j). \tag{2.7}$$

Here, the multi-index $g_{l,j}$ is used to indicate that the j-th polynomial of the l-th basis vector is of order g_{lj}. Finally, the c_l expansion coefficients in Equation (2.6) can be calculated using linear regression.

This approach has been implemented by Rojo Santiago et al. [8]. They have tested the approach on the data of 26 neuro-oncological patients. Amongst their treatment plans are both clinically robust and clinically non-robust plans. Clinically robust plans are plans in which all clinical robust evaluation constraints are met. In clinically non-robust plans, robustness is sacrificed for OAR sparing. They have used PCE to efficiently estimate the dose distribution depending on the uncertainties in 100.000 treatment fractions for each patient. To evaluate the treatment plans, the following standard is used:

$$\text{Voxelwise minimum : CTV D98\%} \geq 95\%.$$

Here, the voxelwise minimum summary evaluation dose is as explained in Section 2.6.1. It was found that for 97.7% of all generated treatment plans, the above standard is satisfied. They also found that for clinically robust plans, the CTV dose is on average 3 percent points higher than the prescribed dose.

It is concluded that this method is reliable as they found that all clinically robust plans were also robust according to the PCE evaluation. Furthermore, if plans were not robust according to PCE evaluation, CTV coverage was insufficient.

For the planning methods in Sections 3, 4 and 5 a voxelwise (minimum) evaluation method, as explained in Section 2.6.1, is used. In addition, for the planning method in Section 3, Polynomial Chaos Expansion (PCE) is used to evaluate the treatment plans, as explained in Section 2.6.2. In Table 2.6.3 an overview is presented of which evaluation methods are used for each of the methods in Section 3 - 5. Note that we also need to evaluate if the maximum dose to the CTV is acceptable. Hereto, we often use the voxelwise maximum evaluation method. Similar to voxelwise minimum, voxelwise maximum is the maximum dose to each voxel over all treatment scenarios.

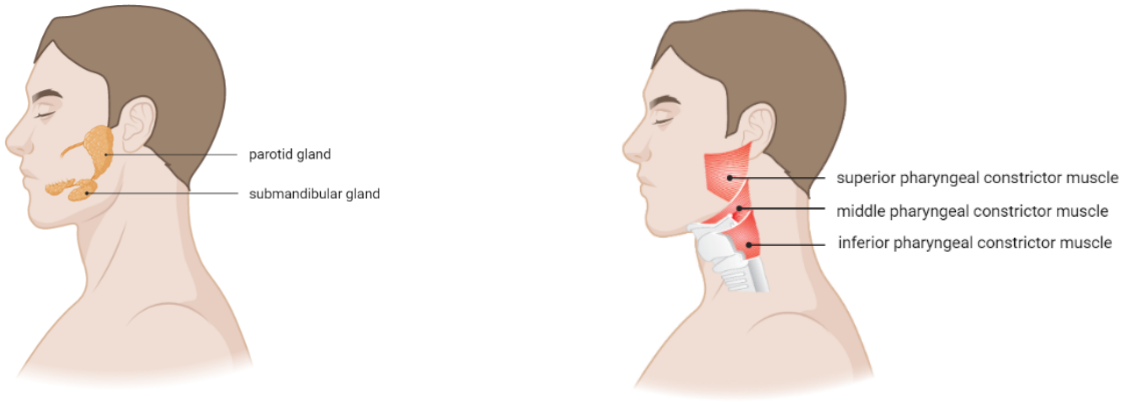
Table 2.6.3: Overview of planning and evaluation methods used in this research.

Planning method	**Evaluation method**
Treatment scenario reduction	voxelwise and PCE
Dose escalation	voxelwise
Variance optimization	voxelwise

2.7 Medical information

In this study, four head and neck cancer patients are included. All patients received IMPT at Holland PTC. Here, we elaborate on some medical information that is relevant for this patient group.

Dose used to irradiate the tumor, also affects healthy tissues. It is crucial to reduce the dose to healthy tissue as much as possible to minimize the risk of radiotherapy induced complications. In head and neck cancer, two main (chronic) complications are xerostomia and dysphagia. Xerostomia, also known as dry mouth syndrome, is caused by malfunctioning of the parotid and submandibular glands. These glands are two of the main salivary glands, their location is highlighted in Figure 2.7.5a. Patients with dysphagia experience difficulty in swallowing. To minimize the risk of dysphagia it is important to spare the oral cavity and the pharyngeal constrictor muscles as much as possible. The latter are often referred to as the constrictor muscles. These muscles are used to move food into the esophagus [23]. In Figure 2.7.5b these muscles are schematically displayed.



(a) Main salivary glands.

(b) Pharyngeal constrictor muscles.

Figure 2.7.5: Anatomical overview of the main salivary glands and the pharyngeal constrictor muscles. Created with BioRender.com

Normal Tissue Complication Probability (NTCP) models have been developed to establish the relationship between dose to the organs at risk and clinical relevant toxicities [24]. Higher NTCP indicates increased risk of a particular complication. The general formula for the NTCP is given by:

$$\text{NTCP} = \frac{1}{1 + e^{-s}}, \tag{2.8}$$

where s is dependent on the complication. For xerostomia s depends, amongst other things, on the mean dose in the parotids and submandibular glands. For dysphagia s is, amongst other things, dependent on the mean dose in the oral cavity and the pharyngeal constrictor muscles. The objective is to obtain NTCPs as low as possible. Note that we do not optimize on the NTCPs directly, as Equation (2.8) is not convex. When a difference of 10%-point in NTCP can be attained between photon therapy and proton therapy, patients are referred to proton therapy. As the difference is often near this 10%-point, a difference of 2%-point in NTCPs can already have significant clinical impact [10].

In this study we focus on the risk of patients getting xerostomia and dysphagia grade II or higher. The grade of the complication indicates the severity. For both xerostomia and dysphagia, four grades are considered, where grade IV represents the most severe level.

2.8 Treatment plan criteria

Here, we list the mean criteria that are used to assess the generated treatment plans.

The CTV of the patients in this study is split into two main parts: CTV_{7000} and CTV_{5425}. CTV_{7000} consists of the primary tumor and is prescribed 70.00 Gy. The other part, CTV_{5425}, which comprises the regions at risk of containing cancerous cells, is prescribed 54.25 Gy. To obtain clinically acceptable treatment plans, several criteria must be met. First, according to the robustness criteria, the D98% CTV_{7000} and D98% CTV_{5425} should be at least 95% of the prescribed dose. That is, D98% $\text{CTV}_{7000} \geq 66.50$ Gy and D98% $\text{CTV}_{5425} \geq 51.54$ Gy. In this study, unless mentioned otherwise, generated treatment plans are scaled such that D98% $\text{CTV}_{7000} \geq 66.50$ Gy and D98% $\text{CTV}_{5425} \geq 51.54$ Gy in the voxelwise minimum scenario (Section 2.6.1) on the planning CT scan. Therefore, the D98% CTV_{7000} and D98% CTV_{5425} are typically evaluated for the repeat CT scan to investigate if robustness can be maintained throughout the fractionated treatment. Another criteria used in this study indicates that D2% $\text{CTV}_{7000} \leq 74.90$ Gy. Here, D2% CTV_{7000} is the maximum dose to 2% of CTV_{7000}. Furthermore, the NTCPs

should be as low as possible. This is achieved by sequential minimization of the mean dose to the OARs. The wish-list used throughout this research is presented in the Appendix A.

3 Treatment scenario reduction

In order to generate robust treatment plans, we want to account for as many treatment scenarios as possible. However, every included treatment scenario increases memory storage as typically all corresponding dose deposition matrices are stored. In addition, the increased size of the optimization problem causes a substantial increase in computation time. To reduce both memory storage and computation time, we investigate the dosimetric impact of the inclusion of fewer treatment scenarios in this section.

3.1 Method

In order to generate robust treatment plans, typically 29 treatment scenarios are taken into account during treatment plan optimization in Erasmus-iCycle. These treatment scenarios include the nominal treatment scenario, where no error is accounted for, and 28 treatment scenarios in which a setup error of $\pm$ 3 mm and a range error of $\pm 3\%$ are considered. A schematic overview of the distribution of the setup errors accounted for using 29 treatment scenarios is depicted in Figure 3.1.1. Note that in this figure, for each point, except the center point which represents the nominal treatment scenario, twice the number of treatment scenarios is considered. That is, once the setup error with positive range error and once the setup error with negative range error.

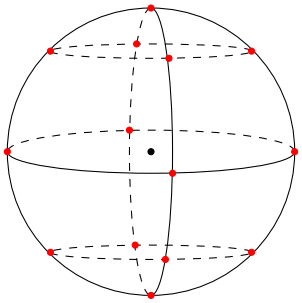


Figure 3.1.1: Distribution of the setup errors in 29 treatment scenarios. The center dot (black) represents the nominal treatment scenario. The other dots (red) represent the setup error used in the other treatment scenarios.

To investigate the impact of the number of treatment scenarios on both the optimization time and treatment plan quality, three test cases are constructed. In each of these test cases, treatment plans are generated while accounting for a specific number of treatment scenarios. In the first test case, a treatment plan is generated by considering only the nominal treatment scenario. Then, for the second test case, 13 out of the 29 original treatment scenarios are included. For the final test case, treatment scenarios are added to the original 29 treatment scenarios to obtain a total of 53 treatment scenarios. A schematic overview of the distribution of the setup errors in 13 and 53 treatment scenarios is shown in Figure 3.1.2.

To investigate the dosimetric impact of the number of treatment scenarios, we compare the treatment plans that are generated with 1, 13, 29 and 53 treatment scenarios. To assess dose delivery to the CTV, we compare D98% CTV_{7000} and D98% CTV_{5425} on the repeat CT scan in the voxelwise minimum scenario, as explained in Section 2.6.1. Here, the voxelwise minimum scenario is based on the 29 treatment scenarios depicted in Figure 3.1.1. Furthermore, we compare the D2% CTV_{7000} for all test cases in the voxelwise maximum scenario on both the planning and repeat CT scan. Finally, the NTCPs of xerostomia and dysphagia are compared for all test cases in the nominal treatment scenario.

Additionally, we evaluate the generated treatment plans with PCE, as explained in Section 2.6.2. Hereto, we sample 10.000 fractionated treatment scenarios from the normal distribution

such that the standard deviation of the setup error is defined by the total error in Table 2.4.1. Furthermore, the standard deviation of the range is 1.5%.

(a) Setup errors in 13 treatment scenarios. (b) Setup errors in 53 treatment scenarios.

Figure 3.1.2: Distribution of the setup errors in 13 and 53 treatment scenarios. The center dot (black) represents the nominal treatment scenario. The other dots (red) represent the setup error of the other treatment scenarios.

3.2 Results

First we present the optimization time for the different test cases in Section 3.2.1. Subsequently, in Section 3.2.2 the dosimetric results of the test cases, evaluated with both voxelwise and PCE evaluation, are presented. As we obtained results with clinically unacceptable results when optimizing only for the nominal treatment scenario, these results are shown in Appendix B. Furthermore, to gain a better understanding of the results, the evaluation results of the unscaled treatment plans are shown in Appendix B.

3.2.1 Optimization time

In Table 3.2.1, the optimization times in minutes when optimizing for 13, 29 and 53 treatment scenarios are listed per patient. We observe some inter patient variation in optimization times. However, for all patients, a reduction in optimization time between 59% and 75% is obtained when reducing the number of treatment scenarios from either 29 to 13 or from 53 to 29.

Table 3.2.1: Optimization time (in minutes) for different number of treatment scenarios per patient.

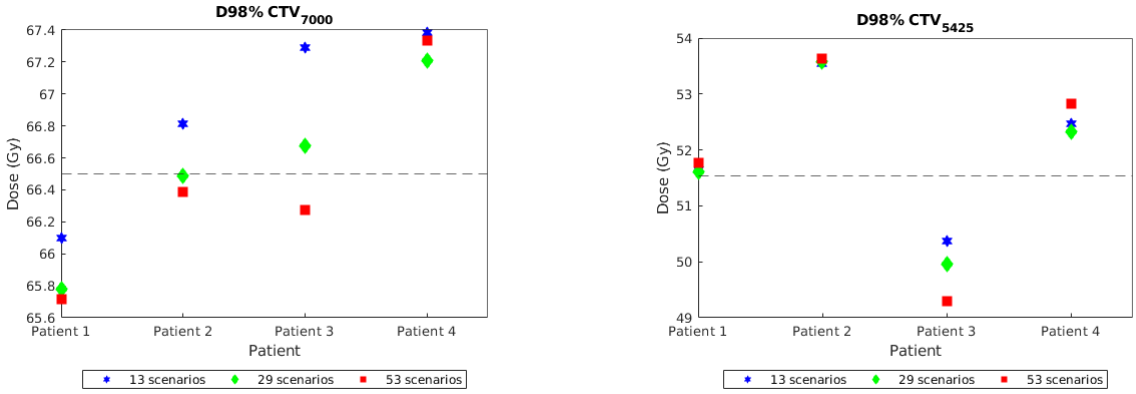
# scenarios	Patient 1	Patient 2	Patient 3	Patient 4
13	351	226	195	197
29	1085	622	614	611
53	3572	2433	1507	2406

3.2.2 Treatment plan quality

In Figure 3.2.1, the D98% on the repeat CT scan is shown in the voxelwise minimum scenario when optimizing for 13, 29 and 53 treatment scenarios. Figure 3.2.1a shows the results for CTV_{7000} and Figure 3.2.1b shows the results for CTV_{5425}. Similarly, in Figure 3.2.2, D98% is shown using PCE evaluation. In addition, the percentage of treatment plans that satisfies the robustness criteria D98% $\text{CTV}_{7000} \geq 66.5$ Gy and D98% $\text{CTV}_{5425} \geq 51.54$ Gy when optimizing for 13, 29 and 53 treatment scenarios is presented, respectively.

According to Figure 3.2.1, the D98% CTV_{7000} is for all patients highest when 13 treatment scenarios are used. The increase in D98% CTV_{7000} when 13 instead of 29 treatment scenarios are used, ranges from 0.2 to 0.6 Gy. However, D98% CTV_{5425} is for three out of four patients highest using 53 treatment scenarios. We also observe that some treatment plans show insufficient D98%.

In these cases, a lower number of treatment scenarios results in improved, even if insufficient, CTV coverage. In Figure 3.2.2 we observe that the results evaluated with PCE show similar behaviour to the results evaluated in the voxelwise minimum scenario. We find that the number of simulated treatment plans with D98% $\text{CTV}_{7000} \geq 66.50$ Gy increases when less treatment scenarios are included. When including 13 treatment scenarios at least 94.7% of the treatment plans receives sufficient D98% CTV_{7000}, while for 53 treatment scenarios, this is 90.1%. For CTV_{5425}, we obtain either improved or similar robustness when optimizing for less treatment scenarios.



(a) D98% CTV_{7000}.

(b) D98% CTV_{5425}.

Figure 3.2.1: D98% CTV_{7000} and D98% CTV_{5425} when optimizing for 13, 29 and 53 treatment scenarios. The results are evaluated on the repeat CT scan in the voxelwise minimum scenario.

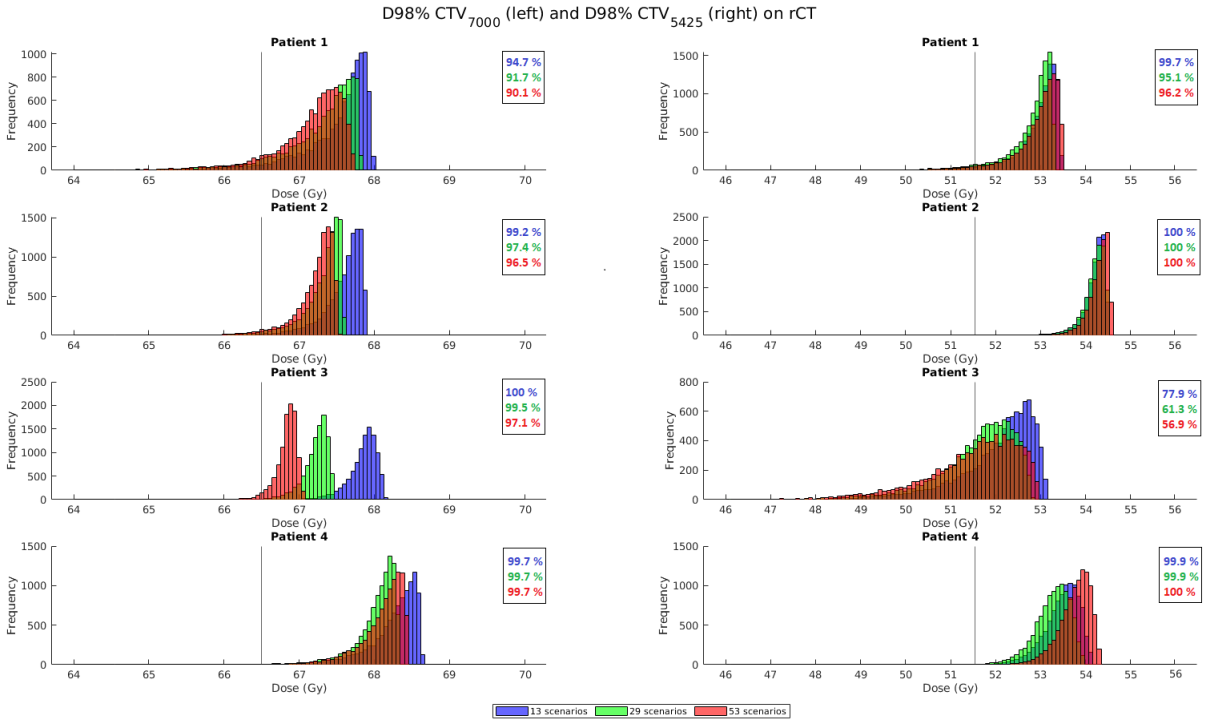
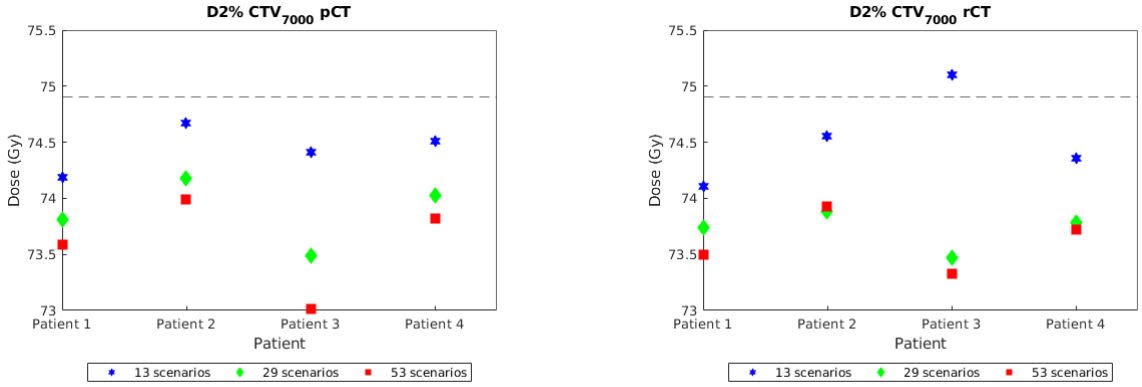


Figure 3.2.2: D98% CTV_{7000} and D98% CTV_{5425} when optimizing for 13, 29 and 53 treatment scenarios. The results are evaluated on the repeat CT scan using polynomial chaos expansion. The values on the top right indicate for each number of treatment scenarios how often the D98% $\text{CTV}_{7000} \geq 66.5$ Gy and D98% $\text{CTV}_{5425} \geq 51.54$ Gy criteria are satisfied.

Figure 3.2.3 shows D2% CTV_{7000} in the voxelwise maximum scenario when optimizing for 13, 29 and 53 treatment scenarios. Figure 3.2.3a shows the results for the planning CT scan and Figure 3.2.3b shows the results for the repeat CT scan. The results that were obtained from PCE evaluation are presented in Figure 3.2.4.

In Figure 3.2.3, we find that D2% CTV_{7000} is highest when optimizing for 13 treatment scenarios. The increase in D2% CTV_{7000} ranges from 0.3 to 1.6 Gy when optimizing for 13 instead of 29 treatment scenarios. Note that it is required that D2% $CTV_{7000} \leq 74.90$ Gy. This criterion is exceeded for patient 3 on the repeat CT scan when using 13 treatment scenarios. When evaluating with PCE, we again find that higher D2% CTV_{7000} is generally obtained when optimizing on fewer treatment scenarios, as shown in Figure 3.2.4. In contrast to voxelwise maximum evaluation, none of the treatment plans exceeds the 74.90 Gy threshold with PCE evaluation.



(a) D2% CTV_{7000} on planning CT scan.

(b) D2% CTV_{7000} on repeat CT scan.

Figure 3.2.3: D2% CTV_{7000} when optimizing for 13, 29 and 53 treatment scenarios. The results are evaluated in the voxelwise maximum scenario on both planning and repeat CT scan.

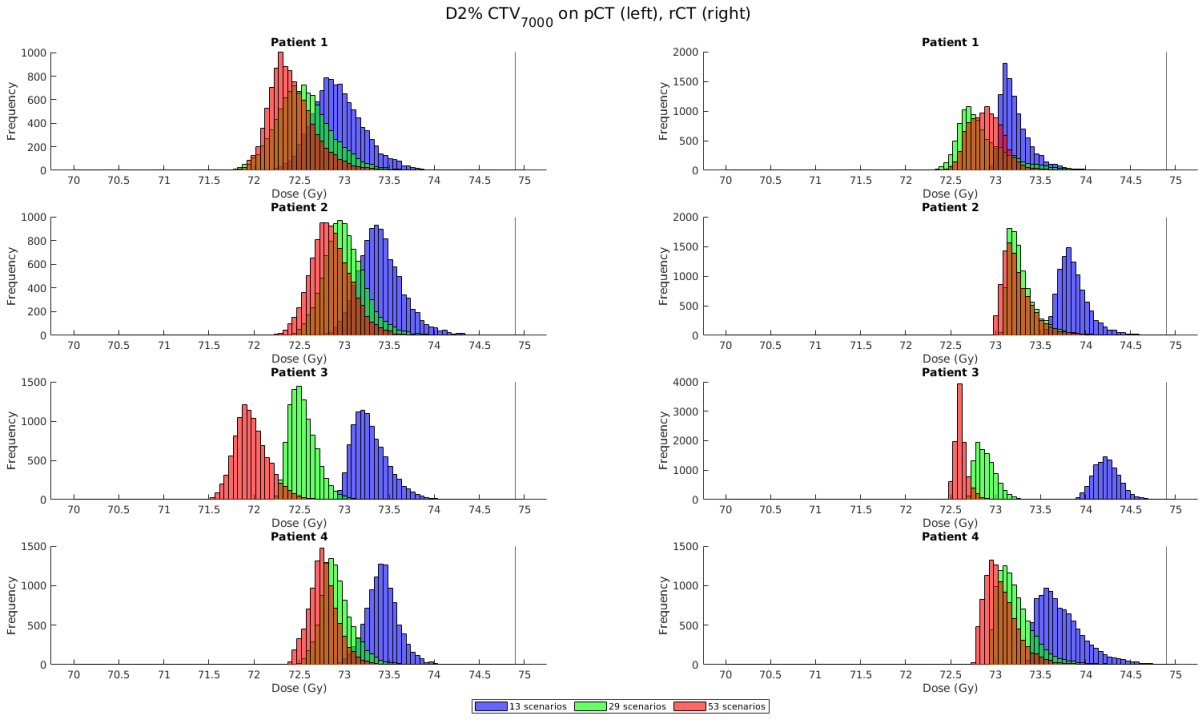
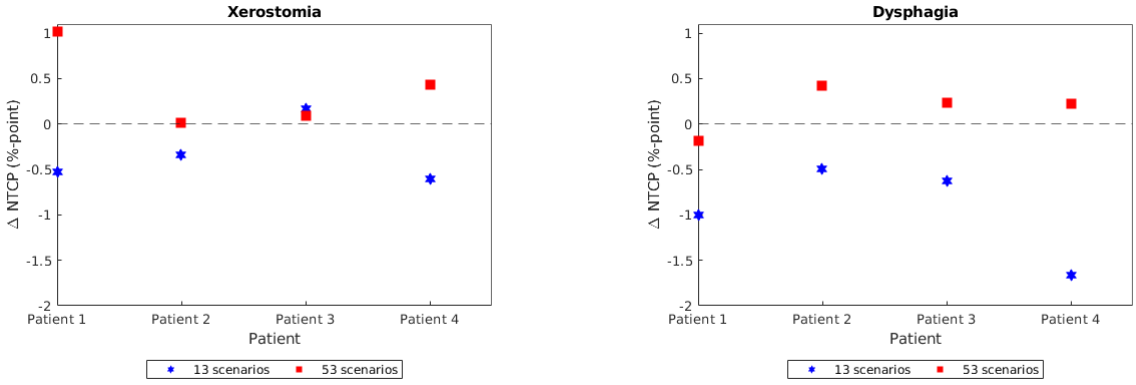


Figure 3.2.4: D2% CTV_{7000} when optimizing for 13, 29 and 53 treatment scenarios. The results are evaluated with polynomial chaos expansion on planning (left) and repeat (right) CT scan.

In Figure 3.2.5, the difference (%-point) in NTCPs in the nominal treatment scenario is shown when optimizing for 13 and 53 instead of 29 treatment scenarios. Figure 3.2.5a shows the results for xerostomia. Similarly, Figure 3.2.5b shows the results for dysphagia. The results that were obtained with PCE evaluation are presented in Figure 3.2.6. Since we obtained similar NTCP results for the planning and repeat CT scans, only the results on the repeat CT scan are shown.

In Figure 3.2.5, we observe a general reduction of 0.3 to 1.7%-point when optimizing for 13 treatment scenarios. However for patient 3, increased NTCP of around 0.2%-point for xerostomia is obtained when optimizing on 13 instead of 29 treatment scenarios. When evaluating with PCE, similar results are obtained. In Figure 3.2.6 we observe a reduction in NTCPs when 13 treatment scenarios are optimized for.



(a) NTCP of xerostomia.

(b) NTCP of dysphagia.

Figure 3.2.5: NTCPs for xerostomia and dysphagia when optimizing for 13 and 53 treatment scenarios with respect to optimizing for 29 treatment scenarios. The results are evaluated in the nominal treatment scenario on the repeat CT scan.

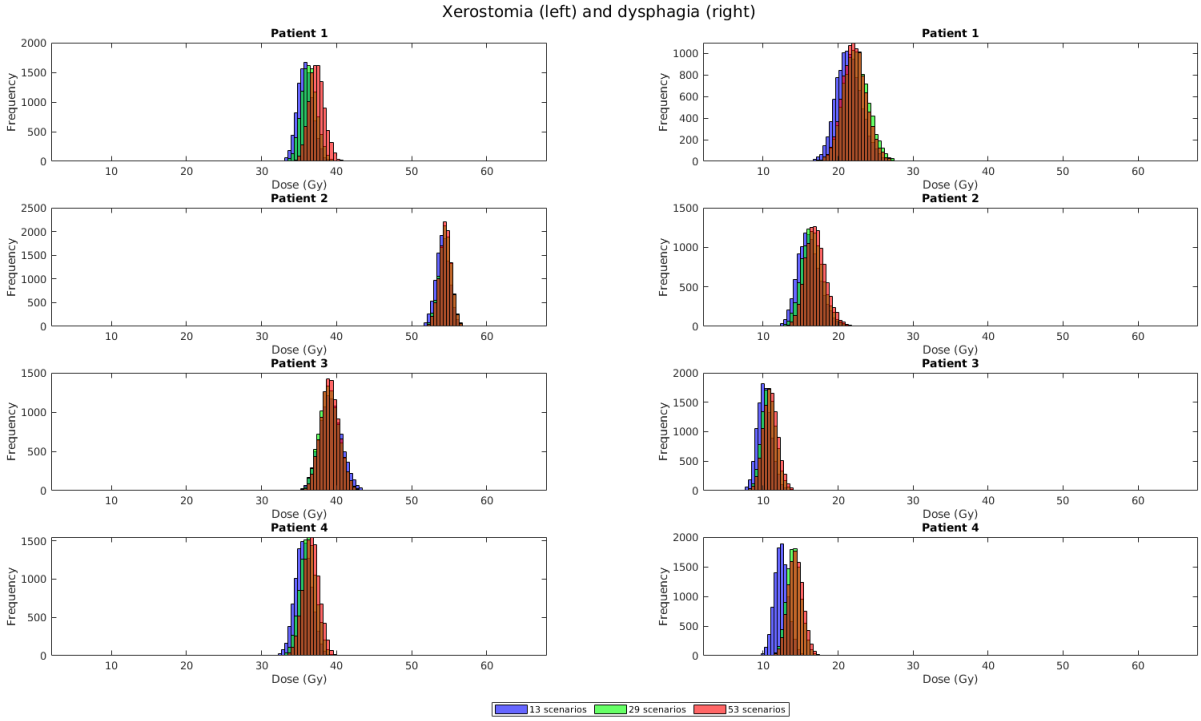


Figure 3.2.6: NTCPs for xerostomia and dysphagia when optimizing for 13, 29 and 53 treatment scenarios. The results are evaluated using polynomial chaos expansion on the repeat CT scan.

3.3 Discussion

In this study, the impact of varying the number of treatment scenarios during optimization is studied. The objective is to reduce optimization time and preserve treatment quality. The effect on the dosimetry of the generated treatment plans is evaluated on the repeat CT scans. To still achieve equal target coverage on the planning CT scan, all treatment plans are scaled to meet the clinical target coverage constraints on 29 treatment scenarios. We find that by reducing the number of treatment scenarios, a significant reduction in optimization time is achieved. However, reducing the number of treatment scenarios also affects the dosimetry of the generated treatment plans.

When comparing the performance of the treatment plans on 13, 29 and 53 treatment scenarios, we find that CTV coverage on the repeat CT scan generally improves when using less treatment scenarios. This effect is mainly observed for CTV_{7000}. Additionally, NTCPs generally improve when reducing the number of treatment scenarios. However, we observe an increase in D2% CTV_{7000} when planning on fewer treatment scenarios. For one patient, the D2% $\text{CTV}_{7000} \leq 74.90$ Gy criterion is violated when optimizing on 13 treatment scenarios in the voxelwise maximum scenario. Note that the voxelwise maximum scenario represents a physically impossible scenario as it comprises the maximum dose to each voxel over 29 treatment scenarios. When evaluating with PCE, none of the treatment plans exceeds the 74.90 Gy threshold.

Evaluation of the dose distribution with PCE is a strong aspect of this study. It allows for assessment in 10000 normally distributed fractionated treatment scenarios rather than in 13, 29 or 53 fixed treatment scenarios. Moreover, unlike the voxelwise minimum scenario with 29 treatment scenarios, the evaluation metric is independent of the chosen treatment scenarios. Another strong aspect of this study is the evaluation of the dose distributions on repeat CT scans. This allows to assess the robustness of the treatment plans under actual inter-fraction anatomical changes.

The treatment plans are scaled to meet exactly the clinical constraints on the planning CT scan in 29 treatment scenarios. Higher scaling factors are required to obtain good CTV coverage when using 13 treatment scenarios whereas the treatment plans on 53 treatment scenarios are scaled down. For most patients on the unscaled treatment plans, D2% CTV_{7000} is relatively similar, as is shown in Appendix B. For different number of treatment scenarios, the D2% CTV_{7000} is within a 0.3 Gy range on the planning CT scan. Due to scaling, the D2% CTV_{7000} is higher on 13 treatment scenarios and lower on 53 treatment scenarios. However, by scaling the treatment plans, lower NTCPs are obtained when less treatment scenarios are used.

Note that only the D2% CTV_{7000} is shown in the results. However, possibly other objectives, such as the near-maximum doses in regions surrounding CTV_{7000} are also dosimetrically impacted. For the sake of simplicity, these are not evaluated in this section.

As we observe a higher D2% CTV_{7000} together with improved CTV coverage and lower NTCPs, the results gave rise to the following hypothesis: increasing the maximum allowable CTV dose, which increases the degrees of freedom of the problem, might result in improved coverage on repeat CT scans and reduced NTCPs. Therefore, we test this hypothesis in Section 4.

In conclusion, reducing the number of treatment scenarios in optimization lead to a significant reduction in the optimization time. Furthermore, CTV coverage and NTCPs of generated treatment plans generally improve on the repeat CT scan, but D2% CTV_{7000} increases. However, D2% CTV_{7000} is still clinically acceptable for all patients when using PCE evaluation.

4 Dose escalation

Increasing robustness settings leads to increased dose to healthy tissues [25]. To reduce this undesired dose, a method called dose escalation is evaluated. In dose escalation, the constraint on the maximum tumor dose is relaxed which increases the degrees of freedom for the problem. As a result, the tumor is expected to receive a higher maximum dose whereas the dose in healthy tissues should be reduced.

In this section we examine if increasing the maximum allowable dose in the tumor, and to which specific parts of the tumor, can be used to improve treatment plan quality. To this end, we first provide results obtained from other experiments with dose escalation in Section 4.1. Then, in Section 4.2 we elaborate on the applied experimental procedure. The corresponding results are presented and discussed in Sections 4.3 and 4.4, respectively.

4.1 Literature

An experiment with dose escalation was previously performed by Petit et al. [26]. They generated different groups of treatment plans. For the first type of plans, a robust optimization strategy is used. This strategy is a mixture between the minimax method and the stochastic programming approach as the treatment plan is optimized for the nominal (most likely) treatment scenario, while ensuring sufficient CTV coverage in the worst-case treatment scenario. Another group of plans is constructed to minimize dose to the OARs by relaxing the maximum tumor dose. Petit et al. have only taken range uncertainties into account, meaning only the nominal, undershoot and overshoot treatment scenarios are considered. The general robust optimization problem used, is as follows:

$$
\begin{aligned}
\text{minimize} \quad & \frac{1}{N_R} \sum_{i=1}^{N_R} d_i^{nom}, \ i \in R \\
\text{subject to} \quad & \min_{i \in T} d_i^{nom} \geq l_T \\
& \min_{i \in T} d_i^{under} \geq l_T \\
& \min_{i \in T} d_i^{over} \geq l_T \\
& \max_{i \in P} d_i^{nom} \leq u_P \\
& \mathbf{x} \geq 0.
\end{aligned}
$$

Here, the mean dose over all N_R voxels inside a ring R around the target T is minimized. Typically, the width of this ring is in the order of a few mm, here 2 mm is used. l_T and u_P represent the lower and upper bound on the target and patient dose, respectively. d^{nom}, d^{under} and d^{over} represent the dose in the nominal, undershoot and overshoot treatment scenario, respectively. Hence, the constraints ensure sufficient target coverage in all of these treatment scenarios. By minimizing the mean dose in R, the aim is to obtain optimal OAR sparing close to the target.

Both types of optimization problems have been tested on a liver patient and two phantom cases. It was found that by applying only robust optimization, without dose escalation, an increased amount of dose is deposited in the region close to the target. However, by allowing dose escalation, this extra dose that was originally delivered close to the target can now be delivered inside the target and thereby spare the surrounding OARs. As a result, the total dose in the patient increases, but more dose is delivered to the target and less dose is delivered to the OARs compared to robust optimization without dose escalation. More specifically, it is found that an increase in maximum tumor dose from 105% to 120% resulted in healthy tissue dose similar to non-robust plans with a maximum tumor dose of 105%. Hence, for the investigated cases, combining robust optimization with an increase in maximum tumor dose leads to lower OAR dose.

Similar research has been performed on this topic for lung patients using photon therapy [27]. Here, it was found that relaxing the maximum constraint on the tumor dose results in better tumor eradication.

4.2 Method

To assess the viability of the dose escalation method, we both relax and remove the maximum dose constraint to different parts of the CTV. These parts, CTV_{7000}, $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{5425\text{shrunk}}$, are defined by the following relations:

$$\text{CTV}_{\text{int10mm}} = \text{CTV}_{5425} \cap \text{CTV}_{7000\text{r0-10mm}},$$
$$\text{CTV}_{5425\text{shrunk}} = \text{CTV}_{5425} \setminus (\text{CTV}_{7000} \cup \text{CTV}_{7000\text{r0-10mm}}),$$

where $\text{CTV}_{7000\text{r0-10mm}}$ is an isotropic ring around CTV_{7000} of 10 mm. For clarification, these parts are schematically depicted in Figure 4.2.1.

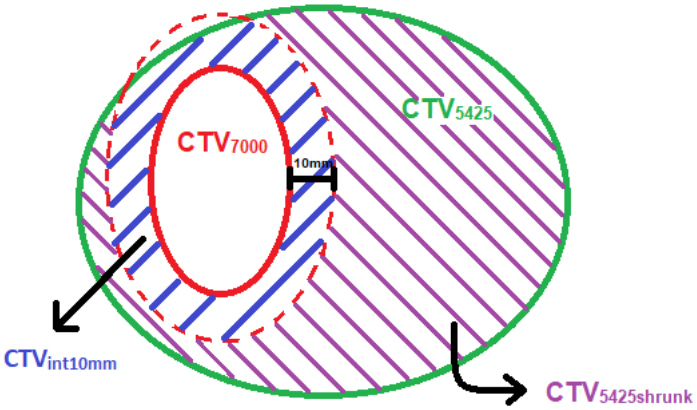


Figure 4.2.1: Schematic overview of CTV_{7000}, $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{5425\text{shrunk}}$.

In order to observe the impact of dose escalation on treatment plan quality, two distinct experiments are constructed, which are compared to a baseline plan. The baseline plan is a treatment plan generated using the conventional planning method in Erasmus-iCycle as described in Section 2.5.1. In the baseline plan, the maximum allowable dose to CTV_{7000}, $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{5425\text{shrunk}}$ are as listed in Table 4.2.1. For the first set of experiments, we test the theoretical potential of dose escalation by applying full dose escalation. In the second set of experiments, we apply partial dose escalation to test if clinically feasible treatment plans can be generated.

Table 4.2.1: Maximum allowable dose (Gy) to CTV_{7000}, $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{5425\text{shrunk}}$ in the baseline plan.

Structure	**Max dose (Gy)**
CTV_{7000}	74.2
$\text{CTV}_{\text{int10mm}}$	57.5
$\text{CTV}_{5425\text{shrunk}}$	57.5

To test if dose escalation indeed reduces dose to healthy tissues, the NTCPs of xerostomia and dysphagia for each test case are compared with respect to the baseline plan in the nominal treatment scenario on the repeat CT scan. The results on the planning CT scan are not shown since they are very similar. Additionally, to assess the robustness of the generated treatment

plans, the D98% CTV_{7000} and D98% CTV_{5425} of all generated treatment plans are evaluated using voxelwise minimum evaluation, as explained in Section 2.6.1. These values are then compared with respect to the baseline plan on the repeat CT scan. Furthermore, to gain an understanding of the results, we compare a representative dose distribution of each experiment to the dose distribution of the baseline plan on the planning CT scan in the voxelwise minimum scenario.

4.2.1 Full dose escalation

To observe the potential of dose escalation, we start by escalating the maximum allowable dose to CTV_{7000}, CTV_{int10mm} and $CTV_{\text{5425shrunk}}$ equally. Specifically, we alter the maximum allowable dose as described in Table 4.2.1 of all structures to 70, 77, 80, 90 and 105 Gy. In addition, we remove the maximum allowable dose constraint. Note that for all cases except for the first case, we relax the maximum dose constraint for all three structures. Whereas in the first case, we actually tighten the maximum dose constraint to CTV_{7000} but not to the other CTV structures.

4.2.2 Partial dose escalation

In the second set of experiments, the aim is to investigate the effect of increasing the dose in CTV_{7000} on one hand and both $CTV_{\text{5425shrunk}}$ and CTV_{int10mm} on the other hand. This is done by increasing the maximum allowable dose of the respective structures by 5, 10, 15 and 20 %.

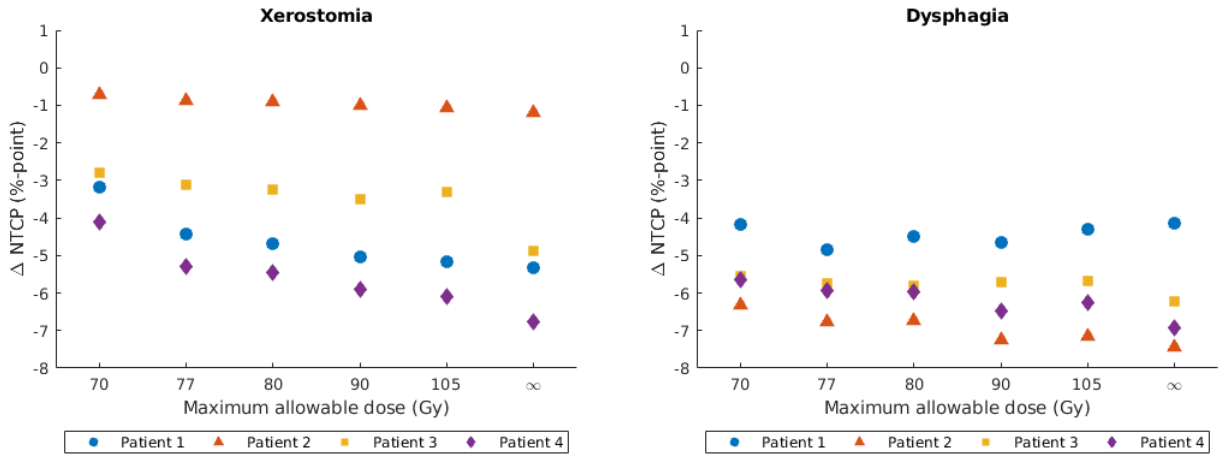
Originally, we only planned on increasing the dose in CTV_{7000}. However, we received poor results. In order to better understand these results, we also investigate increasing the maximum allowable dose in $CTV_{\text{5425shrunk}}$ and CTV_{int10mm}. Note that, according to conversations with radiation oncologists, it is not desirable to increase the maximum dose in $CTV_{\text{5425shrunk}}$ and CTV_{int10mm} as it increases the risk of fibrosis and edema. Note that there are no NTCP models available for these side effects. Hence, this additional experiment is mostly performed to gain a better understanding of the method.

4.3 Results

First we show the results of the full dose escalation experiments in Section 4.3.1. Subsequently, in Section 4.3.2 the results of the partial dose escalation experiments are presented. Furthermore, in Appendix C, additional results regarding the impact of dose escalation on the OARs and CTV are provided.

4.3.1 Full dose escalation

In Figure 4.3.2, the difference in NTCPs (%-point) when increasing the maximum allowable dose to CTV_{7000}, CTV_{int10mm} and $CTV_{\text{5425shrunk}}$ with respect to the baseline plan, is shown. Figure 4.3.2a shows the results for xerostomia. Similarly, Figure 4.3.2b shows the results for dysphagia. We observe a decrease in NTCPs by increasing the maximum allowable dose for all patients compared to the baseline plan. In general, higher maximum allowable doses result in lower NTCPs. The reduction in NTCP ranges from 0.7 up to 7.5%-point. However, for patient 1, the reduction in NTCP for dysphagia decreases with maximum allowable dose to the CTV structures. For this patient, a significant portion of both the oral cavity and the superior pharyngeal constrictor muscle is inside CTV_{7000}, CTV_{int10mm} and $CTV_{\text{5425shrunk}}$, as shown in Figure 4.3.4. As a result, by increasing the maximum allowable dose to these CTV structures, the dose to the oral cavity and the superior pharyngeal muscle constrictor increases. This, in turn, increases the NTCP of dysphagia.

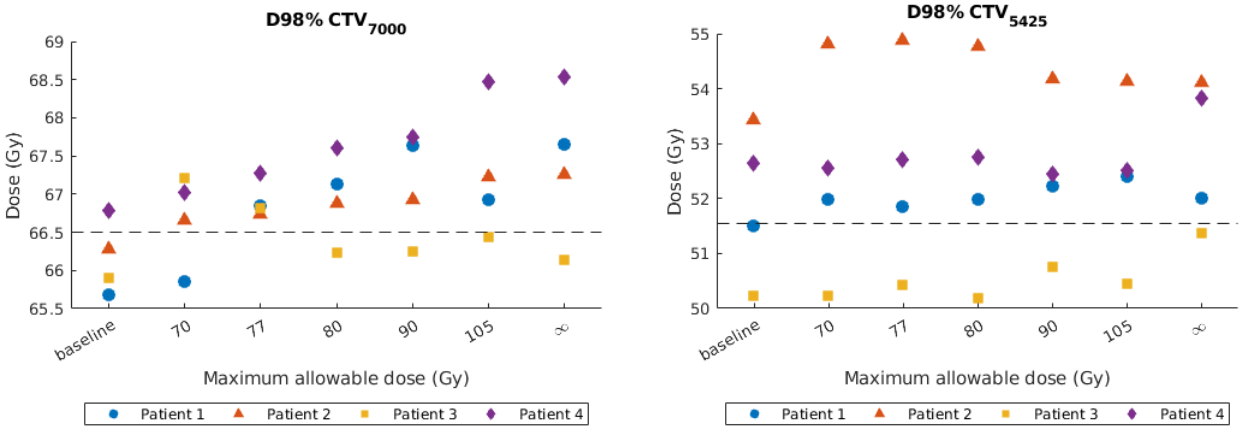


(a) Difference (%-point) in NTCP of xerostomia.

(b) Difference (%-point) in NTCP of dysphagia.

Figure 4.3.2: Difference (%-point) in NTCPs when applying full dose escalation with respect to the baseline plan on the repeat CT scan.

Figure 4.3.3 shows the D98% on the repeat CT scan. Figure 4.3.3a shows the results for CTV_{7000} and Figure 4.3.3b shows the results for CTV_{5425}. We observe mainly for CTV_{7000} an increase in D98% when full dose escalation is applied. An increase of up to approximately 2 Gy can be obtained. Hence, robustness of the plans against anatomical changes is increased.



(a) D98% CTV_{7000} (Gy).

(b) D98% CTV_{5425} (Gy).

Figure 4.3.3: D98% (Gy) on the repeat CT scan for the baseline plan and full dose escalation treatment plans.

Finally, in Figure 4.3.4, a representative example of the difference in dose distribution between the baseline plan and a full dose escalation treatment plan is shown for patient 1. Here, positive values indicate higher delivered dose when compared to the baseline plan. In this case, the maximum allowable dose constraint is completely removed. We observe that the delivered dose mostly increases inside CTV_{7000}, $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{5425\text{shrunk}}$. At most, the increase inside the CTV is 46 Gy when compared to the baseline plan. In the parotids, a reduction of around 25 Gy can be achieved. However, some regions outside CTV_{7000}, $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{5425\text{shrunk}}$ also experience an increased dose delivery. For instance, the dose on the top right (of the image) increases. However, it is important to note that this region is not assigned a maximum dose constraint during optimization. Furthermore, we observe that almost everywhere outside

CTV_{7000}, $CTV_{int10mm}$ and $CTV_{5425shrunk}$, the dose is reduced. Especially in the surrounding OARs, the delivered dose is reduced.

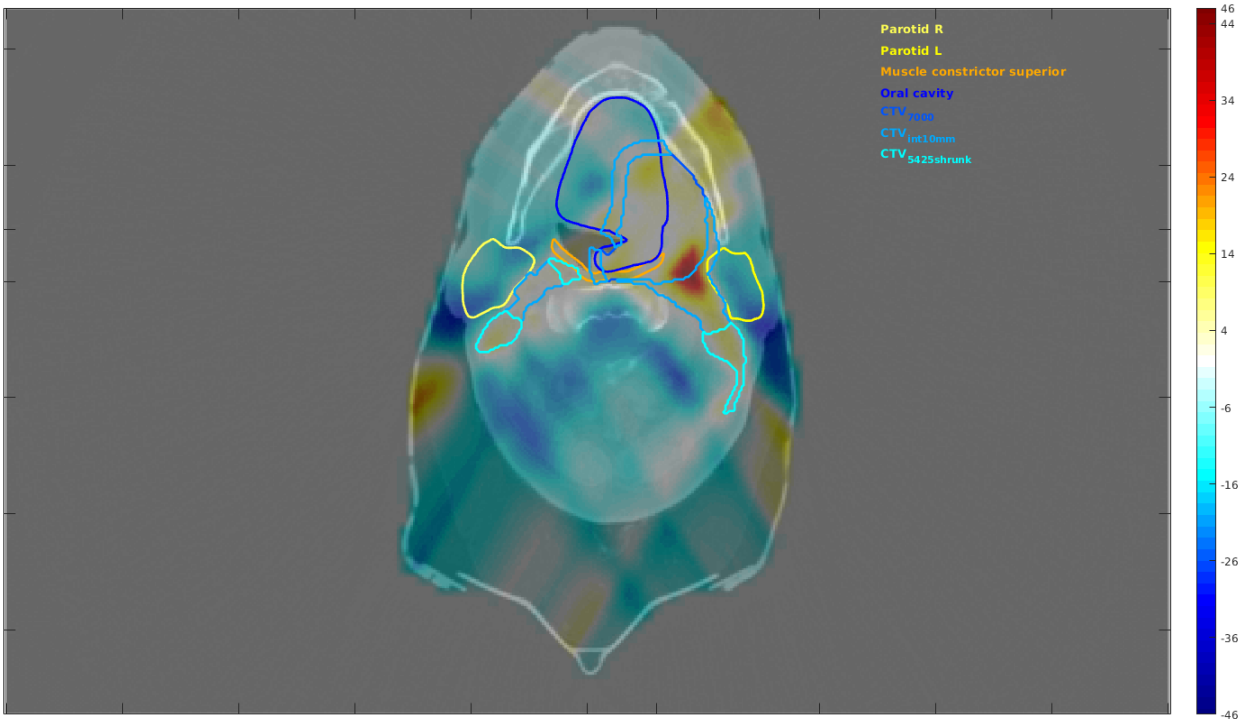


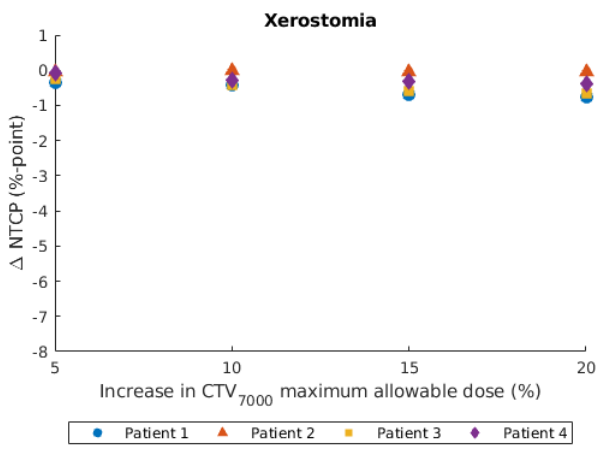
Figure 4.3.4: Difference (Gy) in dose distribution when full dose escalation is applied with respect to the baseline plan.

4.3.2 Partial dose escalation

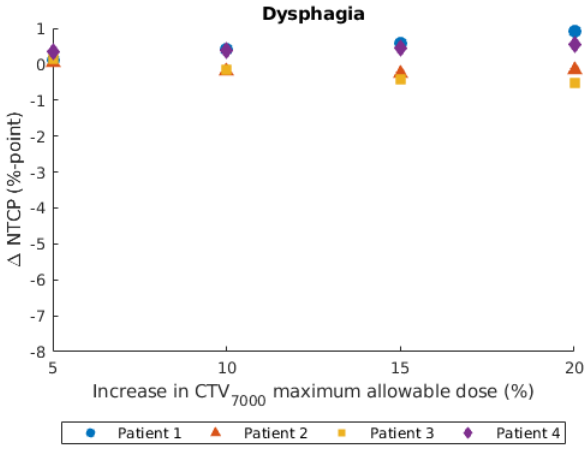
First, we present the results when the maximum allowable dose to only CTV_{7000} is increased and the maximum allowable dose to $CTV_{int10mm}$ and $CTV_{shrunk5425}$ are as in the baseline plan. Then, we present the results when the maximum allowable dose to both $CTV_{int10mm}$ and $CTV_{shrunk5425}$ is increased and the maximum allowable dose to CTV_{7000} remains as in the baseline plan.

Partial dose escalation of CTV_{7000}

Figure 4.3.5 shows the difference in NTCP (%-point) on the repeat CT scan when the maximum allowable dose constraint in CTV_{7000} is relaxed and the maximum allowable dose in $CTV_{5425shrunk}$ and $CTV_{int10mm}$ are as in the baseline plan. Figure 4.3.5a shows the results for xerostomia and Figure 4.3.5b shows the results for dysphagia. Here, the increase in maximum allowable dose is represented in percentage of the conventional maximum allowable dose (Table 4.2.1). In Figure 4.3.5, we observe that increasing the maximum allowable dose to CTV_{7000} results in a decrease in NTCP for xerostomia of up to 0.8%-point. For dysphagia, a decrease up to 0.5%-point in NTCP is obtained for two out of four patients. For the other two patients, an increase up to 0.9%-point in NTCP is obtained as higher maximum dose to CTV_{7000} is allowed.



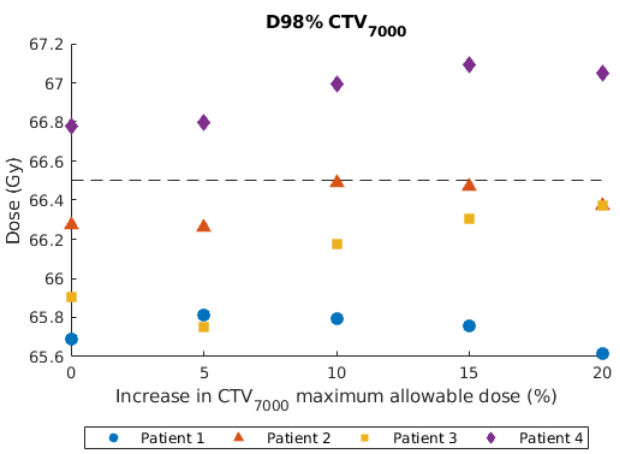
(a) Difference (%-point) in NTCP of xerostomia.



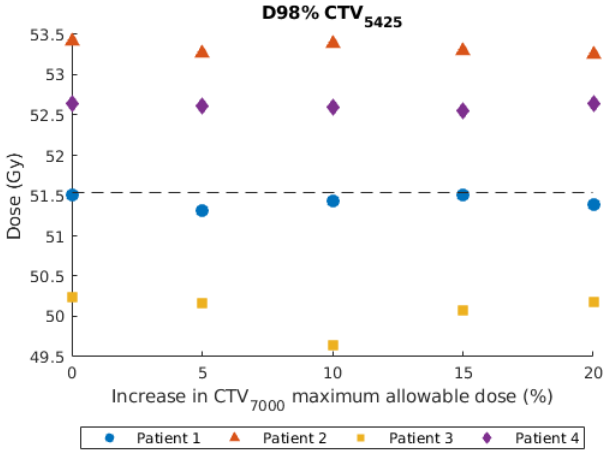
(b) Difference (%-point) in NTCP of dysphagia.

Figure 4.3.5: Difference (%-point) in NTCPs obtained by increasing the maximum allowable dose to CTV_{7000} with respect to the baseline plan, on the repeat CT scan.

In Figure 4.3.6, we show the results for D98% on the repeat CT scan. Figure 4.3.6a shows the results for CTV_{7000}. Similarly, 4.3.6b shows the results for CTV_{5425}. We see that generally the D98% CTV_{7000} increases with increasing maximum allowable dose. The maximum obtained increase is 0.5 Gy. The D98% CTV_{5425} remains relatively constant as higher maximum dose to CTV_{7000} is allowed.



(a) D98% CTV_{7000} (Gy).



(b) D98% CTV_{5425} (Gy).

Figure 4.3.6: D98% (Gy) on the repeat CT scan for the baseline plan and when increasing the maximum allowable dose to CTV_{7000}.

Finally, Figure 4.3.7 shows a representative difference in dose distribution between the baseline plan and the treatment plan by applying partial dose escalation for patient 1. In this case, the maximum allowable dose to CTV_{7000} is increased by 20%. In Figure 4.3.7, positive values indicate regions where the delivered dose is higher compared to the baseline plan. We see that by increasing the maximum allowable dose to CTV_{7000}, the delivered dose to CTV_{7000} increases up to 13 Gy. Outside CTV_{7000}, many parts receive less dose. However, there also exist many parts that receive higher dose.

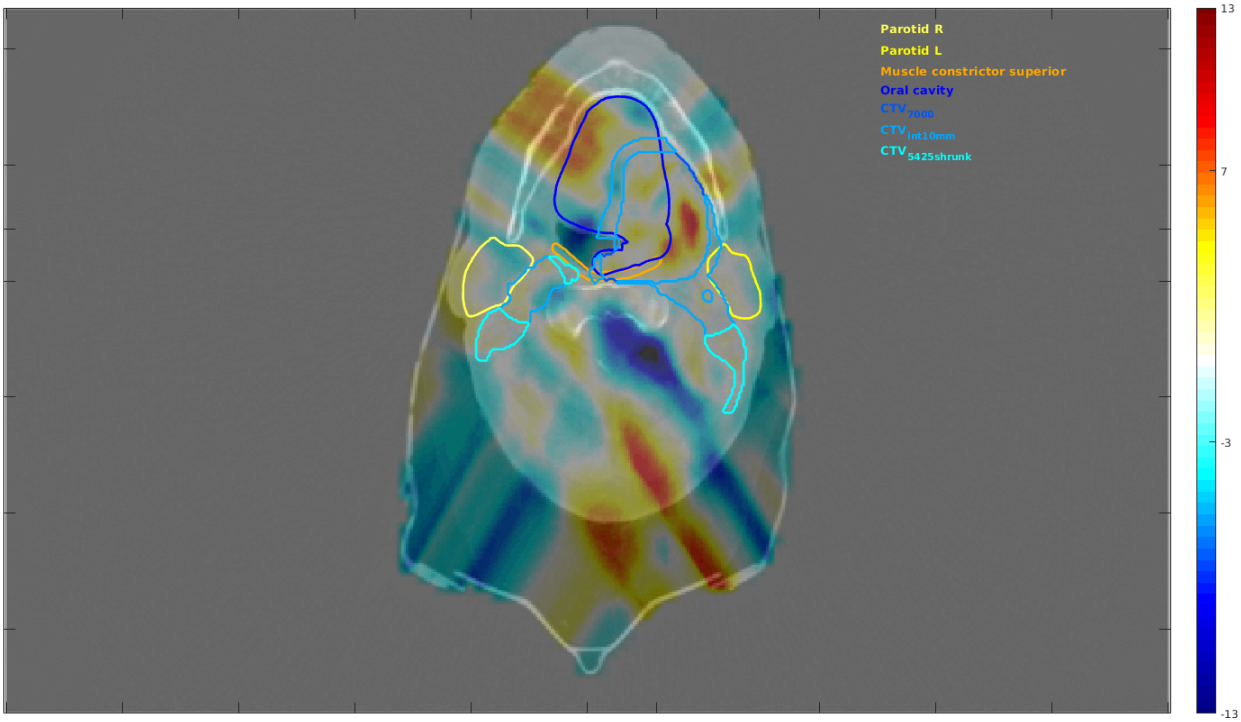
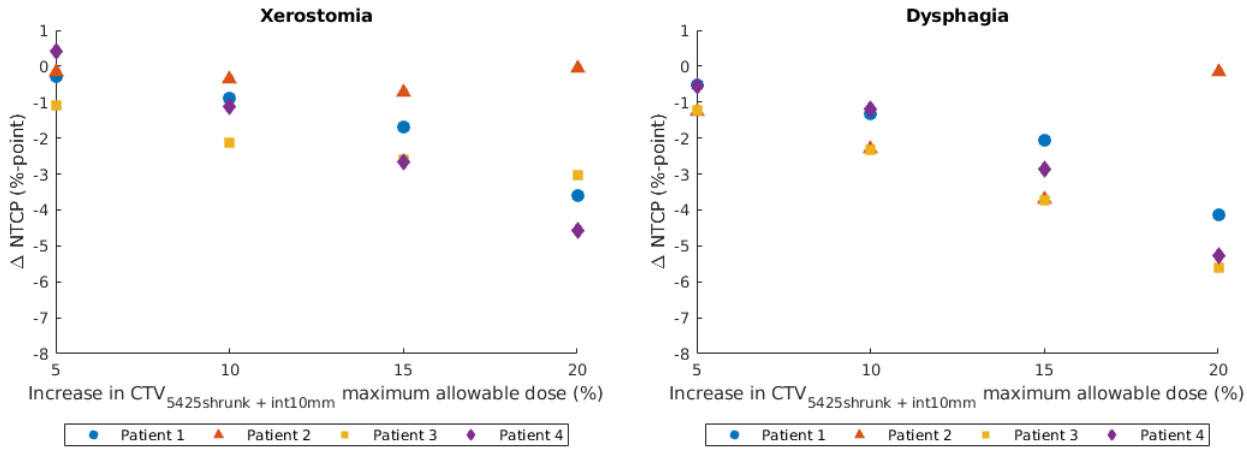


Figure 4.3.7: Difference (Gy) in dose distribution when partial dose escalation is applied with respect to the baseline plan.

Partial dose escalation of $\mathbf{CTV_{int10mm}}$ and $\mathbf{CTV_{shrunk5425}}$

Next, the results for the final experiment are shown. In Figure 4.3.8, the difference in NTCP (%-point) on the repeat CT scan is shown when the maximum dose constraint in $CTV_{5425_{shrunk}}$ and $CTV_{int10mm}$ is relaxed and the maximum dose constraint in CTV_{7000} remains as in the baseline plan. Figure 4.3.8a shows the results for xerostomia. Similarly, Figure 4.3.8b shows the results for dysphagia. We find that the NTCP for both xerostomia and dysphagia decrease with increasing maximum allowable dose. Reduction of up to 5%-point and 6%-point is obtained for xerostomia and dysphagia, respectively. However, for patient 2, increasing the maximum allowable dose by 20%, gives similar NTCP values as in the baseline.

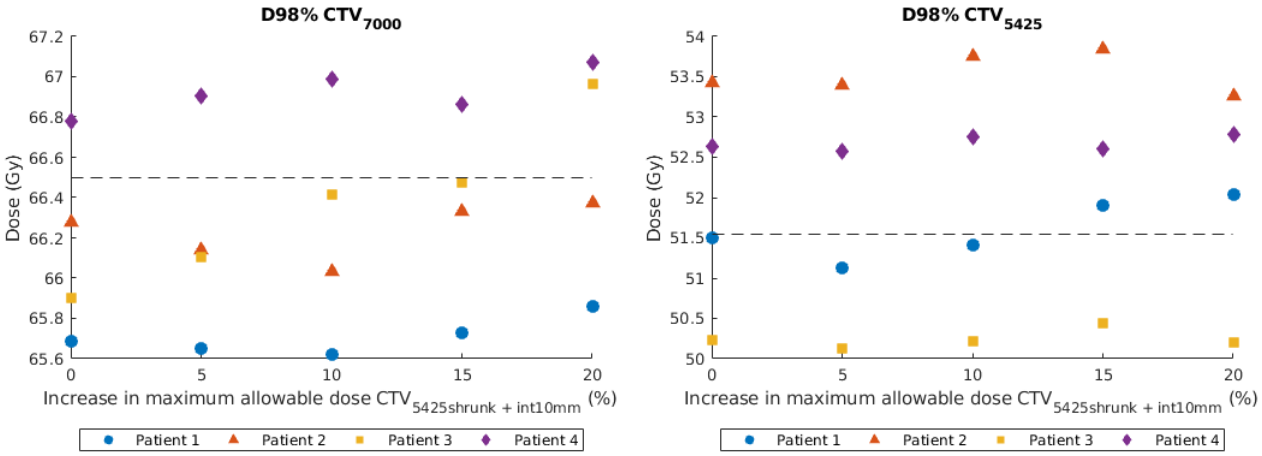


(a) Difference (%-point) in NTCP of xerostomia.

(b) Difference (%-point) in NTCP of dysphagia.

Figure 4.3.8: Difference (%-point) in NTCPs obtained by increasing the maximum allowable dose to $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{\text{5425shrunk}}$ with respect to the baseline plan, on the repeat CT.

In Figure 4.3.9, we show the D98% in both CTV_{7000} and CTV_{5425} on the repeat CT scan. In Figure 4.3.9a, D98% CTV_{7000} is shown. Similarly, in Figure 4.3.9b, D98% CTV_{5425} is shown. We observe that D98% CTV_{7000} is generally highest when increasing the maximum allowable dose by 20%. The increase is up to 1 Gy. Furthermore, the increase in maximum allowable dose does not appear to have a significant impact on the D98% CTV_{5425}.



(a) D98% CTV_{7000} (Gy).

(b) D98% CTV_{5425} (Gy).

Figure 4.3.9: D98% (Gy) on the repeat CT scan for the baseline plan and when increasing the maximum allowable dose to $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{\text{5425shrunk}}$.

In Figure 4.3.10, a representative difference in dose distribution between the baseline plan and the plan obtained with partial dose escalation is shown for patient 1. In the case shown here, the maximum allowable dose to $\text{CTV}_{\text{5425shrunk}}$ and $\text{CTV}_{\text{int10mm}}$ is increased by 20%. Here, positive values indicate regions where more dose is delivered in the partial dose plan with respect to the baseline plan. We observe that inside $\text{CTV}_{\text{5425shrunk}}$ and $\text{CTV}_{\text{int10mm}}$ the delivered dose, increases up to 5 Gy. We also observe that in many regions outside $\text{CTV}_{\text{5425shrunk}}$ and $\text{CTV}_{\text{int10mm}}$, reduced dose delivery of around 25 Gy is obtained when applying partial dose escalation.

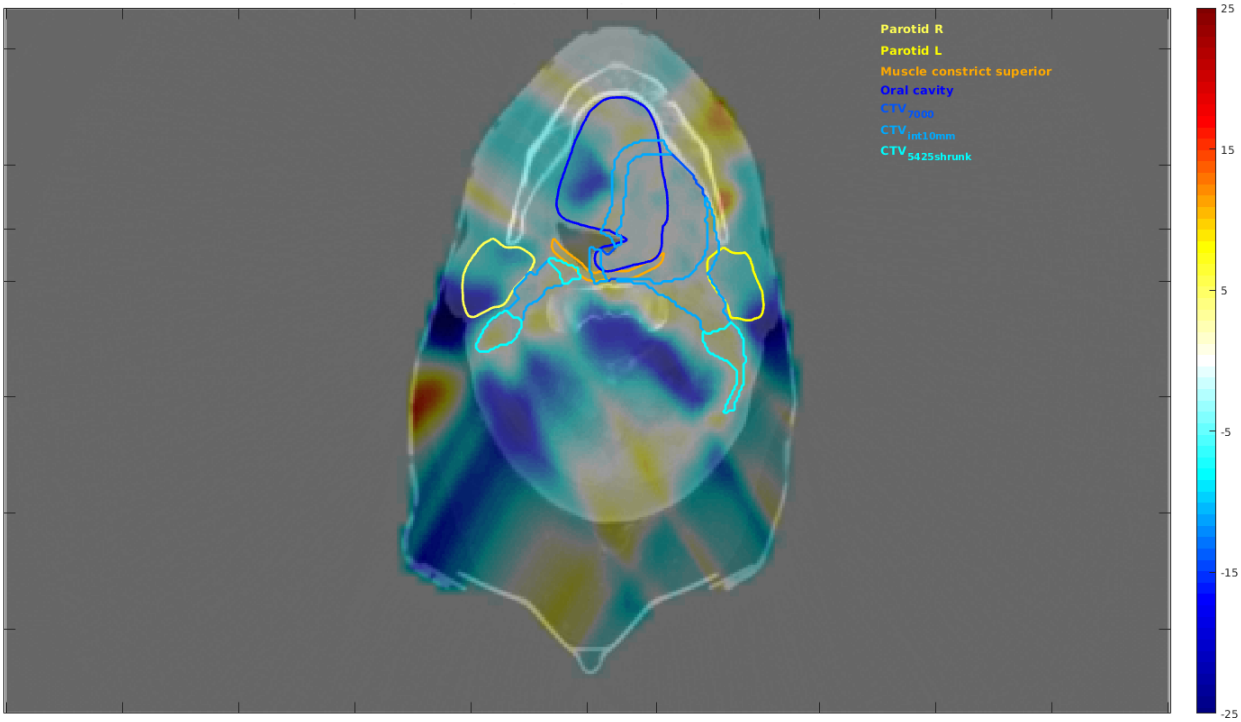


Figure 4.3.10: Difference (Gy) in dose distribution when partial dose escalation is applied with respect to the baseline plan.

4.4 Discussion

In this study, the impact of increasing the maximum allowable dose to different parts of the CTV is investigated. The objective is to preserve CTV coverage and reduce NTCPs. The results are evaluated on the repeat CT scans. To still achieve equal target coverage on the planning CT scan, all treatment plans are scaled to meet the clinical target coverage constraints on 29 treatment scenarios.

When increasing the maximum allowable dose in CTV_{7000}, $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{5425\text{shrunk}}$ simultaneously, a considerable reduction in NTCPs on both the planning and repeat CT scan is obtained. Hence, this experiment demonstrates the potential of dose escalation in reducing the NTCPs. In addition, increasing the maximum allowable dose in only $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{5425\text{shrunk}}$ shows significant NTCP reduction. However, by applying dose escalation only to CTV_{7000}, poor results are obtained as the decrease in NTCPs is much lower than obtained by increasing the maximum allowable dose in $\text{CTV}_{\text{shrunk5425}}$ and $\text{CTV}_{\text{int10mm}}$. From this, we find that the primary contributing factor in reducing NTCPs is allowing a higher maximum dose to $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{5425\text{shrunk}}$.

Furthermore, improved robustness is obtained on the repeat CT scan when applying full dose escalation with respect to the baseline plan. In contrast, when applying partial dose escalation, robustness is more comparable to the baseline plan.

In all treatment plans, an overall reduction in dose delivered to tissues surrounding the CTV is obtained when applying dose escalation to either CTV_{7000}, $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{5425\text{shrunk}}$ or $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{5425\text{shrunk}}$. These results are in line with the anticipated outcomes of dose escalation. However, the reduction is again less when only increasing the maximum allowable dose to CTV_{7000}.

In conclusion, dose escalation is an efficient method to reduce the dose to surrounding healthy tissues and maintaining robustness. The most significant contribution to these results is obtained by increasing the maximum allowable dose to $CTV_{int10mm}$ and $CTV_{5425shrunk}$. However, as this increases the risk of edema and fibrosis, it is best to refrain from increasing the dose to these particular CTV parts. For further research, it would be interesting to develop NTCP models for edema and fibrosis. This enables fair assessment of the risk of side effects.

Furthermore, by applying dose escalation only to CTV_{7000}, a small reduction in NTCP can be obtained for similar robustness. Therefore, a suggestion for further research, is to investigate if there exists a specific group of patients for whom escalating the dose to only CTV_{7000} results in decent NTCP reduction.

Another suggestion for further research focuses on the relation between findings in Sections 3 and 4. In Section 3, we found higher D2% and lower NTCPs when optimizing on 13 treatment scenarios and performing a scaling. To show that increased D2% results in improved robustness and reduced NTCPs, one could test if for the same D2% CTV_{7000} that is obtained when planning on 13 treatment scenarios, similar reduction in NTCP and improvement in CTV coverage can be obtained.

5 Variance optimization

Generation of robust treatment plans is time consuming as it involves optimization of many treatment scenarios. To reduce the optimization time, while maintaining robustness, variance optimization is applied [28]. In variance optimization, rather than constraining the minimum target dose in multiple treatment scenarios, only the minimum target dose in the nominal treatment scenario is constrained. In order to preserve practical robustness, an additional objective is added. This objective consists of two terms. The first term is used to minimize the variance in delivered dose of a voxel over all treatment scenarios. The second term minimizes the difference between the prescribed dose and mean dose of the voxels over all treatment scenarios. This way, the delivered dose in a voxel inside the CTV should be similar and sufficient in every treatment scenario.

In this section the feasibility of variance optimization is investigated. To this end, we first derive the mathematical formula to the optimization problem in Section 5.1. Then, treatment plans are generated with and without variance optimization and the results of both types of treatment plans are presented in Section 5.2. Finally, the results are discussed in Section 5.3.

5.1 Method

With variance optimization, we aim to find pencil beam intensities $\mathbf{x}$ such that the delivered dose $\mathbf{d}(\mathbf{x}) \in \mathbb{R}^n$ is close to the prescribed dose $\mathbf{d}^* \in \mathbb{R}^n$. To this end, we consider the following least squares objective function:

$$F(\mathbf{d}(\mathbf{x})) = (\mathbf{d}(\mathbf{x}) - \mathbf{d}^*)^T(\mathbf{d}(\mathbf{x}) - \mathbf{d}^*). \tag{5.1}$$

Let $\mathbf{d}(\mathbf{x})$ be a random variable with $K = \{\mathbf{d}_1(\mathbf{x}), ..., \mathbf{d}_K(\mathbf{x})\}$ the set of possible outcomes. Recall from Section 2.5 that $\mathbf{d}_k(\mathbf{x}) = D_k\mathbf{x}$ is the dose deposited in treatment scenario k. Then, the expectation of $\mathbf{d}(\mathbf{x})$ is given by

$$\begin{aligned}\bar{\mathbf{d}}(\mathbf{x}) &= \frac{1}{|K|}\sum_{k\in K}\mathbf{d}_k(\mathbf{x}) \\ &= \frac{1}{|K|}\sum_{k\in K}D_k\mathbf{x} \\ &= \bar{D}\mathbf{x}.\end{aligned}$$

Here, the mean dose deposition matrix $\bar{D}$ is the element wise average of the dose deposition matrices of all treatment scenarios.

To make sure Equation (5.1) is minimized over all possible outcomes of $\mathbf{d}(\mathbf{x})$, we want to minimize the expectation of the least squares objective function. This expectation is determined as follows:

$$\begin{aligned}\mathbb{E}\left[F(\mathbf{d}(\mathbf{x}))\right] &= \mathbb{E}\left[(\mathbf{d}(\mathbf{x}) - \mathbf{d}^*)^T(\mathbf{d}(\mathbf{x}) - \mathbf{d}^*)\right] \\ &= \mathbb{E}\left[\sum_{i=1}^{n}(d_i - d_i^*)\cdot(d_i - d_i^*)\right] \\ &= \sum_{i=1}^{n}\mathbb{E}\left[(d_i - d_i^*)\cdot(d_i - d_i^*)\right] &(5.2)\\ &= \sum_{i=1}^{n}\left(C_{ii} + \mathbb{E}\left[d_i - d_i^*\right]\mathbb{E}[d_i - d_i^*]\right) &(5.3)\\ &= \mathrm{tr}(C) + (\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)^T(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*), &(5.4)\end{aligned}$$

with C the covariance matrix of $\mathbf{d}(\mathbf{x}) - \mathbf{d}^*$. We derived Equation (5.3) from Equation (5.2) by applying the covariance formula which states that for two random variables X, Y

$$\mathbb{E}[XY^T] = cov(X, Y) + \mathbb{E}[X]\mathbb{E}[Y^T]. \tag{5.5}$$

Next, we separately compute $\mathrm{tr}(C)$ and $(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)^T(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)$ to obtain an expression for $\mathbb{E}\left[F(\mathbf{d}(\mathbf{x}))\right]$ in terms of D_k and $\mathbf{x}$. We begin with the computation of $\mathrm{tr}(C)$.

Computation of tr(C)

We use the formula presented in Equation (5.5) to obtain an expression for the covariance matrix C. Since we have $X = Y = \mathbf{d}(\mathbf{x}) - \mathbf{d}^*$, we obtain

$$\mathbb{E}\left[(\mathbf{d}(\mathbf{x}) - \mathbf{d}^*)(\mathbf{d}(\mathbf{x}) - \mathbf{d}^*)^T\right] = C + \mathbb{E}\left[\mathbf{d}(\mathbf{x}) - \mathbf{d}^*\right]\mathbb{E}\left[\mathbf{d}(\mathbf{x}) - \mathbf{d}^*\right]^T.$$

Then, C can be computed as follows:

$$\begin{aligned}
C &= \mathbb{E}\left[(\mathbf{d}(\mathbf{x}) - \mathbf{d}^*)(\mathbf{d}(\mathbf{x}) - \mathbf{d}^*)^T\right] - \mathbb{E}\left[\mathbf{d}(\mathbf{x}) - \mathbf{d}^*\right]\mathbb{E}[\mathbf{d}(\mathbf{x}) - \mathbf{d}^*]^T \\
&= \mathbb{E}\left[\left(\mathbf{d}(\mathbf{x}) - \mathbf{d}^*\right)\left(\mathbf{d}(\mathbf{x}) - \mathbf{d}^*\right)^T\right] - \left(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*\right)\left(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*\right)^T \\
&= \mathbb{E}\left[\mathbf{d}(\mathbf{x})\mathbf{d}(\mathbf{x})^T\right] - \mathbb{E}\left[\mathbf{d}(\mathbf{x})\mathbf{d}^{*T}\right] - \mathbb{E}\left[\mathbf{d}^*\mathbf{d}(\mathbf{x})^T\right] + \mathbb{E}\left[\mathbf{d}^*\mathbf{d}^{*T}\right] \\
&\qquad - \left(\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right]\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right]^T - \mathbb{E}\left[\mathbf{d}(\mathbf{x})\right]\mathbf{d}^{*T} - \mathbf{d}^*\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right]^T + \mathbf{d}^*\mathbf{d}^{*T}\right) \\
&= \mathbb{E}\left[\mathbf{d}(\mathbf{x})\mathbf{d}(\mathbf{x})^T\right] - \mathbb{E}\left[\mathbf{d}(\mathbf{x})\right]\mathbf{d}^{*T} - \mathbf{d}^*\mathbb{E}\left[\mathbf{d}(\mathbf{x})^T\right] + \mathbf{d}^*\mathbf{d}^{*T} \\
&\qquad - \left(\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right]\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right]^T - \mathbb{E}\left[\mathbf{d}(\mathbf{x})\right]\mathbf{d}^{*T} - \mathbf{d}^*\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right]^T + \mathbf{d}^*\mathbf{d}^{*T}\right) \\
&= \mathbb{E}\left[\mathbf{d}(\mathbf{x})\mathbf{d}(\mathbf{x})^T\right] - \mathbb{E}\left[\mathbf{d}(\mathbf{x})\right]\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right]^T \\
&= \mathbb{E}\left[D\mathbf{x}\mathbf{x}^T D^T\right] - \mathbb{E}\left[D\mathbf{x}\right]\mathbb{E}\left[D\mathbf{x}\right]^T \\
&= \mathbb{E}\left[D\mathbf{x}\mathbf{x}^T D^T\right] - \mathbb{E}\left[D\right]\mathbf{x}\mathbf{x}^T\mathbb{E}\left[D\right]^T \qquad (5.6) \\
&= \frac{1}{|K|}\sum_{k\in K} D_k\mathbf{x}\mathbf{x}^T D_k^T - \bar{D}\mathbf{x}\mathbf{x}^T\bar{D}^T, \qquad (5.7)
\end{aligned}$$

where in the last step we used

$$\mathbb{E}\left[D\mathbf{x}\mathbf{x}^T D^T\right] = \frac{1}{|K|}\sum_{k\in K} D_k\mathbf{x}\mathbf{x}^T D_k^T.$$

The verification of this step is provided in Appendix D. Now that we have an expression for the covariance matrix C, we can compute $\mathrm{tr}(C)$. We find that

$$\begin{aligned}
\mathrm{tr}(C) &= \mathrm{tr}\left(\frac{1}{|K|}\sum_{k\in K} D_k\mathbf{x}\mathbf{x}^T D_k^T - \bar{D}\mathbf{x}\mathbf{x}^T\bar{D}^T\right) \\
&= \frac{1}{|K|}\mathrm{tr}\left(\sum_{k\in K} D_k\mathbf{x}\mathbf{x}^T D_k^T\right) - \mathrm{tr}\left(\bar{D}\mathbf{x}\mathbf{x}^T\bar{D}^T\right) \\
&= \frac{1}{|K|}\sum_{k\in K}\mathrm{tr}\left(D_k\mathbf{x}\mathbf{x}^T D_k^T\right) - \mathrm{tr}\left(\bar{D}\mathbf{x}\mathbf{x}^T\bar{D}^T\right) \qquad (5.8) \\
&= \frac{1}{|K|}\sum_{k\in K}\mathbf{x}^T D_k^T D_k\mathbf{x} - \mathbf{x}^T\bar{D}^T\bar{D}\mathbf{x} \qquad (5.9) \\
&= \frac{1}{|K|}\sum_{k\in K}\mathbf{x}^T D_k^T D_k\mathbf{x} - \frac{1}{|K|^2}\sum_{k\in K}\mathbf{x}^T {D_k}^T\sum_{k\in K} D_k\mathbf{x}. \qquad (5.10)
\end{aligned}$$

Here, we derived Equation (5.9) from Equation (5.8) by applying the cyclic property of the trace.

Computation of $(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)^T(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)$

Next we compute an expression in terms of D_k and $\mathbf{x}$ for the second term of Equation (5.4). We find that:

$$\begin{aligned}(\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right] - \mathbf{d}^*)^T \left(\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right] - \mathbf{d}^*\right) &= \mathbb{E}\left[\mathbf{d}(\mathbf{x})\right]^T \mathbb{E}\left[\mathbf{d}(\mathbf{x})\right] - \mathbf{d}^{*T}\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right] - \mathbb{E}\left[\mathbf{d}(\mathbf{x})\right]^T \mathbf{d}^* + \mathbf{d}^{*T}\mathbf{d}^* \\ &= \mathbb{E}\left[\mathbf{d}(\mathbf{x})\right]^T \mathbb{E}\left[\mathbf{d}(\mathbf{x})\right] - 2\mathbf{d}^{*T}\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right] + \mathbf{d}^{*T}\mathbf{d}^* \\ &= \frac{1}{|K|^2}\sum_{k\in K}\mathbf{x}^T D_k^T \sum_{k\in K} D_k\mathbf{x} - \frac{2}{|K|}\mathbf{d}^{*T}\sum_{k\in K} D_k\mathbf{x} + \mathbf{d}^{*T}\mathbf{d}^*.\end{aligned}$$

Computation of $\mathbb{E}\left[F\left(\mathbf{d}(\mathbf{x})\right)\right]$

Now that we have an expression for $\mathrm{tr}\left(C\right)$ and $(\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right] - \mathbf{d}^*)^T \left(\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right] - \mathbf{d}^*\right)$, we can write $\mathbb{E}\left[F\left(\mathbf{d}(\mathbf{x})\right)\right]$ in terms of D_k and $\mathbf{x}$. Hence, we find that the expectation of the least squares objective function is equivalent to

$$\begin{aligned}\mathbb{E}\left[F\left(\mathbf{d}(\mathbf{x})\right)\right] &= \mathrm{tr}\left(C\right) + (\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right] - \mathbf{d}^*)^T \left(\mathbb{E}\left[\mathbf{d}(\mathbf{x})\right] - \mathbf{d}^*\right) \\ &= \frac{1}{|K|}\sum_{k\in K}\mathbf{x}^T D_k^T D_k\mathbf{x} - \frac{1}{|K|^2}\sum_{k\in K}\mathbf{x}^T {D_k}^T \sum_{k\in K} D_k\mathbf{x} + \frac{1}{|K|^2}\sum_{k\in K}\mathbf{x}^T D_k^T \sum_{k\in K} D_k\mathbf{x} \\ &\quad - \frac{2}{|K|}\mathbf{d}^{*T}\sum_{k\in K} D_k\mathbf{x} + \mathbf{d}^{*T}\mathbf{d}^* \\ &= \frac{1}{|K|}\sum_{k\in K}\mathbf{x}^T D_k^T D_k\mathbf{x} - \frac{2}{|K|}\mathbf{d}^{*T}\sum_{k\in K} D_k\mathbf{x} + \mathbf{d}^{*T}\mathbf{d}^* \\ &= \mathbf{x}^T\left(\frac{1}{|K|}\sum_{k\in K} D_k^T D_k\right)\mathbf{x} - \frac{2}{|K|}\mathbf{d}^{*T}\sum_{k\in K} D_k\mathbf{x} + \mathbf{d}^{*T}\mathbf{d}^*.\end{aligned}$$

Note that we can write this quadratic optimization problem into the canonical form:

$$\mathbb{E}\left[F(\mathbf{d}(\mathbf{x}))\right] = \frac{1}{2}\mathbf{x}^T A\mathbf{x} + \mathbf{b}^T\mathbf{x} + c. \tag{5.11}$$

Here,

$$\begin{aligned}A &= 2\cdot\frac{1}{|K|}\sum_{k\in K} D_k^T D_k, \\ \mathbf{b} &= -\frac{2}{|K|}\left(\mathbf{d}^{*T}\sum_{k\in K} D_k\right)^T, \\ c &= \mathbf{d}^{*T}\mathbf{d}.\end{aligned}$$

The aim of variance optimization is to find $\mathbf{x}$ such that Equation (5.11) is minimized. Note that by minimizing $\mathrm{tr}(C)$, the variance of the dose to each voxel over different treatment scenarios is minimized. The second term, $(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)^T(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)$, is used to minimize the difference between the mean dose over all treatment scenarios and the prescribed dose.

Experimental setup

To investigate the feasibility of variance optimization, four sets of experiments are constructed which are compared to a baseline plan. The baseline plan is a treatment plan generated using the conventional robust planning method in Erasmus-iCycle, as described in Section 2.5.1. The baseline plan is optimized using 29 treatment scenarios. In the first set of experiments, we apply full variance optimization to minimize $\mathrm{tr}(C) + (\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)^T(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)$. For the second set of experiments, partial variance optimization is applied in which we only optimize $\mathrm{tr}(C)$. The

latter approach has also been proposed by [28]. For both experiments, we include 29 treatment scenarios.

As we received poor dosimetric results using both full and partial variance optimization, two additional sets of experiments are constructed. The third set is implemented to further exploit the potential of partial variance optimization. In this third experiment, referred to as constrained partial variance optimization, we constrain the NTCPs obtained in the baseline plan and subsequently optimize $\text{tr}(C)$. This way, we test if similar CTV doses as in the baseline plan can be obtained for similar NTCPs. Finally, in the fourth set of experiments we apply modified full variance optimization. This method is used to analyze the results obtained from full variance optimization in order to gain a better understanding of full variance optimization. To this end, we optimize $\text{tr}(C) + \frac{|\text{tr}(C)|}{|\mathbb{E}[F(\mathbf{d}(\mathbf{x}))]|} \cdot (\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)^T(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)$. An overview of the terms that are optimized in each of the experiments, is provided in Table 5.1.1.

Table 5.1.1: Overview of terms to be optimized for in each experiment.

Experiment	**Term to be optimized**
Full variance optimization	$\text{tr}(C) + (\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)^T(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)$
Partial variance optimization	$\text{tr}(C)$
Constrained partial variance optimization	$\text{tr}(C)$
Modified full variance optimization	$\text{tr}(C) + \frac{\lvert\text{tr}(C)\rvert}{\lvert\mathbb{E}[F(\mathbf{d}(\mathbf{x}))]\rvert} \cdot (\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)^T(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)$

In all experiments, the method is applied to the voxels of CTV_{7000}, $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{5425\text{shrunk}}$ simultaneously (Figure 4.2.1). Hence, for each of these structures, an objective is added that minimizes (part of) Equation (5.11).

To investigate the impact of partial and full variance optimization on optimization time, we compare the optimization times with respect to the baseline plan. Additionally, to assess the dosimetric impact, we evaluate D2% CTV_{7000} in the voxelwise maximum scenario and the NTCPs in the nominal treatment scenario. All results are evaluated on the planning CT scan and compared with respect to the baseline plan. Note that the treatment plans are scaled such that D98% $\text{CTV}_{7000} \geq 66.5$ Gy and D98% $\text{CTV}_{5425} \geq 51.54$ Gy on the planning CT scan. Hence, we do not show D98% CTV_{7000} and D98% CTV_{5425} as they do not provide information on the difference between the methods in treatment plan quality.

5.2 Results

In Section 5.2.1 we present the optimization times for the baseline, full and partial variance optimization plans. The dosimetric results of the full and partial variance optimization experiments are presented in Section 5.2.2. Then, in Section 5.2.3 the dosimetric results of constrained partial variance optimization are presented. Finally, in Section 5.2.4 results of additional tests on the functioning of full variance optimization are presented.

5.2.1 Optimization time

The optimization times for the baseline, full and partial variance optimization plans are listed in Table 5.2.2. We observe that the optimization time reduces when applying either full or partial variance optimization. When applying partial variance optimization, reduction in optimization times of 85% to 99% are achieved compared to the baseline plan. We also observe some variability in optimization time between patients. The optimization time for patient 1 using full variance optimization is high when compared to the other patients. This is due to divergence during optimization of the treatment plan. As a result, part of the optimization process starts over at a new starting point which causes an increase in optimization time.

Table 5.2.2: Optimization time in minutes for baseline, partial and full variance optimization treatment plans.

Plan	**Optimization time (min)**			
	Patient 1	Patient 2	Patient 3	Patient 4
Baseline	804	686	456	693
Full variance optimization	120	33	29	26
Partial variance optimization	7	4	5	24

5.2.2 Full and partial variance optimization

In Figure 5.2.1, the D2% CTV_{7000} is shown for the baseline, full and partial variance optimization plans. Note that D2% CTV_{7000} should be below 74.90 Gy. This criterion is satisfied for both the baseline and partial variance optimization plans. However, when applying full variance optimization, we find that D2% CTV_{7000} always exceeds the 74.90 Gy threshold.

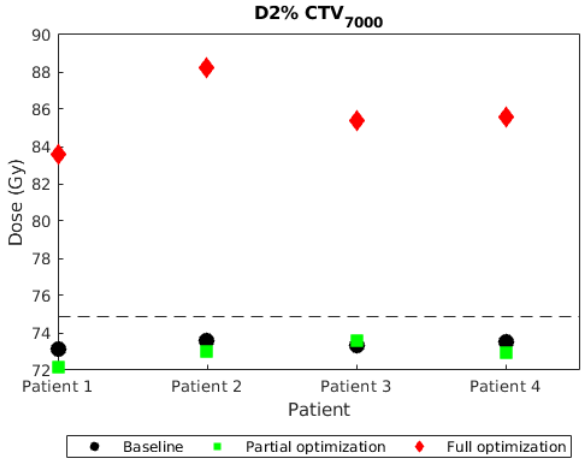
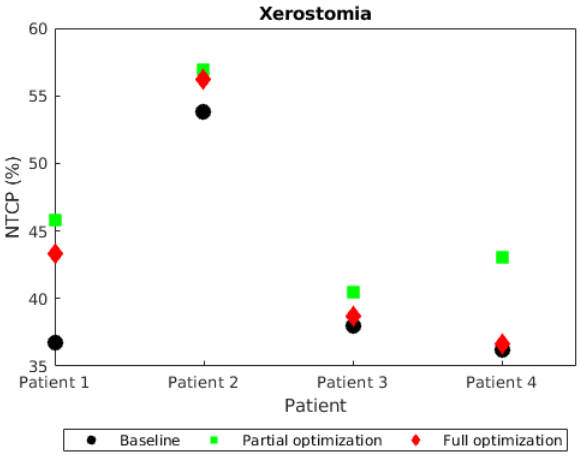
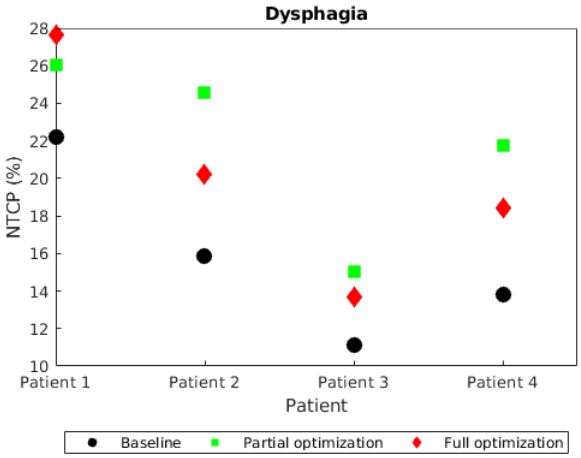


Figure 5.2.1: D2% CTV_{7000} in the baseline, full and partial variance optimization treatment plans.

In Figure 5.2.2, the NTCPs of xerostomia and dysphagia are shown for the baseline, full and partial variance optimization plans. Figure 5.2.2a shows the results for xerostomia. Similarly, Figure 5.2.2b shows the results for dysphagia. Both NTCPs increase in the partial and full variance treatment plan when compared to the baseline plan. The increase is most significant when using partial variance optimization where an increase of up to 9%-point is obtained in the risk of both xerostomia and dysphagia.



(a) NTCP Xerostomia.

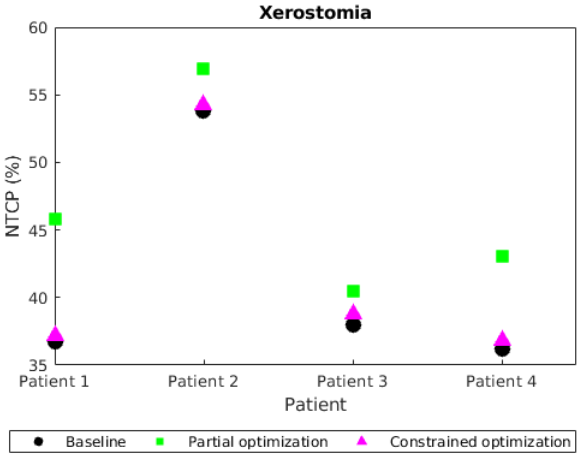


(b) NTCP Dysphagia.

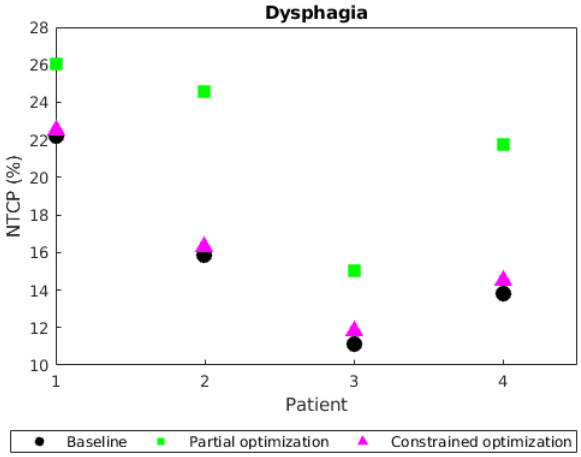
Figure 5.2.2: NTCPs for xerostomia and dysphagia in the baseline, full and partial variance optimization treatment plans.

5.2.3 Constrained partial variance optimization

In this experiment, we investigate if for similar CTV coverage, similar NTCPs can be obtained when compared to the baseline plan. To this end, we constrain the NTCPs obtained for the unscaled baseline plan. After optimization, the treatment plan is again scaled to fulfill the robustness criteria. In Figure 5.2.3, we show the NTCPs that are obtained after scaling of the treatment plans. The results for xerostomia and dysphagia are shown in Figures 5.2.3a and 5.2.3b, respectively. For reference, the results of the partial variance optimization are also shown. In general, the NTCPs in the baseline plan and constrained variance optimization plan are very similar. This indicates that for equal robustness, similar NTCPs can be obtained.



(a) NTCP for xerostomia.



(b) NTCP for dysphagia.

Figure 5.2.3: NTCPs for xerostomia and dysphagia in the baseline, partial and partial constrained variance optimization treatment plans.

In Figure 5.2.4, the D2% CTV_{7000} is shown for baseline, partial and partial constrained variance optimization plans. In all cases, we find that D2% $\text{CTV}_{7000} \leq 74.90$ Gy, as desired. Hence, for NTCPs similar to the baseline plan, suitable D2% CTV_{7000} is obtained.

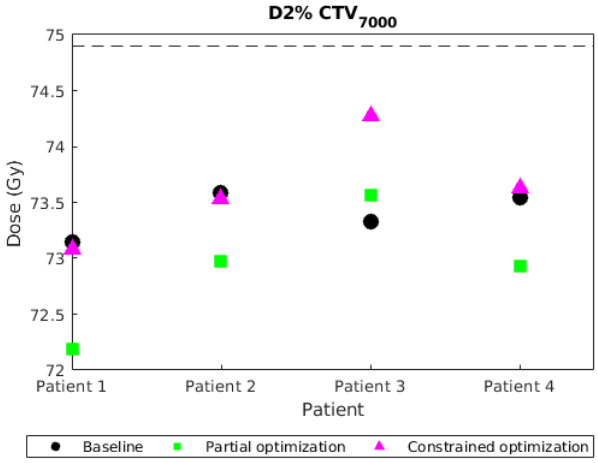


Figure 5.2.4: D2% CTV_{7000} in the baseline and partial constrained variance optimization treatment plans.

5.2.4 Modified full variance optimization

In Section 5.2.2 we obtained poor D2% CTV_{7000} results using full variance optimization. To better understand these results, we analyze the contribution of the individual terms of

$$\mathbb{E}\left[F(\mathbf{d}(\mathbf{x}))\right] = \text{tr}(C) + \left(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*\right)^T \left(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*\right)$$

in the optimization process. The values of the separate terms, as well as the complete formula for both the baseline and full variance optimization plan of patient 1 are listed in Table 5.2.3. For the other patients, similiar results were obtained.

Table 5.2.3: Values obtained with baseline and full variance optimization for patient 1.

	$\text{tr}(C)$	$\left(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*\right)^T \left(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*\right)$	$\mathbb{E}\left[F(\mathbf{d}(\mathbf{x}))\right]$
Baseline	$1.6 \cdot 10^4$	$4.5 \cdot 10^5$	$4.6 \cdot 10^5$
Full variance optimization	$5.5 \cdot 10^4$	$3.1 \cdot 10^5$	$3.7 \cdot 10^5$

We observe that by applying full variance optimization, $\mathbb{E}\left[F(\mathbf{d}(\mathbf{x}))\right]$ decreases. However, the term $\left(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*\right)^T \left(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*\right)$ decreases whereas $\text{tr}(C)$ shows an increase when compared to the baseline plan. As a result, in the full variance optimization plan, the mean delivered dose is closer to the prescribed dose but the variance in voxel dose increases when compared to the baseline plan. This causes the difference in minimum and maximum dose to a particular voxel over different treatment scenarios to increase.

A potential explanation for this behaviour is that $\left(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*\right)^T \left(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*\right)$ is one order of magnitude larger than $\text{tr}(C)$. As a result, $\left(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*\right)^T \left(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*\right)$ dominates the minimization process. To test this hypothesis, we apply an adjusted version of full variance optimization. We minimize

$$\text{tr}(C) + \frac{1.6 \cdot 10^4}{4.6 \cdot 10^5} \cdot \left(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*\right)^T \left(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*\right)$$

such that both terms are of the same order. Consequently, both terms are minimized rather than only the second term. The results obtained with this modified version of full variance optimization are shown in Figure 5.2.5. Here, we did not scale the treatment plan such that D98% $\text{CTV}_{7000} \geq 66.5$ Gy and $\text{CTV}_{5425} \geq 51.54$ Gy. Instead, we compare the unscaled D98% CTV_{7000}, D98% CTV_{5425} and D2% CTV_{7000} to the baseline, full and partial variance optimization method.

We observe that when applying full variance optimization, low D98% and high D2% are obtained. Hence, reducing the order of $(\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)^T (\mathbb{E}[\mathbf{d}(\mathbf{x})] - \mathbf{d}^*)$, results in improved CTV coverage and improved D2% CTV_{7000} when compared to full variance optimization. However, the NTPs (not shown) increased.

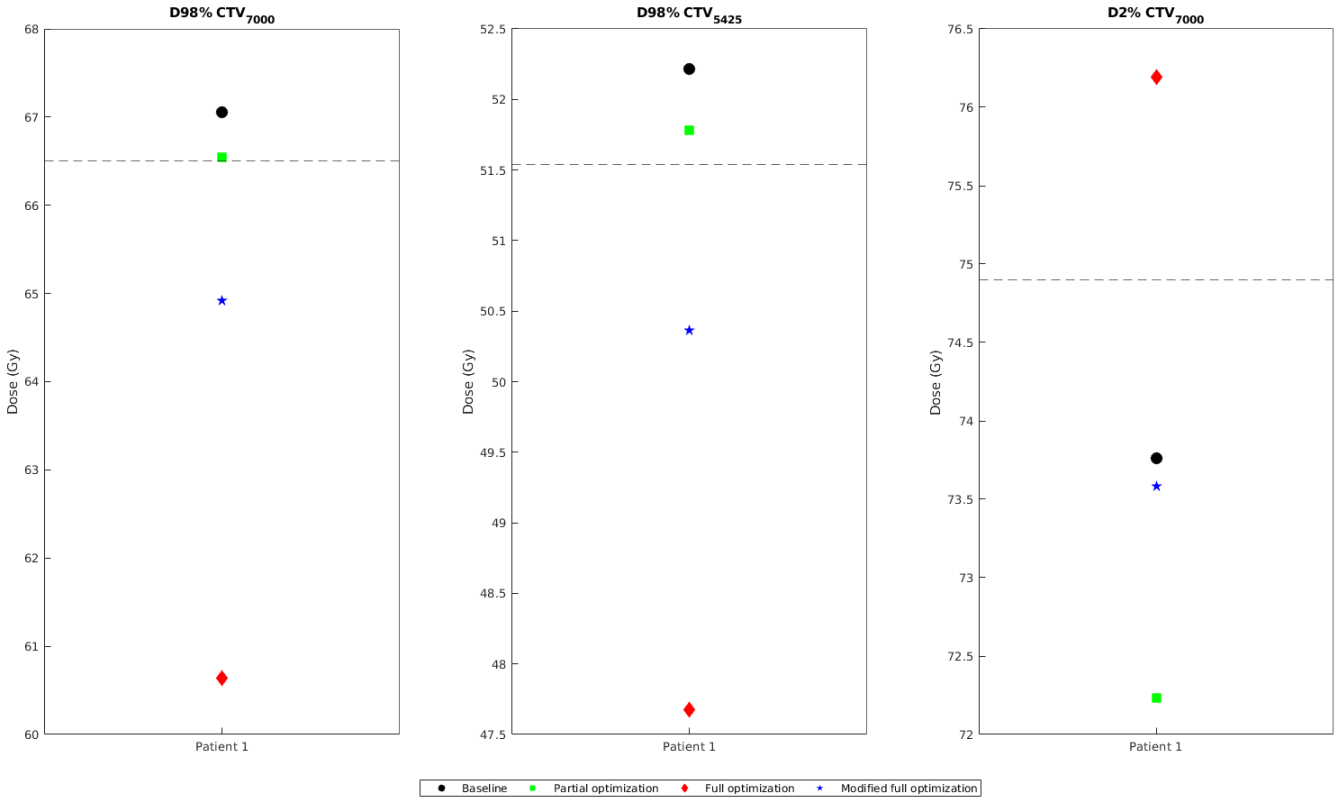


Figure 5.2.5: unscaled D98% CTV_{7000}, D98% CTV_{5425} and D2% CTV_{7000} of the baseline, partial and full variance optimization as well as the modified full variance optimization plan.

5.3 Discussion

In this study we investigate if variance optimization can be applied to reduce optimization time and preserve treatment plan quality. This method only constraints the minimum target coverage in the nominal treatment scenario and adds an additional objective to obtain practical robustness.

Both full and partial variance optimization result in a significant reduction in optimization time when compared to the conventional planning method. As the optimization time is independent of the number of treatment scenarios, many treatment scenarios can be included.

However, when applying full variance optimization, we find that the D2% CTV_{7000} is too high. Also, the NTCPs for both xerostomia and dysphagia increase. It seems that full variance optimization prioritizes minimizing the difference between the mean dose and the prescribed dose over minimizing the variance of the dose in different treatment scenarios. As a result, the minimum dose to the CTV generally decreases whereas the maximum dose to these voxels increases. Consequently, high D2% CTV_{7000} is obtained.

When applying partial variance optimization, we again notice an increase in NTCPs. Nevertheless, applying constrained partial variance optimization shows that similar NTCPs as in the baseline plan can be obtained for similar CTV coverage and suitable D2% CTV_{7000}. Hence, with partial variance optimization, it should be possible to generate treatment plans with al-

most similar dosimetric results in significantly less time. As it is still unknown how to obtain these results, a suggestion for further research would be to find a trade off between NTCPs and CTV coverage. Alternatively, it would be interesting to develop models that predict the NTCPs. When constraining the predicted NTCPs, partial variance optimization could be used to reduce optimization time while preserving treatment plan quality.

Originally, we also generated a baseline, full and partial variance optimization plan using 13 treatment scenarios. As similar results were achieved, we only present the results on 29 treatment scenarios.

Due to increased NTCPs and/or high D2% CTV_{7000}, variance optimization is inferior to the original planning method in Erasmus-iCycle. Hence, we provide some suggestions for further research. First, as full variance optimization causes the difference between minimum and maximum delivered dose to increase, the expectation of the penalized least squares objective function could be optimized. With this method, voxels that are at high risk of deviating from the prescribed dose can be assigned higher penalty which might improve the dose delivery to the CTV.

Another suggestion for further research is to test the method while including more treatment scenarios. As increasing the number of treatment scenarios does not affect the optimization time. These results could also be evaluated using polynomial chaos expansion.

In conclusion, by applying variance optimization results, significant reduction in optimization time up to 99% can be obtained. However, to obtain robustness, increased D2% CTV_{7000} and/or increased NTCPs are obtained. Therefore, variance optimization is currently not applicable in clinical practice.

6 Conclusion

In this section, we summarize the most important conclusions. Furthermore, we formulate answers to our research questions. We begin with answering the first research question:

> Can we reduce optimization time while maintaining treatment plan quality?

To answer this question, two methods were investigated. In Section 3, it is investigated if treatment plans can be generated while accounting for fewer treatment scenarios. It is found that reducing the number of treatment scenarios decreases the optimization time up to 75%. Additionally, treatment plan quality can be preserved when planning on 13 instead of 29 treatment scenarios. Therefore, it is suggested to plan on 13 instead of 29 treatment scenarios using Erasmus-iCycle. The second method that is investigated to reduce the optimization time of treatment plan generation, is variance optimization, as presented in Section 5. It was found that with this method, the optimization time can be reduced up to 99% of the original optimization time. However, to obtain good target coverage, the risk of two main complications, xerostomia and dysphagia, increases. Furthermore, often too high near-maximum doses are obtained. As a result, this method is currently not applicable in clinical practice and suggestions for further research were provided.

Next, we answer the second research question:

> Can we reduce the risk of side effects?

In Section 4 we investigated if allowing higher maximum dose to the tumor can be used to reduce the risk of side effects. We found that by increasing the maximum allowable dose to the entire tumor, the risk of dysphagia and xerostomia can be lowered up to 7.5%-point. However, according to radiation oncologists, it is safest to only increase the maximum allowable dose to the tumor core. If not, the risk of edema and fibrosis increases. By only allowing higher maximum dose to the tumor core, the extent to which the risk of side effects decreases is limited (<1%-point). However, it is recommended to investigate if, for some patients, the potential increase in risk of edema and fibrosis outweigh the significantly reduced risk of xerostomia and dysphagia.

It is important to note that all methods investigated in this thesis have been tested on four patients. As a results, our conclusions are based on a small patient population. Therefore, it is recommended to test each method for more patients.

References

[1] Globocan, "All cancers excluding non-melanoma skin cancer." https://gco.iarc.fr/today/data/factsheets/cancers/40-All-cancers-excluding-non-melanoma-skin-cancer-fact-sheet.pdf. Accessed on 30-12-2022.

[2] H. Y. Pan, B. G. Haffty, B. P. Falit, T. A. Buchholz, L. D. Wilson, S. M. Hahn, and B. D. Smith, "Supply and demand for radiation oncology in the united states: Updated projections for 2015 to 2025," *International Journal of Radiation Oncology* Biology* Physics*, vol. 96, no. 3, pp. 493–500, 2016.

[3] P. T. C.-O. Group, "Particle therapy facilities in clinical operation," 2020.

[4] "Provision healthcare." `https://provisionhealthcare.com/about-proton-therapy/advantages-of-proton/`. Accessed: 10-11-2022.

[5] S. Breedveld, D. Craft, R. Van Haveren, and B. Heijmen, "Multi-criteria optimization and decision-making in radiotherapy," *European Journal of Operational Research*, vol. 277, no. 1, pp. 1–19, 2019.

[6] Z. Perko, D. Lathouwers, and M. Hoogeman, "Medical physics of photon and proton therapy, lecture series delft university of technology," 2022.

[7] A. E. Nahum, "The radiobiology of hypofractionation," *Clinical oncology*, vol. 27, no. 5, pp. 260–269, 2015.

[8] J. Rojo-Santiago, S. J. Habraken, D. Lathouwers, A. M. Romero, Z. Perkó, and M. S. Hoogeman, "Accurate assessment of a dutch practical robustness evaluation protocol in clinical pt with pencil beam scanning for neurological tumors," *Radiotherapy and Oncology*, vol. 163, pp. 121–127, 2021.

[9] X. Liang, R. B. Mailhot Vega, Z. Li, D. Zheng, N. Mendenhall, and J. A. Bradley, "Dosimetric consequences of image guidance techniques on robust optimized intensity-modulated proton therapy for treatment of breast cancer," *Radiation oncology*, vol. 15, pp. 1–9, 2020.

[10] M. Oud, S. Breedveld, M. Giżyńska, M. Kroesen, S. Hutschemaekers, S. Habraken, S. Petit, Z. Perkó, B. Heijmen, and M. Hoogeman, "An online adaptive plan library approach for intensity modulated proton therapy for head and neck cancer," *Radiotherapy and Oncology*, vol. 176, pp. 68–75, 2022.

[11] E. W. Korevaar, S. J. Habraken, D. Scandurra, R. G. Kierkels, M. Unipan, M. G. Eenink, R. J. Steenbakkers, S. G. Peeters, J. D. Zindler, M. Hoogeman, *et al.*, "Practical robustness evaluation in radiotherapy–a photon and proton-proof alternative to ptv-based plan evaluation," *Radiotherapy and Oncology*, vol. 141, pp. 267–274, 2019.

[12] S. McGowan, N. Burnet, and A. Lomax, "Treatment planning optimisation in proton therapy," *The British journal of radiology*, vol. 86, no. 1021, pp. 20120288–20120288, 2013.

[13] V. T. Taasti, L. P. Muren, K. Jensen, J. B. B. Petersen, J. Thygesen, A. Tietze, C. Grau, and D. C. Hansen, "Comparison of single and dual energy ct for stopping power determination in proton therapy of head and neck cancer," *Physics and imaging in radiation oncology*, vol. 6, pp. 14–19, 2018.

[14] J. Unkelbach, T. Bortfeld, B. C. Martin, and M. Soukup, "Reducing the sensitivity of impt treatment plans to setup errors and range uncertainties via probabilistic treatment planning," *Medical physics*, vol. 36, no. 1, pp. 149–163, 2009.

[15] J. Unkelbach, M. Alber, M. Bangert, R. Bokrantz, T. C. Chan, J. O. Deasy, A. Fredriksson, B. L. Gorissen, M. Van Herk, W. Liu, *et al.*, "Robust radiotherapy planning," *Physics in Medicine & Biology*, vol. 63, no. 22, p. 22TR02, 2018.

[16] J. Unkelbach and H. Paganetti, "Robust proton treatment planning: physical and biological optimization," in *Seminars in radiation oncology*, vol. 28, pp. 88–96, Elsevier, 2018.

[17] W. Chen, J. Unkelbach, A. Trofimov, T. Madden, H. Kooy, T. Bortfeld, and D. Craft, "Including robustness in multi-criteria optimization for intensity-modulated proton therapy," *Physics in Medicine & Biology*, vol. 57, no. 3, p. 591, 2012.

[18] J. Unkelbach, T. C. Chan, and T. Bortfeld, "Accounting for range uncertainties in the optimization of intensity modulated proton therapy," *Physics in Medicine & Biology*, vol. 52, no. 10, p. 2755, 2007.

[19] S. Breedveld, P. R. Storchi, P. W. Voet, and B. J. Heijmen, "icycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar imrt plans," *Medical physics*, vol. 39, no. 2, pp. 951–963, 2012.

[20] S. Breedveld, B. van den Berg, and B. Heijmen, "An interior-point implementation developed and tuned for radiation therapy treatment planning," *Computational Optimization and Applications*, vol. 68, no. 2, pp. 209–242, 2017.

[21] Z. Perkó, S. R. Van der Voort, S. Van De Water, C. M. Hartman, M. Hoogeman, and D. Lathouwers, "Fast and accurate sensitivity analysis of impt treatment plans using polynomial chaos expansion," *Physics in Medicine & Biology*, vol. 61, no. 12, p. 4646, 2016.

[22] D. Xiu and G. E. Karniadakis, "The wiener–askey polynomial chaos for stochastic differential equations," *SIAM journal on scientific computing*, vol. 24, no. 2, pp. 619–644, 2002.

[23] E. N. Marieb and K. Hoehn, "Human anatomy and physiology. human anatomy," 2015.

[24] N. V. voor Radiotherapie en Oncologie (NVRO), "Landelijk indicatie protocol protonentherapie," 2019.

[25] S. Van De Water, I. van Dam, D. R. Schaart, A. Al-Mamgani, B. J. Heijmen, and M. S. Hoogeman, "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," *Radiotherapy and Oncology*, vol. 120, no. 1, pp. 56–62, 2016.

[26] S. Petit, J. Seco, and H. Kooy, "Increasing maximum tumor dose to manage range uncertainties in impt treatment planning," *Physics in Medicine & Biology*, vol. 58, no. 20, p. 7329, 2013.

[27] S. F. Petit, S. Breedveld, J. Unkelbach, D. den Hertog, and M. Balvert, "Robust dose-painting-by-numbers vs. nonselective dose escalation for non-small cell lung cancer patients," *Medical physics*, vol. 48, no. 6, pp. 3096–3108, 2021.

[28] N. Wahl and H.Wieser. (2022, June 27 - July 2). Scenario-free probabilistic proton dose optimization using expected dose influence and total variance [Conference presentation abstract]. PTCOG 2022 Miami, United States. https://www.czech-in.org/cmPortalv15/Searchable/ptcog60/config/ nodetails#abstractdetails/0001130800.

Appendix

A Wish-list

Table A.1: Wish-list used in this study.

Constraints					
	Structure	**Type**	**Goal**	**Sufficient**	**Robust**
	CTV_{7000}	min	67.9 Gy	67.9 Gy	29 scenarios
	$CTV_{int10mm}$	min	53.2 Gy	53.2 Gy	29 scenarios
	$CTV_{shrunk5425}$	min	53.2 Gy	53.2 Gy	29 scenarios
Objectives					
Priority	**Structure**	**Type**	**Goal**	**Sufficient**	**Robust**
1	CTV_{7000}	minimize max	74.2 Gy	74.2 Gy	29 scenarios
1	$CTV_{int10mm}$	minimize max	57.5 Gy	57.5 Gy	29 scenarios
1	$CTV_{shrunk5425}$	minimize max	57.5 Gy	57.5 Gy	29 scenarios
2	$CTV_{combined\ ring\ 0\text{-}10\ mm}$	minimize max	57.5 Gy	57.5 Gy	No
2	$CTV_{combined\ ring\ 10\text{-}15\ mm}$	minimize max	57.5 Gy	57.5 Gy	No
3	$\mathbf{x}_i$	minimize max	35.5	35.5	No
4	Brain stem	minimize max	30 Gy	1 Gy	29 scenarios
4	Spinal cord	minimize max	30 Gy	1 Gy	29 scenarios
5	Brain	minimize mean	1 Gy	1 Gy	29 scenarios
6	Parotid	minimize mean	1 Gy	1 Gy	No
7	Submandibular gland	minimize mean	1 Gy	1 Gy	No
8	Constrictor muscle superior	minimize mean	1 Gy	1 Gy	No
8	Constrictor muscle middle	minimize mean	1 Gy	1 Gy	No
9	Oral cavity	minimize mean	1 Gy	1 Gy	No
12	Muscle esophagus inlet	minimize mean	1 Gy	1 Gy	No
10	Larynx subglottic	minimize mean	1 Gy	1 Gy	No
10	Glottic area	minimize mean	1 Gy	1 Gy	No
11	Constrictor muscle inferior	minimize mean	1 Gy	1 Gy	No
11	Cricopharyngeus	minimize mean	1 Gy	1 Gy	No
13	$CTV_{combined\ ring\ 0\text{-}10\ mm}$	minimize max	1 Gy	1 Gy	No
13	$CTV_{combined\ ring\ 10\text{-}15\ mm}$	minimize max	1 Gy	1 Gy	No
13	$CTV_{combined\ ring\ 15\text{-}25\ mm}$	minimize max	1 Gy	1 Gy	No
14	$\sum_{i=1}^{m} \mathbf{x}_i$	minimize mean	1	1	No

Here, $\mathbf{x}_i$ denotes the $i-$th element of $\mathbf{x}$ which is the vector that contains the pencil beam intensities.

B Treatment scenario reduction

Here, we present some additional results. In Section B.1, we present the optimization time and the dosimetric results, evaluated with both voxelwise and PCE, when optimizing for 1, 13, 29 and 53 treatment scenarios. Additionally, in Section B.2, the D98% CTV_{7000}, D98% CTV_{5425} and D2% CTV_{7000} are shown for the unscaled treatment plans on the planning CT scan when optimizing on 1, 13, 29 and 53 treatment scenarios.

B.1 Optimizing nominal treatment scenario

Optimization times

Here we present the optimization time for plans generated using the nominal treatment scenario as well as 13, 29 and 53 treatment scenarios. The results are listed in Table B.1.

Table B.1: Optimization times (in minutes) for different number of treatment scenarios.

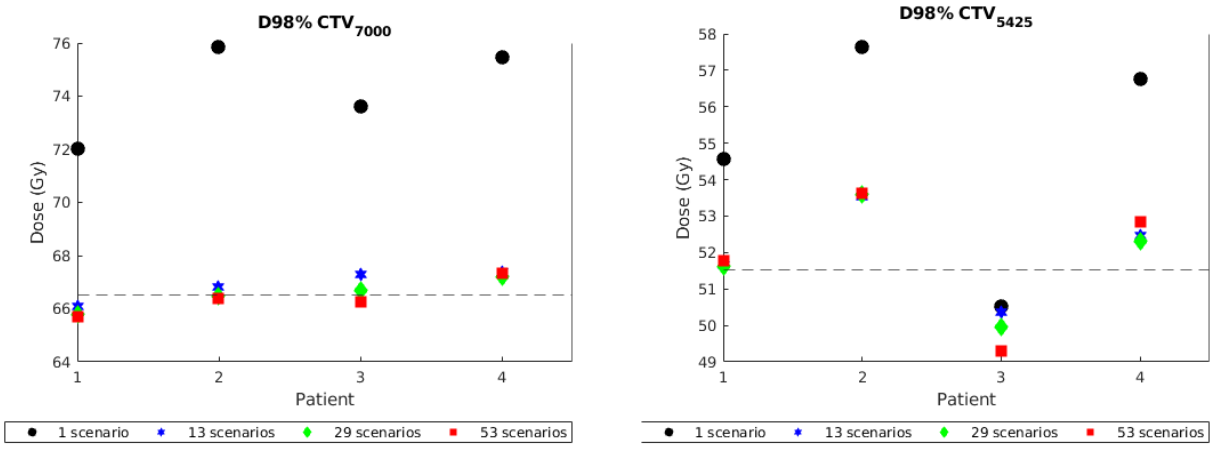
# scenarios	Patient 1	Patient 2	Patient 3	Patient 4
1	22	21	14	16
13	351	226	195	197
29	1085	622	614	611
53	3572	2433	1507	2406

Optimizing nominal treatment scenario

In Figure B.1 the D98% on the repeat CT scan is shown in the voxelwise minimum scenario when optimizing for 1, 13, 29 and 53 treatment scenarios. Figure B.1a shows the results for CTV_{7000} and Figure B.1b shows the results for CTV_{5425}. Similarly, in Figure B.2, D98% is shown using PCE evaluation. According to Figure B.1, both D98% CTV_{7000} and D98% CTV_{5425} increase significantly when planning only on the nominal treatment scenario. In Figure B.2, we find similar results when using PCE evaluation.

In Figure B.3, the D2% CTV_{7000} is shown when optimizing for 1, 13, 29 and 53 treatment scenarios. The results are obtained in the voxelwise maximum scenario. Figure B.3a shows the results on the planning CT scan. Similarly, B.3b shows the results on the repeat CT scan. In addition, in Figure B.4, the same results are shown when PCE evaluation is applied. For both evaluation methods, we find that the D2% CTV_{7000} drastically increases when planning on the nominal treatment scenario when compared to planning on 13, 29 or 53 treatment scenarios. In all cases, the D2% $CTV_{7000} \leq 74.90$ Gy is exceeded.

Finally, in Figure B.5, the difference (%-point) in NTCPs of xerostomia and dysphagia are shown when optimizing on 1, 13, 29 and 53 treatment scenarios on the repeat CT scan. Figure B.5a shows the results for xerostomia and B.5b shows the results for dysphagia. Furthermore, in Figure B.6 the results are shown when PCE evaluation is applied. We observe for both xerostomia and dysphagia that the evaluation methods are in line with each other. However, there is no trend. For some patients, the NTCP increases whereas for the others, it decreases.



(a) D98% CTV_{7000} on the repeat CT scan.

(b) D98% CTV_{5425} on the repeat CT scan.

Figure B.1: D98% CTV_{7000} and D98% CTV_{5425} for 1, 13, 29 and 53 treatment scenarios on the repeat CT scan. Evaluated with voxelwise minimum evaluation.

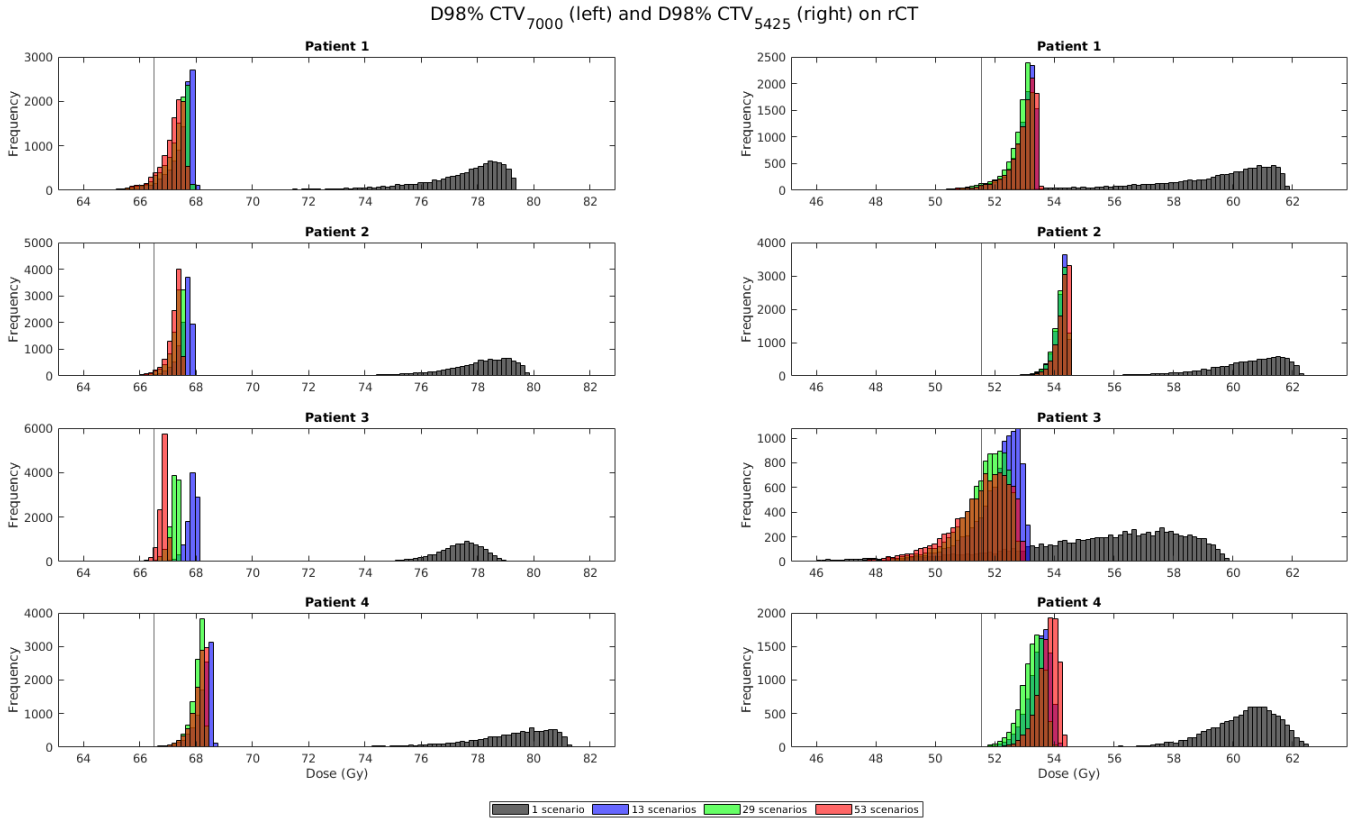
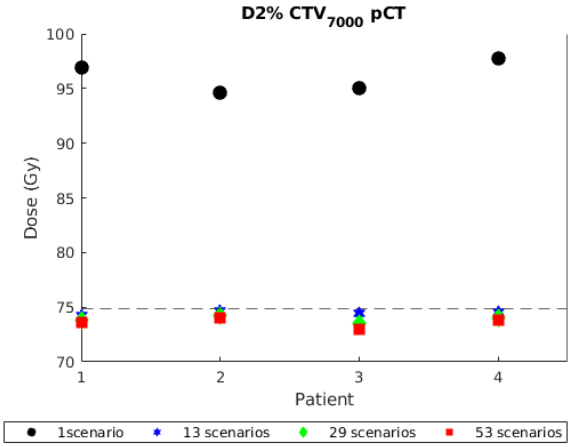
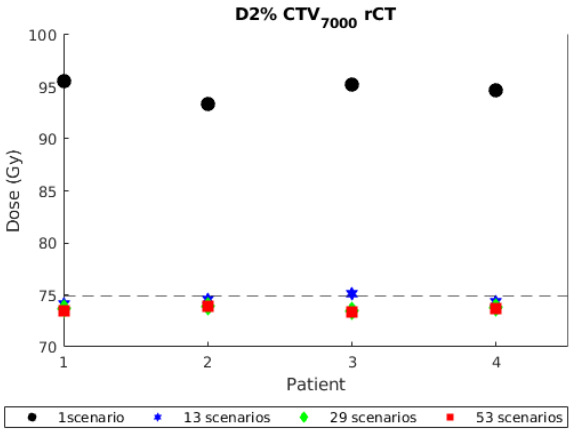


Figure B.2: D98% CTV_{7000} and D98% CTV_{5425} for 1, 13, 29 and 53 treatment scenarios on the repeat CT scan. Evaluated using polynomial chaos expansion.



(a) D2% CTV_{7000} on planning CT scan.



(b) D2% CTV_{7000} on repeat CT scan.

Figure B.3: D2% CTV_{7000} for 1, 13, 29 and 53 treatment scenarios on both the planning and repeat CT scan. Evaluated with voxelwise maximum evaluation.

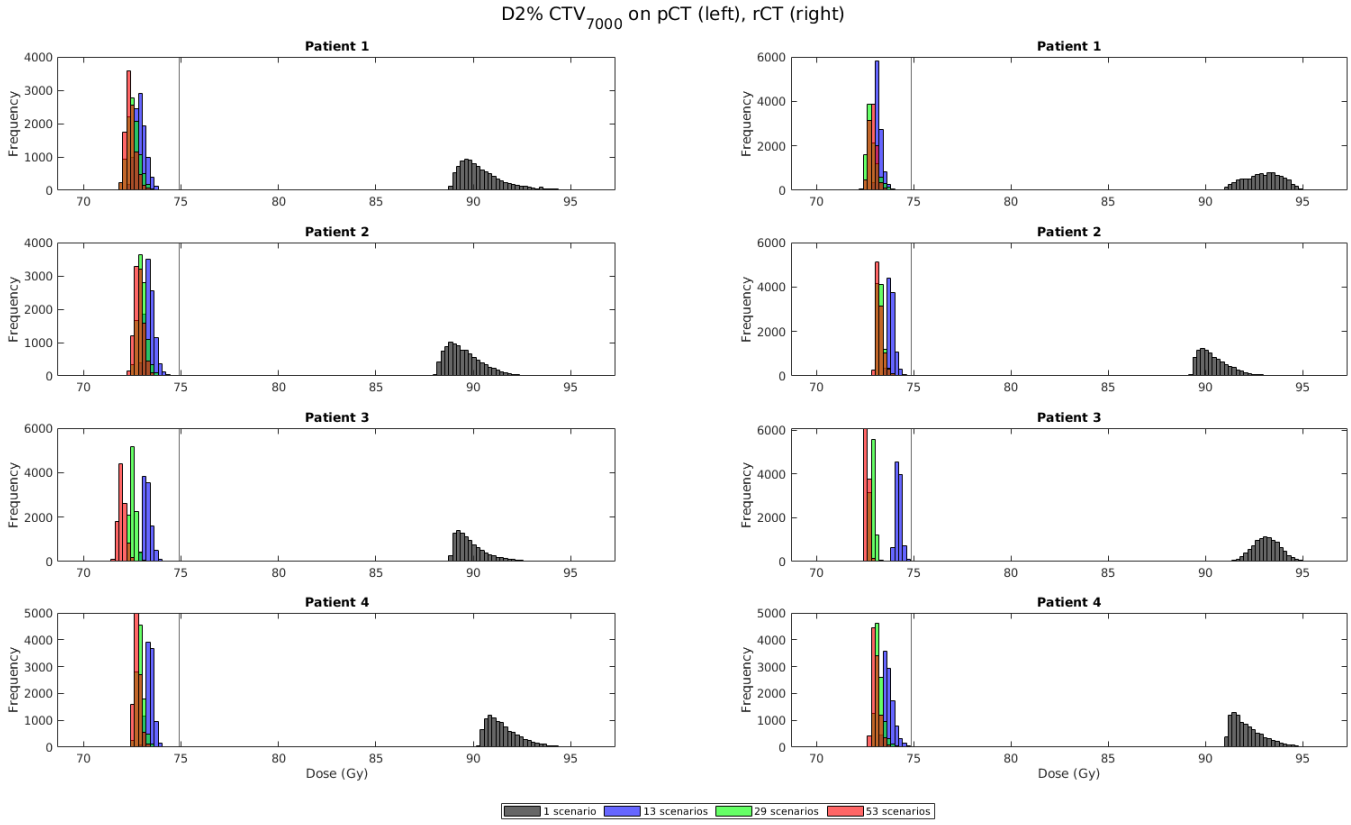
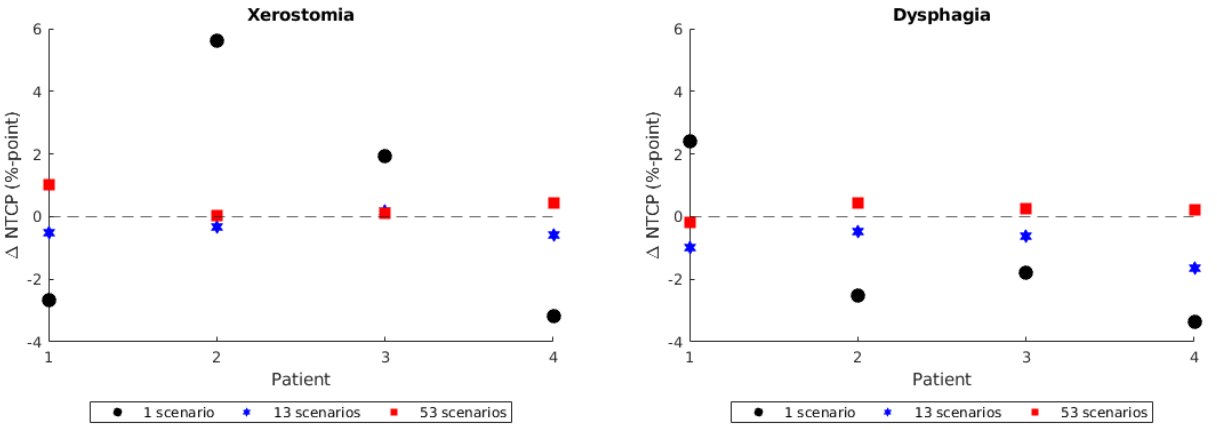


Figure B.4: D2% CTV_{7000} for 1, 13, 29 and 53 treatment scenarios on both the planning and repeat CT scan. Evaluated using polynomial chaos expansion.



(a) NTCP of xerostomia on the repeat CT scan.

(b) NTCP of dysphagia on the repeat CT scan.

Figure B.5: NTCPs of xerostomia and dysphagia when using 1, 13 and 53 treatment scenarios with respect to using 29 treatment scenarios.

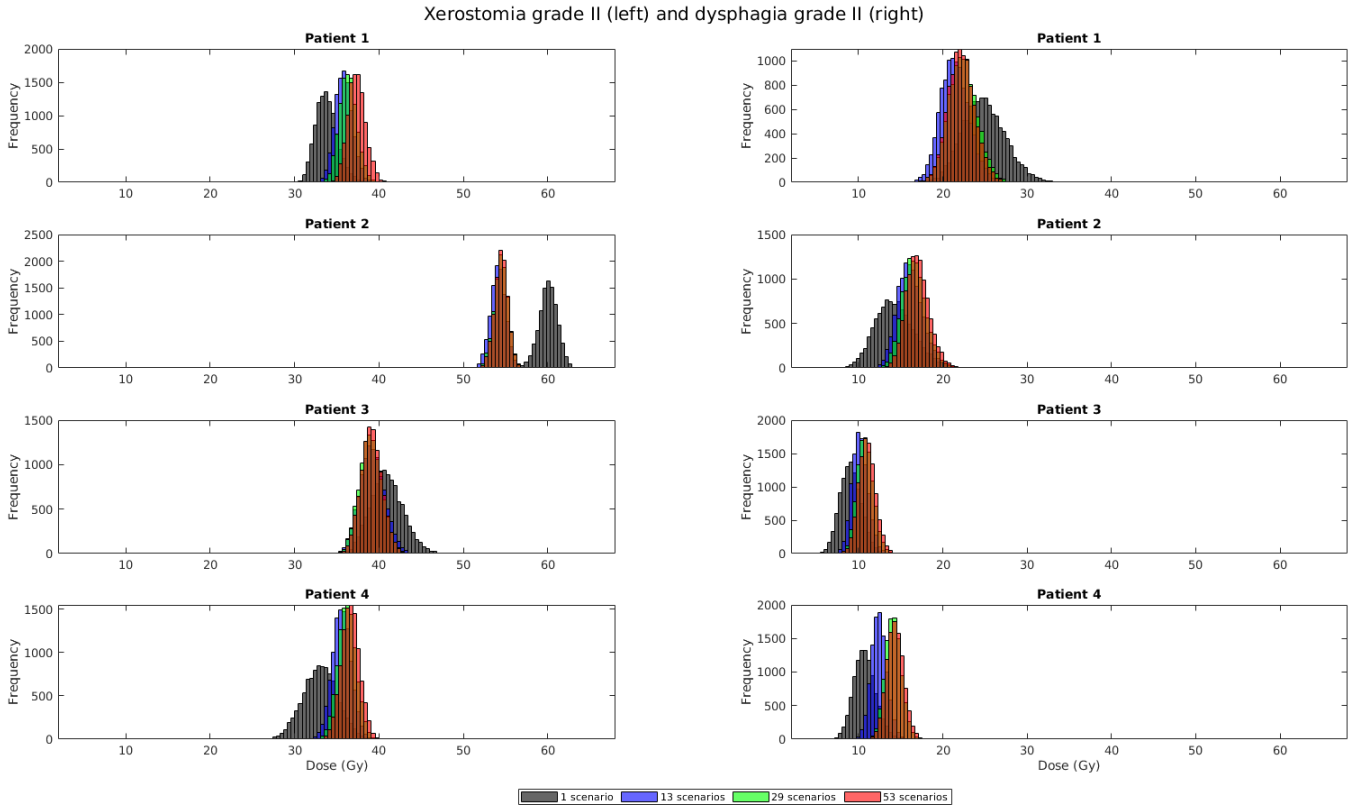
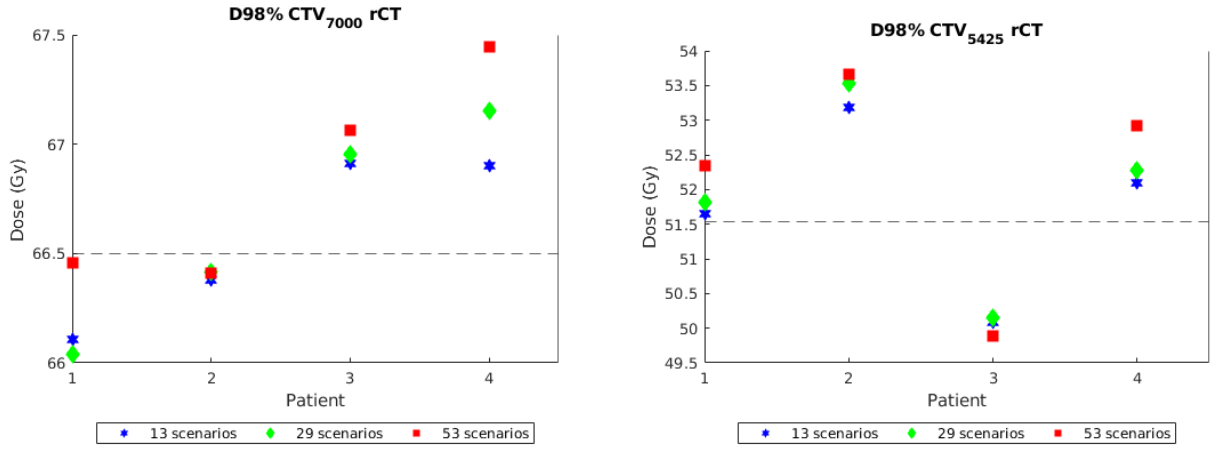


Figure B.6: NTCPs of xerostomia and dysphagia for 1, 13, 29 and 53 treatment scenarios on the repeat CT scan. Evaluated using polynomial chaos expansion.

B.2 Unscaled treatment plans

Here, we show the results of the unscaled treatment plans. In Figure B.7, D98% on the repeat CT scan is shown using voxelwise minimum evaluation on 29 treatment scenarios. Figure B.7a shows the results for CTV_{7000}. Similarly, in Figure B.7b the results for CTV_{5425} are shown. In Figure B.8 the results are shown when PCE is used to evaluate the treatment plans. Additionally, in Figure B.9b, D2% CTV_{7000} is shown on the planning and repeat CT scan. Similarly, in Figure

B.10 D2% CTV_{7000} is shown using PCE evaluation. We observe that D98% CTV_{7000} and D98% CTV_{5425} are lower when planning on the fewer treatment scenarios whereas D2% CTV_{7000} is relatively similar.



(a) Unscaled D98% CTV_{7000} on repeat CT scan.

(b) Unscaled D98% CTV_{5425} on repeat CT scan.

Figure B.7: Unscaled D98% CTV_{7000} and CTV_{5425} for 13, 29 and 53 treatment scenarios on the repeat CT scan using voxelwise minimum evaluation.

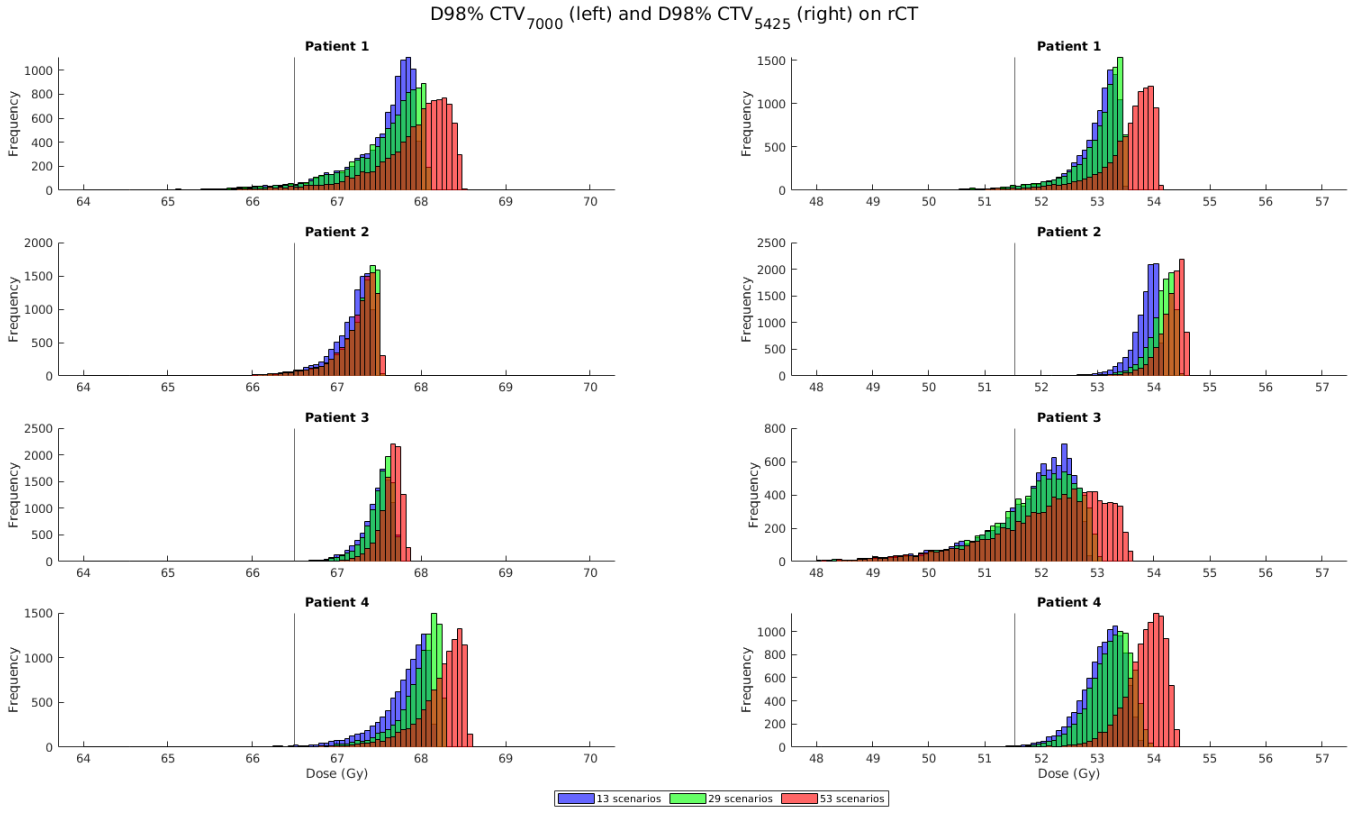
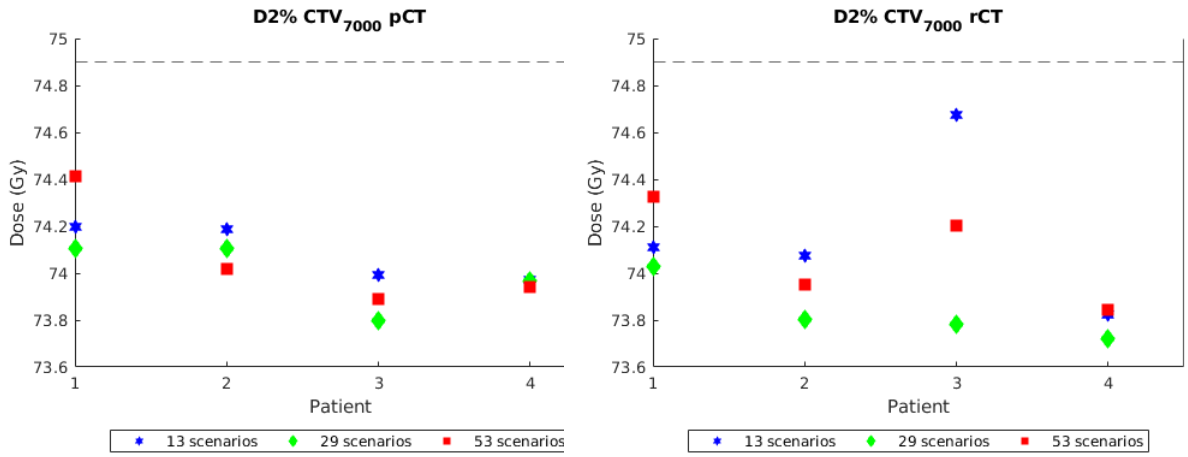


Figure B.8: Unscaled D98% CTV_{7000} and D98% CTV_{5425} for 13, 29 and 53 treatment scenarios on the repeat CT scan. Evaluated using polynomial chaos expansion.



(a) Unscaled D2% CTV_{7000} on planning CT scan. (b) Unscaled D2% CTV_{7000} on repeat CT scan.

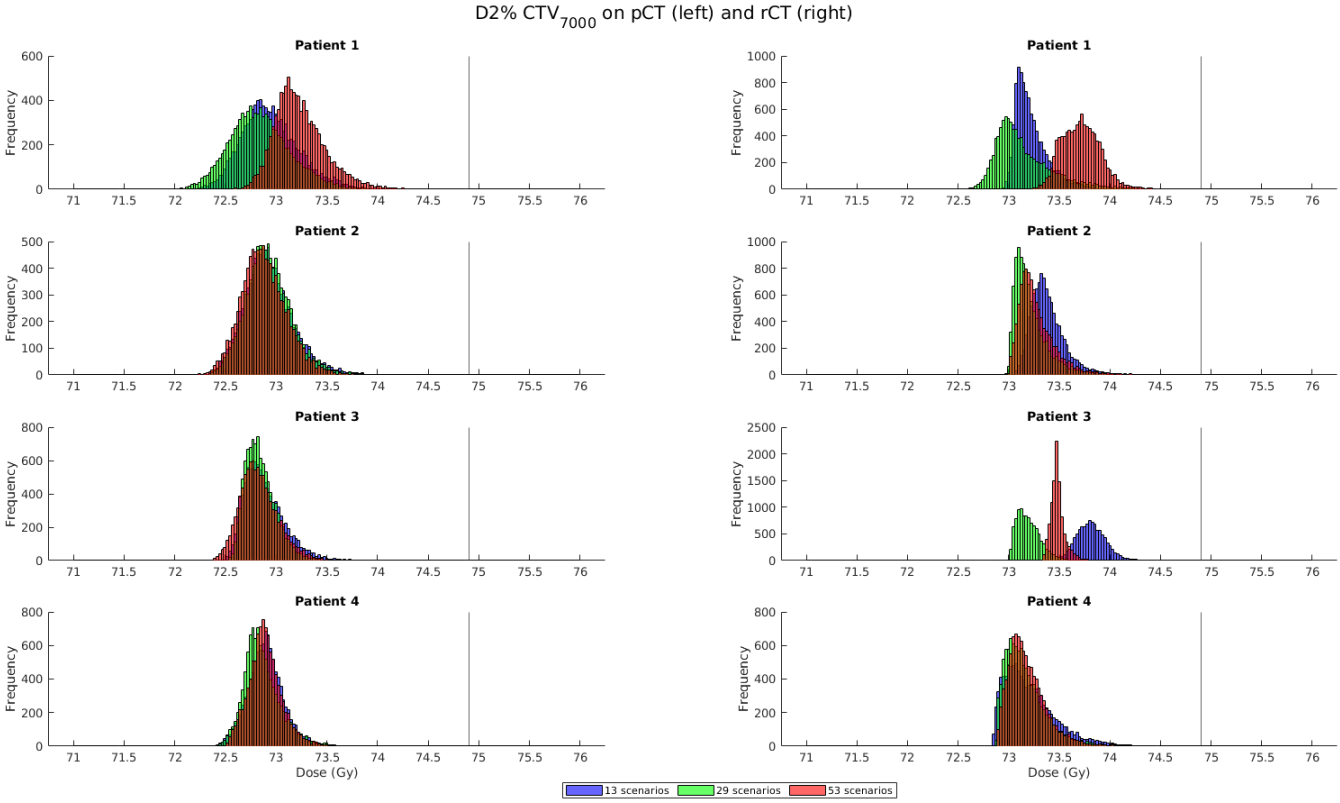


Figure B.10: Unscaled D2% CTV_{7000} for 13, 29 and 53 treatment scenarios on the planning and repeat CT scan using PCE evaluation.

C Dose escalation

Here, we show the dosimetric impact to other CTV structures and OARs when compared to the baseline plan. These other CTV parts are isotropic rings around $\text{CTV}_{\text{combined}}$, which is defined as

$$\text{CTV}_{\text{combined}} = \text{CTV}_{5425} \cup \text{CTV}_{7000} \cup \text{CTV}_{7000\text{r0-10mm}},$$

where $\text{CTV}_{7000\text{r0-10mm}}$ is an isotropic ring around CTV_{7000} of 10 mm. For clarification, the structure $\text{CTV}_{\text{combined}}$ is schematically depicted in Figure C.1. To investigate the dosimetric effect on tissues surrounding the parts where dosis escalation is applied, we compare the delivered dose in three different rings around $\text{CTV}_{\text{combined}}$. More specifically, the delivered dose in a ring of 0-10 mm, 10-15 mm and 15-25 mm around $\text{CTV}_{\text{combined}}$ is compared to the delivered dose in the baseline plan. Furthermore, we investigate the delivered dose to the brainstem, spinal cord, esophagus inlet muscle, the subglottic larynx, the glottic area and the cricopharyngeus.

For the CTV structures, except CTV_{7000}, $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{\text{shrunk5425}}$, the max and mean values are computed in the nominal treatment scenario. For CTV_{7000}, $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{\text{shrunk5425}}$, the maximum value is computed over all 29 treatment scenarios.

For all OARs, except the spinal cord, brainstem and brain, the mean dose is computed over all voxels in the nominal treatment scenario. The mean value of the spinal cord, brainstem and brain are computed using all 29 treatment scenarios.

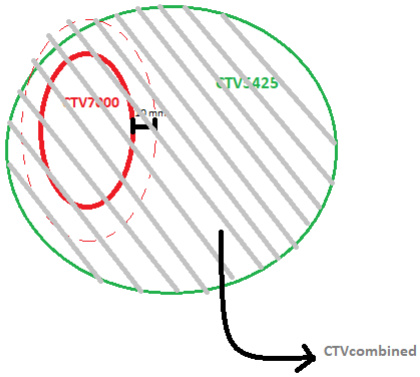


Figure C.1: Sketch of $\text{CTV}_{\text{combined}}$.

C.1 Full dose escalation

We investigate the dosimetric consequences of full dose escalation on several CTV parts and OARs. The dose these parts receive when compared to the baseline plan from the multicriterial optimization process are presented in Figures C.2 and C.3, respectively.

In Figure C.2, we observe that by increasing the maximum allowable dose to CTV_{7000}, $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{5425\text{shrunk}}$, the mean delivered dose in all rings around the $\text{CTV}_{\text{combined}}$ decreases. However, the maximum delivered dose to the ring of 0-10 around $\text{CTV}_{\text{combined}}$ increases compared to the baseline plan.

In Figure C.3, we note that all structures, except the superior pharyngeal constrictor muscle in patient 1, receive lower or similar dose compared to the baseline plan. For most OARs, the difference stabilizes for higher values of the maximum allowable dose.

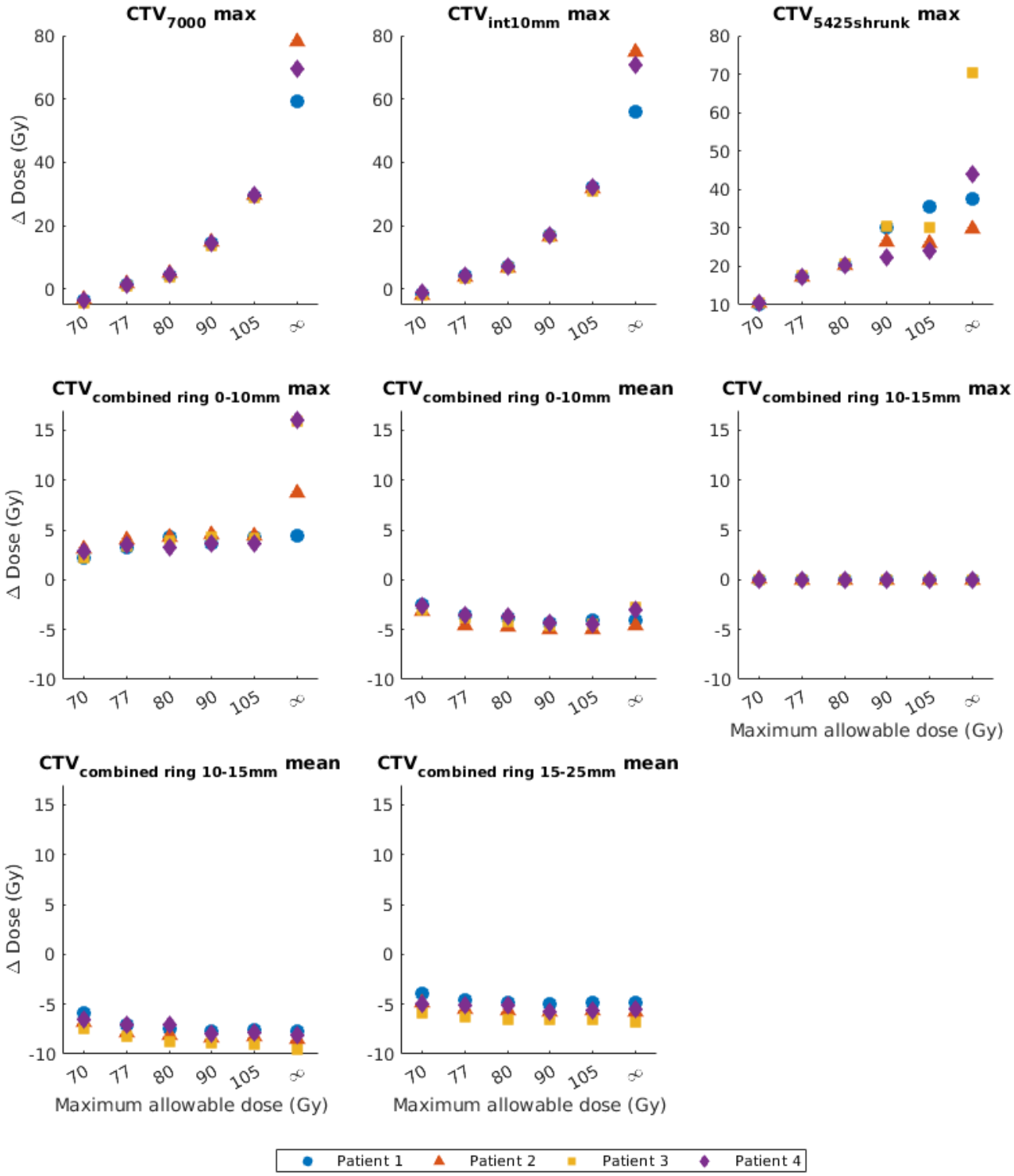


Figure C.2: Difference (Gy) in CTV dose using full dose escalation with respect the baseline plan.

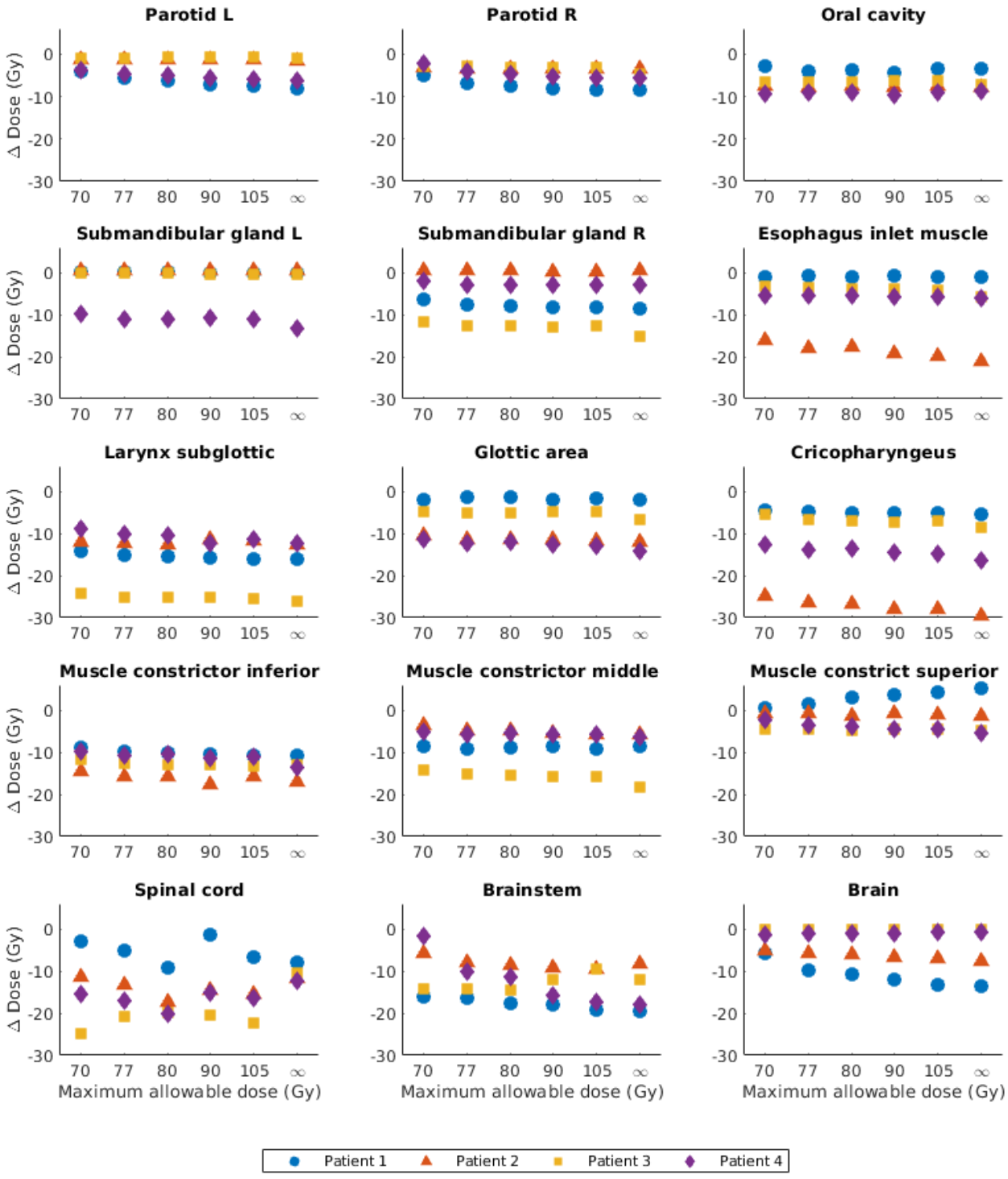


Figure C.3: Difference (Gy) in OAR dose using full dose escalation with respect to the baseline plan.

C.2 Partial dose escalation

We show the difference in dose delivered to several CTV parts and OARs as obtained from the multicriterial optimization for increased maximum allowable dose to CTV_{7000} when compared to the baseline. The results are presented in Figures C.4 and C.5, respectively.

In Figure C.4, we observe that by increasing the maximum allowable dose to CTV_{7000}, especially the dose in this region increases. The areas closest to this region, $\text{CTV}_{\text{int10mm}}$ and $\text{CTV}_{5425\text{shrunk}}$ also experiences a small increase in maximum delivered dose, generally. In the rings around $\text{CTV}_{\text{combined}}$, the maximum and mean dose are relatively similar to the maximum and mean dose in the baseline plan.

The dose delivered to the OARs is much more similar to the baseline plan than in the full dose escalation experiments, as can observed in Figure C.5.

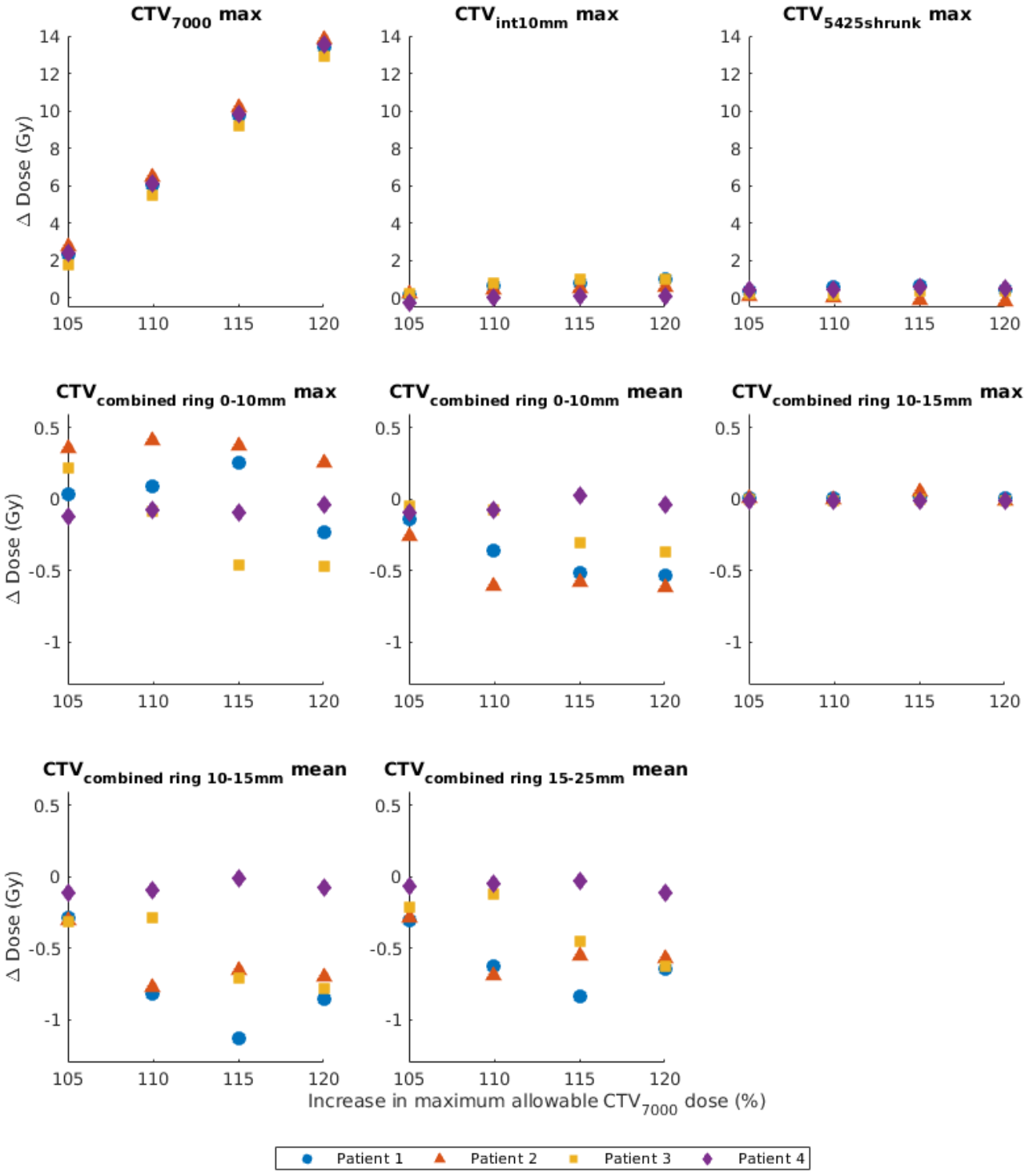


Figure C.4: Difference (Gy) in CTV dose when increasing maximum allowable dose CTV_{7000} with respect to the baseline plan.

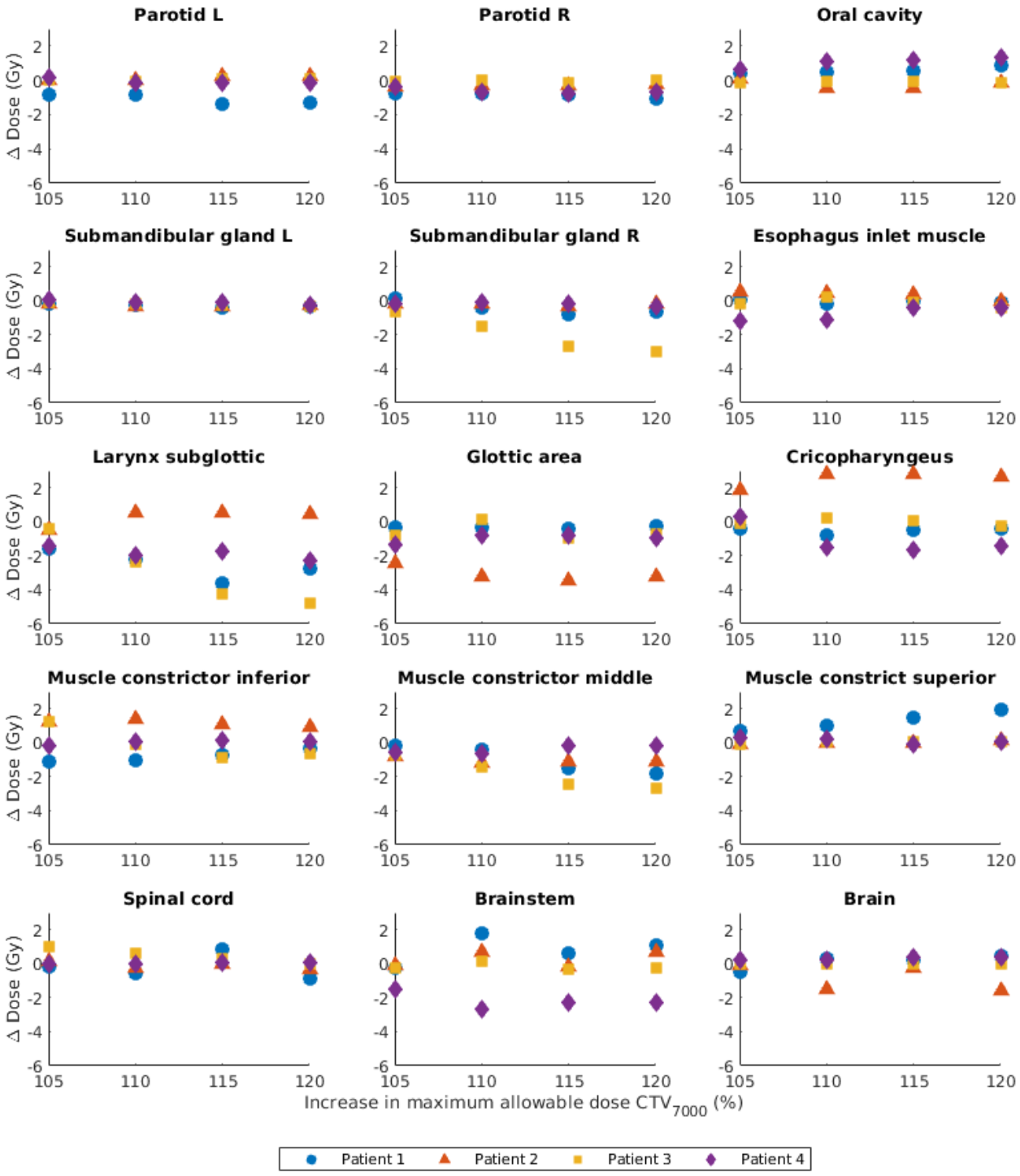


Figure C.5: Difference in OAR dose (Gy) when increasing the maximum allowable CTV_{7000} dose with respect to the baseline plan.

Finally, we investigate the mean and maximum dose to several CTV parts and OARs, as obtained from the multicriterial optimization when applying dose escalation to $CTV_{int10mm}$ and $CTV_{5425shrunk}$. The results are shown in Figures C.6 and C.7, respectively.

In Figure C.6 we observe that especially the dose to $CTV_{5425shrunk}$ increases compared to the baseline plan. Whereas for increasing values of the maximum allowable dose to both $CTV_{int10mm}$ and $CTV_{5425shrunk}$, the maximum dose to CTV_{7000} and $CTV_{int10mm}$ decreases slightly. In the combined ring 0-10 mm, the maximum delivered dose increases with increasing values of the maximum allowable dose. The other rings generally receive a lower mean dose.

Most OARs receive similar or a decreased amount of dose when increasing the maximum allowable dose when compared to the baseline plan, as can be observed in Figure C.7.

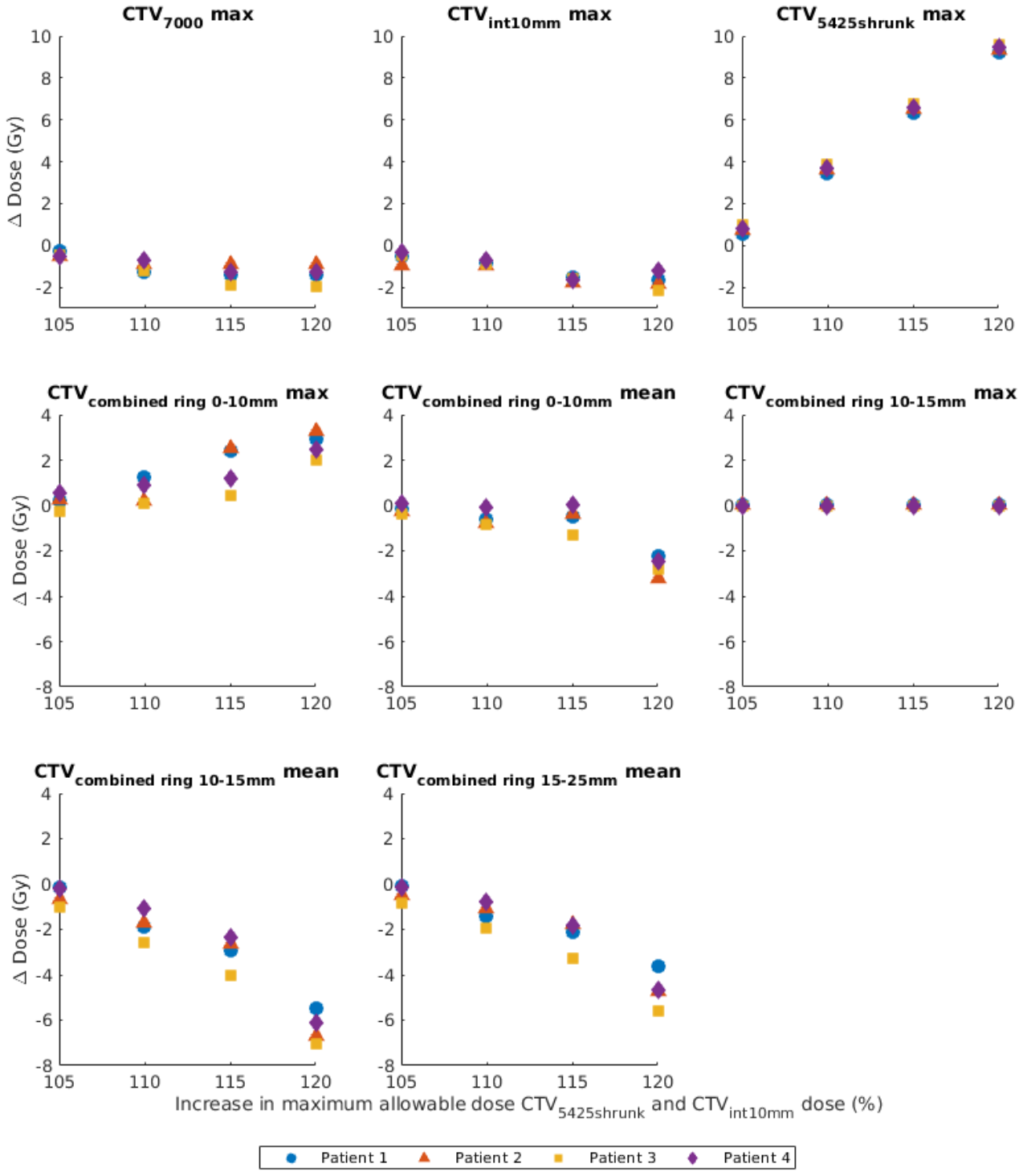


Figure C.6: Difference (Gy) in CTV dose when increasing the maximum allowable dose to $CTV_{5425shrunk}$ and $CTV_{int10mm}$ with respect to the baseline plan.

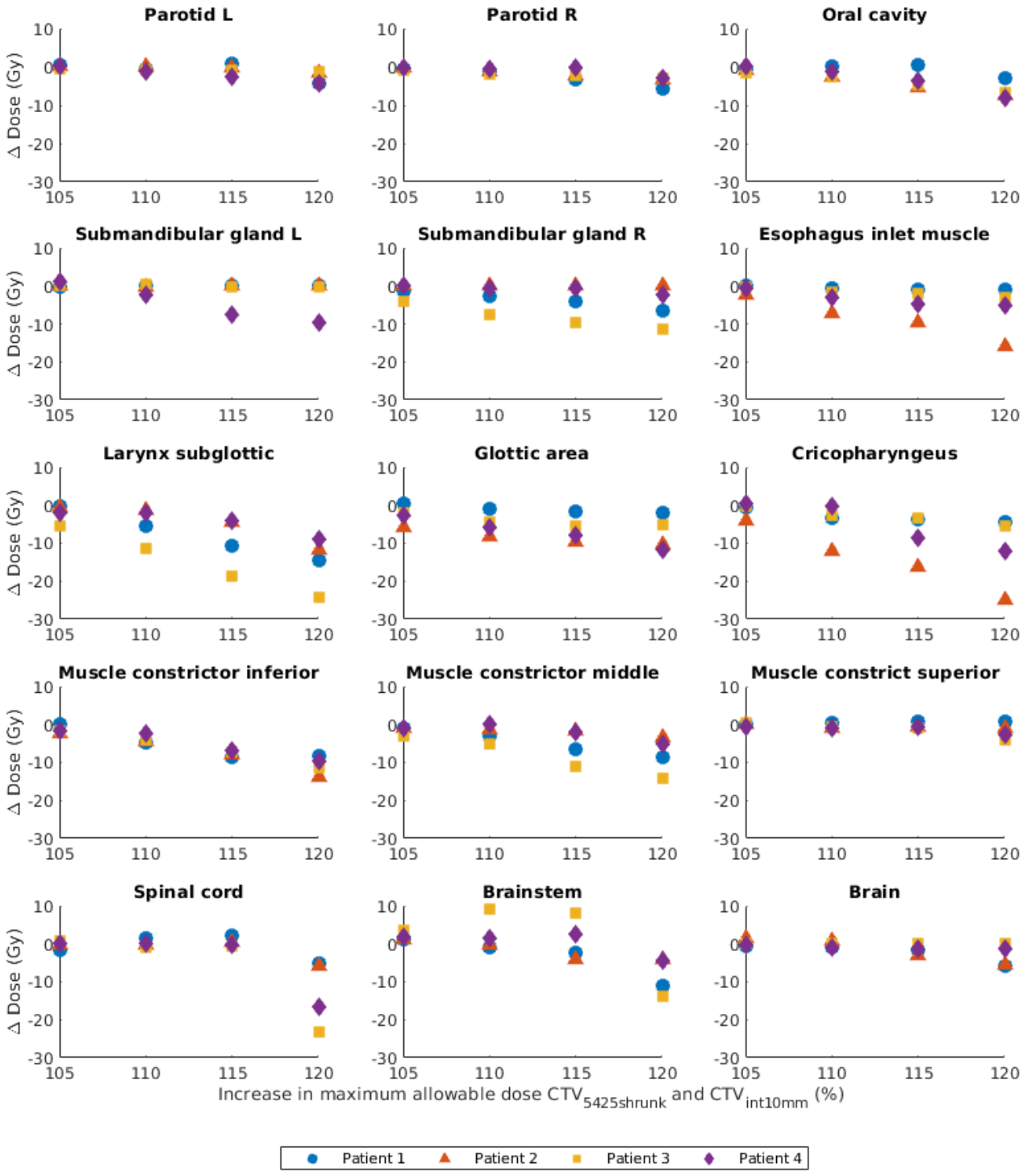


Figure C.7: Difference (Gy) in OAR dose obtained when increasing the maximum allowable dose to $CTV_{5425shrunk}$ and $CTV_{int10mm}$ with respect to the baseline plan.

D Variance optimization

We derived Equation (5.7) from Equation (5.6) by applying the following relation:

$$\mathbb{E}\left[D\mathbf{x}\mathbf{x}^T D^T\right] = \frac{1}{|K|}\sum_{k\in K} D_k \mathbf{x}\mathbf{x}^T D_k^T.$$

Here, we provide verification for this step. We defined the mean of multiple matrices as the matrix of the element wise averages. Therefore, to compute the expected value of several matrices, we might as well compute the expected value of each entry of the matrix. Hence, it suffices to show what the expected value of entry iz of $D\mathbf{x}\mathbf{x}^T D^T$ is. We have

$$(D\mathbf{x}\mathbf{x}^T D^T)_{iz} = \sum_j \sum_p D_{ij} x_j x_p D_{zp}.$$

The expected value of $(D\mathbf{x}\mathbf{x}^T D^T)_{iz}$ is computed as follows:

$$\begin{aligned}
\mathbb{E}\left[(D\mathbf{x}\mathbf{x}^T D^T)_{iz}\right] &= \mathbb{E}\left[\sum_j \sum_p D_{ij} x_j x_p D_{zp}\right], \\
&= \mathbb{E}\left[\sum_j \sum_p x_j x_p D_{ij} D_{zp}\right] \\
&= \sum_j \sum_p \mathbb{E}\left[x_j x_p D_{ij} D_{zp}\right] \\
&= \sum_j \sum_p x_j x_p \mathbb{E}\left[D_{ij} D_{zp}\right] \\
&= \sum_j \sum_p x_j x_p \left(\frac{1}{|K|}\sum_{k\in K} (D_k)_{ij}(D_k)_{zp}\right) \\
&= \sum_j \sum_p \left(\frac{1}{|K|}\sum_{k\in K} (D_k)_{ij} x_j x_p (D_k)_{zp}\right) \\
&= \left(\frac{1}{|K|}\sum_{k\in K} D_k \mathbf{x}\mathbf{x}^T D_k^T\right)_{iz}.
\end{aligned}$$

Hence, we conclude that

$$\mathbb{E}[D\mathbf{x}\mathbf{x}^T D^T] = \frac{1}{|K|}\sum_{k\in K} D_k \mathbf{x}\mathbf{x}^T D_k^T.$$